Lectures in Labor Economics

Daron Acemoglu

David Autor

Contents

Part 1

Introduction to Human Capital Investments

CHAPTER 1

The Basic Theory of Human Capital

1. General Issues

One of the most important ideas in labor economics is to think of the set of marketable skills of workers as a form of capital in which workers make a variety of investments. This perspective is important in understanding both investment incentives, and the structure of wages and earnings.

Loosely speaking, human capital corresponds to any stock of knowledge or characteristics the worker has (either innate or acquired) that contributes to his or her "productivity". This definition is broad, and this has both advantages and disadvantages. The advantages are clear: it enables us to think of not only the years of schooling, but also of a variety of other characteristics as part of human capital investments. These include school quality, training, attitudes towards work, etc. Using this type of reasoning, we can make some progress towards understanding some of the differences in earnings across workers that are not accounted by schooling differences alone.

The disadvantages are also related. At some level, we can push this notion of human capital too far, and think of every difference in remuneration that we observe in the labor market as due to human capital. For example, if I am paid less than another Ph.D., that must be because I have lower "skills" in some other dimension that's not being measured by my years of schooling—this is the famous (or infamous) *unobserved heterogeneity* issue. The presumption that all pay differences are related to skills (even if these skills are unobserved to the economists in the standard data sets) is not a bad place to start when we want to impose a conceptual structure on empirical wage distributions, but there are many notable exceptions, some of which will be discussed later. Here it is useful to mention three:

(1) Compensating differentials: a worker may be paid less in money, because he is receiving part of his compensation in terms of other (hard-to-observe) characteristics of the job, which may include lower effort requirements, more pleasant working conditions, better amenities etc.
(2) Labor market imperfections: two workers with the same human capital may be paid different wages because jobs differ in terms of their productivity and pay, and one of them ended up matching with the high productivity job, while the other has matched with the low productivity one.
(3) Taste-based discrimination: employers may pay a lower wage to a worker because of the worker's gender or race due to their prejudices.

In interpreting wage differences, and therefore in thinking of human capital investments and the incentives for investment, it is important to strike the right balance between assigning earning differences to unobserved heterogeneity, compensating wage differentials and labor market imperfections.

2. Uses of Human Capital

The standard approach in labor economics views human capital as a set of skills/characteristics that increase a worker's productivity. This is a useful starting place, and for most practical purposes quite sufficient. Nevertheless, it may be useful to distinguish between some complementary/alternative ways of thinking of human capital. Here is a possible classification:

(1) The Becker view: human capital is directly useful in the production process. More explicitly, human capital increases a worker's productivity in all tasks, though possibly differentially in different tasks, organizations, and situations. In this view, although the role of human capital in the production process may be quite complex, there is a sense in which we can think of it as represented (representable) by a unidimensional object, such as the stock

of knowledge or skills, h, and this stock is directly part of the production function.

(2) The Gardener view: according to this view, we should not think of human capital as unidimensional, since there are many many dimensions or types of skills. A simple version of this approach would emphasize mental vs. physical abilities as different skills. Let us dub this the Gardener view after the work by the social psychologist Howard Gardener, who contributed to the development of multiple-intelligences theory, in particular emphasizing how many geniuses/famous personalities were very "unskilled" in some other dimensions.

(3) The Schultz/Nelson-Phelps view: human capital is viewed mostly as the capacity to adapt. According to this approach, human capital is especially useful in dealing with "disequilibrium" situations, or more generally, with situations in which there is a changing environment, and workers have to adapt to this.

(4) The Bowles-Gintis view: "human capital" is the capacity to work in organizations, obey orders, in short, adapt to life in a hierarchical/capitalist society. According to this view, the main role of schools is to instill in individuals the "correct" ideology and approach towards life.

(5) The Spence view: observable measures of human capital are more a signal of ability than characteristics independently useful in the production process.

Despite their differences, the first three views are quite similar, in that "human capital" will be valued in the market because it increases firms' profits. This is straightforward in the Becker and Schultz views, but also similar in the Gardener view. In fact, in many applications, labor economists' view of human capital would be a mixture of these three approaches. Even the Bowles-Gintis view has very similar implications. Here, firms would pay higher wages to educated workers because these workers will be more useful to the firm as they will obey orders better and will be more reliable members of the firm's hierarchy. The Spence view is different from the others, however, in that observable measures of human capital may be rewarded because they are signals about some other characteristics of workers. We will discuss different implications of these views below.

3. Sources of Human Capital Differences

It is useful to think of the possible sources of human capital differences before discussing the incentives to invest in human capital:

(1) Innate ability: workers can have different amounts of skills/human capital because of innate differences. Research in biology/social biology has documented that there is some component of IQ which is genetic in origin (there is a heated debate about the exact importance of this component, and some economists have also taken part in this). The relevance of this observation for labor economics is twofold: (i) there is likely to be heterogeneity in human capital even when individuals have access to the same investment opportunities and the same economic constraints; (ii) in empirical applications, we have to find a way of dealing with this source of differences in human capital, especially when it's likely to be correlated with other variables of interest.

(2) Schooling: this has been the focus of much research, since it is the most easily observable component of human capital investments. It has to be borne in mind, however, that the R^2 of earnings regressions that control for schooling is relatively small, suggesting that schooling differences account for a relatively small fraction of the differences in earnings. Therefore, there is much more to human capital than schooling. Nevertheless, the analysis of schooling is likely to be very informative if we presume that the same forces that affect schooling investments are also likely to affect non-schooling investments. So we can infer from the patterns of schooling investments what may be happening to non-schooling investments, which are more difficult to observe.

(3) School quality and non-schooling investments: a pair of identical twins who grew up in the same environment until the age of 6, and then completed the same years of schooling may nevertheless have different amounts of human capital. This could be because they attended different schools with varying qualities, but it could also be the case even if they went to the same school. In this latter case, for one reason or another, they may have chosen to make different investments in other components of their human capital (one may have worked harder, or studied especially for some subjects, or because of a variety of choices/circumstances, one may have become more assertive, better at communicating, etc.). Many economists believe that these "unobserved" skills are very important in understanding the structure of wages (and the changes in the structure of wages). The problem is that we do not have good data on these components of human capital. Nevertheless, we will see different ways of inferring what's happening to these dimensions of human capital below.

(4) Training: this is the component of human capital that workers acquire after schooling, often associated with some set of skills useful for a particular industry, or useful with a particular set of technologies. At some level, training is very similar to schooling in that the worker, at least to some degree, controls how much to invest. But it is also much more complex, since it is difficult for a worker to make training investments by himself. The firm also needs to invest in the training of the workers, and often ends up bearing a large fraction of the costs of these training investments. The role of the firm is even greater once we take into account that training has a significant "matching" component in the sense that it is most useful for the worker to invest in a set of specific technologies that the firm will be using in the future. So training is often a joint investment by firms and workers, complicating the analysis.

(5) Pre-labor market influences: there is increasing recognition among economists that peer group effects to which individuals are exposed before they join the labor market may also affect their human capital significantly. At some level, the analysis of these pre-labor market influences may be "sociological". But it also has an element of investment. For example, an altruistic parent deciding where to live is also deciding whether her offspring will be exposed to good or less good pre-labor market influences. Therefore, some of the same issues that arise in thinking about the theory of schooling and training will apply in this context too.

4. Human Capital Investments and The Separation Theorem

Let us start with the partial equilibrium schooling decisions and establish a simple general result, sometimes referred to as a "separation theorem" for human capital investments. We set up the basic model in continuous time for simplicity.

Consider the schooling decision of a single individual facing exogenously given prices for human capital. Throughout, we assume that there are perfect capital markets. The separation theorem referred to in the title of this section will show that, with perfect capital markets, schooling decisions will maximize the net present discounted value of the individual. More specifically, consider an individual with an instantaneous utility function $u(c)$ that satisfies the standard neoclassical assumptions. In particular, it is strictly increasing and strictly concave. Suppose that the individual has a planning horizon of T (where $T = \infty$ is allowed), discounts the future at the rate $\rho > 0$ and faces a constant flow rate of death equal to $\nu \geq 0$. Standard arguments imply that the objective function of this individual at time $t = 0$ is

$$\max \int_0^T \exp(-(\rho + \nu) t) u(c(t)) dt. \tag{1.1}$$

Suppose that this individual is born with some human capital $h(0) \geq 0$. Suppose also that his human capital evolves over time according to the differential equation

$$\dot{h}(t) = G(t, h(t), s(t)), \tag{1.2}$$

where $s(t) \in [0,1]$ is the fraction of time that the individual spends for investments in schooling, and $G : \mathbb{R}_+^2 \times [0,1] \to \mathbb{R}_+$ determines how human capital evolves as a function of time, the individual's stock of human capital and schooling decisions. In addition, we can impose a further restriction on schooling decisions, for example,

$$s(t) \in \mathcal{S}(t), \tag{1.3}$$

where $\mathcal{S}(t) \subset [0,1]$ and may be useful to model constraints of the form $s(t) \in \{0,1\}$, which would correspond to the restriction that schooling must be full-time (or other such restrictions on human capital investments).

The individual is assumed to face an exogenous sequence of wage per unit of human capital given by $[w(t)]_{t=0}^{T}$, so that his labor earnings at time t are

$$W(t) = w(t)[1 - s(t)][h(t) + \omega(t)],$$

where $1 - s(t)$ is the fraction of time spent supplying labor to the market and $\omega(t)$ is non-human capital labor that the individual may be supplying to the market at time t. The sequence of non-human capital labor that the individual can supply to the market, $[\omega(t)]_{t=0}^{T}$, is exogenous. This formulation assumes that the only margin of choice is between market work and schooling (i.e., there is no leisure).

Finally, let us assume that the individual faces a constant (flow) interest rate equal to r on his savings. Using the equation for labor earnings, the lifetime budget constraint of the individual can be written as

$$\int_0^T \exp(-rt)\, c(t)\, dt \leq \int_0^T \exp(-rt)\, w(t)[1 - s(t)][h(t) + \omega(t)]\, dt. \tag{1.4}$$

The Separation Theorem, which is the subject of this section, can be stated as follows:

Theorem 1.1. ***(Separation Theorem)*** *Suppose that the instantaneous utility function $u(\cdot)$ is strictly increasing. Then the sequence $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to the maximization of (1.1) subject to (1.2), (1.3) and (1.4) if and only if $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ maximizes*

$$\int_0^T \exp(-rt)\, w(t)[1 - s(t)][h(t) + \omega(t)]\, dt \tag{1.5}$$

subject to (1.2) and (1.3), and $[\hat{c}(t)]_{t=0}^{T}$ *maximizes (1.1) subject to (1.4) given* $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$. *That is, human capital accumulation and supply decisions can be* separated *from consumption decisions.*

Proof. To prove the "only if" part, suppose that $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ does not maximize (1.5), but there exists $\hat{c}(t)$ such that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to (1.1). Let the value of (1.5) generated by $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ be denoted Y. Since $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ does not maximize (1.5), there exists $[s(t), h(t)]_{t=0}^{T}$ reaching a value of (1.5), $Y' > Y$. Consider the sequence $[c(t), s(t), h(t)]_{t=0}^{T}$, where $c(t) = \hat{c}(t) + \varepsilon$. By the hypothesis that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to (1.1), the budget constraint (1.4) implies

$$\int_0^T \exp(-rt)\,\hat{c}(t)\,dt \leq Y.$$

Let $\varepsilon > 0$ and consider $c(t) = \hat{c}(t) + \varepsilon$ for all t. We have that

$$\begin{aligned} \int_0^T \exp(-rt)\,c(t)\,dt &= \int_0^T \exp(-rt)\,\hat{c}(t)\,dt + \frac{[1-\exp(-rT)]}{r}\varepsilon. \\ &\leq Y + \frac{[1-\exp(-rT)]}{r}\varepsilon. \end{aligned}$$

Since $Y' > Y$, for ε sufficiently small, $\int_0^T \exp(-rt)\,c(t)\,dt \leq Y'$ and thus $[c(t), s(t), h(t)]_{t=0}^{T}$ is feasible. Since $u(\cdot)$ is strictly increasing, $[c(t), s(t), h(t)]_{t=0}^{T}$ is strictly preferred to $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$, leading to a contradiction and proving the "only if" part.

The proof of the "if" part is similar. Suppose that $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ maximizes (1.5). Let the maximum value be denoted by Y. Consider the maximization of (1.1) subject to the constraint that $\int_0^T \exp(-rt)\,c(t)\,dt \leq Y$. Let $[\hat{c}(t)]_{t=0}^{T}$ be a solution. This implies that if $[c'(t)]_{t=0}^{T}$ is a sequence that is strictly preferred to $[\hat{c}(t)]_{t=0}^{T}$, then $\int_0^T \exp(-rt)\,c'(t)\,dt > Y$. This implies that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ must be a solution to the original problem, because any other $[s(t), h(t)]_{t=0}^{T}$ leads to a value of (1.5) $Y' \leq Y$, and if $[c'(t)]_{t=0}^{T}$ is strictly preferred to $[\hat{c}(t)]_{t=0}^{T}$, then $\int_0^T \exp(-rt)\,c'(t)\,dt > Y \geq Y'$ for any Y' associated with any feasible $[s(t), h(t)]_{t=0}^{T}$. $\square$

The intuition for this theorem is straightforward: in the presence of perfect capital markets, the best human capital accumulation decisions are those that maximize the lifetime budget set of the individual. It can be shown that this theorem does not hold when there are imperfect capital markets. Moreover, this theorem also fails to hold when leisure is an argument of the utility function of the individual. Nevertheless, it is a very useful benchmarkas a starting point of our analysis.

5. Schooling Investments and Returns to Education

We now turn to the simplest model of schooling decisions in partial equilibrium, which will illustrate the main tradeoffs in human capital investments. The model presented here is a version of Mincer's (1974) seminal contribution. This model also enables a simple mapping from the theory of human capital investments to the large empirical literature on returns to schooling.

Let us first assume that $T = \infty$, which will simplify the expressions. The flow rate of death, ν, is positive, so that individuals have finite expected lives. Suppose that (1.2) and (1.3) are such that the individual has to spend an interval S with $s(t) = 1$—i.e., in full-time schooling, and $s(t) = 0$ thereafter. At the end of the schooling interval, the individual will have a schooling level of

$$h(S) = \eta(S),$$

where $\eta(\cdot)$ is an increasing, continuously differentiable and concave function. For $t \in [S, \infty)$, human capital accumulates over time (as the individual works) according to the differential equation

$$\dot{h}(t) = g_h h(t), \tag{1.6}$$

for some $g_h \geq 0$. Suppose also that wages grow exponentially,

$$\dot{w}(t) = g_w w(t), \tag{1.7}$$

with boundary condition $w(0) > 0$.

Suppose that

$$g_w + g_h < r + \nu,$$

so that the net present discounted value of the individual is finite. Now using Theorem 1.1, the optimal schooling decision must be a solution to the following maximization problem

$$\max_S \int_S^\infty \exp\left(-\left(r+\nu\right)t\right) w\left(t\right) h\left(t\right) dt. \tag{1.8}$$

Now using (1.6) and (1.7), this is equivalent to:

$$\max_S \frac{\eta\left(S\right) w\left(0\right) \exp\left(-\left(r+\nu-g_w\right)S\right)}{r+\nu-g_h-g_w}. \tag{1.9}$$

Since $\eta\left(S\right)$ is concave, the objective function in (1.9) is strictly concave. Therefore, the unique solution to this problem is characterized by the first-order condition

$$\frac{\eta'\left(S^*\right)}{\eta\left(S^*\right)} = r+\nu-g_w. \tag{1.10}$$

Equation (1.10) shows that higher interest rates and higher values of ν (corresponding to shorter planning horizons) reduce human capital investments, while higher values of g_w increase the value of human capital and thus encourage further investments.

Integrating both sides of this equation with respect to S, we obtain

$$\ln\eta\left(S^*\right) = \text{constant} + \left(r+\nu-g_w\right)S^*. \tag{1.11}$$

Now note that the wage earnings of the worker of age $\tau \geq S^*$ in the labor market at time t will be given by

$$W\left(S,t\right) = \exp\left(g_w t\right)\exp\left(g_h\left(t-S\right)\right)\eta\left(S\right).$$

Taking logs and using equation (1.11) implies that the earnings of the worker will be given by

$$\ln W\left(S^*,t\right) = \text{ constant} + \left(r+\nu-g_w\right)S^* + g_w t + g_h\left(t-S^*\right),$$

where $t-S$ can be thought of as worker experience (time after schooling). If we make a cross-sectional comparison across workers, the time trend term $g_w t$, will also go into the constant, so that we obtain the canonical Mincer equation where, in the cross section, log wage earnings are proportional to schooling and experience.

Written differently, we have the following cross-sectional equation

$$\ln W_j = \text{ constant} + \gamma_s S_j + \gamma_e \text{experience}, \tag{1.12}$$

where j refers to individual j. Note however that we have not introduced any source of heterogeneity that can generate different levels of schooling across individuals. Nevertheless, equation (1.12) is important, since it is the typical empirical model for the relationship between wages and schooling estimated in labor economics.

The economic insight provided by this equation is quite important; it suggests that the functional form of the Mincerian wage equation is not just a mere coincidence, but has economic content: the opportunity cost of one more year of schooling is foregone earnings. This implies that the benefit has to be commensurate with these foregone earnings, thus should lead to a proportional increase in earnings in the future. In particular, this proportional increase should be at the rate $(r + \nu - g_w)$.

Empirical work using equations of the form (1.12) leads to estimates for γ in the range of 0.06 to 0.10. Equation (1.12) suggests that these returns to schooling are not unreasonable. For example, we can think of the annual interest rate r as approximately 0.10, ν as corresponding to 0.02 that gives an expected life of 50 years, and g_w corresponding to the rate of wage growth holding the human capital level of the individual constant, which should be approximately about 2%. Thus we should expect an estimate of γ around 0.10, which is consistent with the upper range of the empirical estimates.

6. A Simple Two-Period Model of Schooling Investments and Some Evidence

Let us now step back and illustrate these ideas using a two-period model and then use this model to look at some further evidence. In period 1 an individual (parent) works, consumes c, saves s, decides whether to send their offspring to school, $e = 0$ or 1, and then dies at the end of the period. Utility of household i is given as:

$$\ln c_i + \ln \hat{c}_i \tag{1.13}$$

where $\hat{c}$ is the consumption of the offspring. There is heterogeneity among children, so the cost of education, θ_i varies with i. In the second period skilled individuals (those with education) receive a wage w_s and an unskilled worker receives w_u.

First, consider the case in which there are no credit market problems, so parents can borrow on behalf of their children, and when they do so, they pay the same interest rate, r, as the rate they would obtain by saving. Then, the decision problem of the parent with income y_i is to maximize (1.13) with respect to e_i, c_i and $\hat{c}_i$, subject to the budget constraint:

$$c_i + \frac{\hat{c}_i}{1+r} \leq \frac{w_u}{1+r} + e_i \frac{w_s - w_u}{1+r} + y_i - e_i \theta_i$$

Note that e_i does not appear in the objective function, so the education decision will be made simply to maximize the budget set of the consumer. This is the essence of the Separation Theorem, Theorem 1.1 above. In particular, here parents will choose to educate their offspring only if

$$\theta_i \leq \frac{w_s - w_u}{1+r} \tag{1.14}$$

One important feature of this decision rule is that a greater skill premium as captured by $w_s - w_u$ will encourage schooling, while the higher interest rate, r, will discourage schooling (since schooling is a form of investment with upfront costs and delayed benefits).

In practice, this solution may be difficult to achieve for a variety of reasons. First, there is the usual list of informational/contractual problems, creating credit constraints or transaction costs that introduce a wedge between borrowing and lending rates (or even make borrowing impossible for some groups). Second, in many cases, it is the parents who make part of the investment decisions for their children, so the above solution involves parents borrowing to finance both the education expenses and also part of their own current consumption. These loans are then supposed to be paid back by their children. With the above setup, this arrangement works since parents are fully altruistic. However, if there are non-altruistic parents, this will create obvious problems.

Therefore, in many situations credit problems might be important. Now imagine the same setup, but also assume that parents cannot have negative savings, which is a simple and severe form of credit market problems. This modifies the constraint set as follows

$$\begin{aligned} c_i &\leq y_i - e_i\theta_i - s_i \\ s_i &\geq 0 \\ \hat{c}_i &\leq w_u + e_i\left(w_s - w_u\right) + \left(1+r\right)s \end{aligned}$$

First note that for a parent with $y_i - e_i\theta_i > w_s$, the constraint of nonnegative savings is not binding, so the same solution as before will apply. Therefore, credit constraints will only affect parents who needed to borrow to finance their children's education.

To characterize the solution to this problem, let us look at the utilities from investinging and not investing in education of a parent. Also to simplify the discussion let us focus on parents who would not choose positive savings, that is, those parents with $\left(1+r\right)y_i \leq w_u$. The utilities from investing and not investing in education are given, respectively, by $U(e = 1 \mid y_i, \theta_i) = \ln(y_i - \theta_i) + \ln w_s$, and $U(e = 0 \mid y_i, \theta_i) = \ln y_i + \ln w_u$. Comparison of these two expressions implies that parents with

$$\theta_i \leq y_i \frac{w_s - w_u}{w_s}$$

will invest in education. It is then straightforward to verify that:

(1) This condition is more restrictive than (1.14) above, since $\left(1+r\right)y_i \leq w_u < w_s$.

(2) As income increases, there will be more investment in education, which contrasts with the non-credit-constrained case.

One interesting implication of the setup with credit constraints is that the skill premium, $w_s - w_u$, still has a positive effect on human capital investments. However, in more general models with credit constraints, the conclusions may be more nuanced. For example, if $w_s - w_u$ increases because the unskilled wage, w_u, falls, this may reduce the income level of many of the households that are marginal for the education decision, thus discourage investment in education.

7. Evidence on Human Capital Investments and Credit Constraints

This finding, that income only matters for education investments in the presence of credit constraints, motivates investigations of whether there are significant differences in the educational attainment of children from different parental backgrounds as a test of the importance of credit constraints on education decisions. In addition, the empirical relationship between family income and education is interesting in its own right.

A typical regression would be along the lines of

$$\text{schooling=controls} + \alpha \cdot \log \text{parental income}$$

which leads to positive estimates of α, consistent with credit constraints. The problem is that there are at least two alternative explanations for why we may be estimating a positive α:

(1) Children's education may also be a consumption good, so rich parents will "consume" more of this good as well as other goods. If this is the case, the positive relationship between family income and education is not evidence in favor of credit constraints, since the "separation theorem" does not apply when the decision is not a pure investment (enters directly in the utility function). Nevertheless, the implications for labor economics are quite similar: richer parents will invest more in their children's education.

(2) The second issue is more problematic. The distribution of costs and benefits of education differ across families, and are likely to be correlated with income. That is, the parameter θ_i in terms of the model above will be

correlated with y_i, so a regression of schooling on income will, at least in part, capture the direct effect of different costs and benefits of education.

One line of attack to deal with this problem has been to include other characteristics that could proxy for the costs and benefits of education, or attitudes toward education. The interesting finding here is that when parents' education is also included in the regression, the role of income is substantially reduced.

Does this mean that credit market problems are not important for education? Does it mean that parents' income does not have a direct affect on education? Not necessarily. In particular, there are two reasons for why such an interpretation may not be warranted.

(1) First, parents' income may affect the quality rather than the quantity of education. This may be particularly important in the U.S. context where the choice of the neighborhood in which the family lives appears to have a major effect on the quality of schooling. This implies that in the United States high income parents may be "buying" more human capital for their children, not by sending them to school for longer, but by providing them with better schooling.
(2) Parental income is often measured with error, and has a significant transitory component, so parental education may be a much better proxy for permanent income than income observations in these data sets. Therefore, even when income matters for education, all its effect my load on the parental education variable.

Neither problem is easy to deal with, but there are possible avenues. First, we could look at the incomes of children rather than their schooling as the outcome variable. To the extent that income reflects skills (broadly defined), it will incorporate unobserved dimensions of human capital, including school quality. This takes us to the literature on intergenerational mobility. The typical regression here is

$$\text{log child income} = \text{controls} + \alpha \cdot \text{log parental income} \tag{1.15}$$

Regressions of this sort were first investigated by Becker and Tomes. They found relatively small coefficients, typically in the neighborhood of 0.3 (while others, for example Behrman and Taubman estimated coefficients as low as 0.2). This means that if your parents are twice as rich as my parents, you will typically have about 30 to 40 percent higher income than me. With this degree of intergenerational dependence, differences in initial conditions will soon disappear. In fact, your children will be typically about 10 percent (α^2 percent) richer than my children. So this finding implies that we are living in a relatively "egalitarian" society.

To see this more clearly, consider the following simple model:

$$\ln y_t = \mu + \alpha \ln y_{t-1} + \varepsilon_t$$

where y_t is the income of t-th generation, and ε_t is serially independent disturbance term with variance σ_ε^2. Then the long-term variance of log income is:

$$\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\alpha^2} \tag{1.16}$$

Using the estimate of 0.3 for α, equation (1.16) implies that the long-term variance of log income will be approximately 10 percent higher than σ_ε^2, so the long-run income distribution will basically reflect transitory shocks to dynasties' incomes and skills, and not inherited differences.

Returning to the interpretation of α in equation (1.15), also note that a degree of persistence in the neighborhood of 0.3 is not very different from what we might expect to result simply from the inheritance of IQ between parents and children, or from the children's adoption of cultural values favoring education from their parents. As a result, these estimates suggest that there is a relatively small effect of parents income on children's human capital.

This work has been criticized, however, because there are certain simple biases, stacking the cards against finding large estimates of the coefficient α. First, measurement error will bias the coefficient α towards zero. Second, in typical panel data sets, we observe children at an early stage of their life cycles, where differences in earnings may be less than at later stages, again biasing α downward. Third, income mobility may be very nonlinear, with a lot of mobility among middle income families, but very little at the tails. Work by Solon and Zimmerman has dealt with the first two problems. They find that controlling for these issues increases the degree of persistence substantially to about 0.45 or even 0.55. The next figure shows Solon's baseline estimates.

Table 4—OLS and IV Estimates of ρ for Various Single-Year Income Measures in 1967

Income measure	OLS	IV	Sample si
Log earnings	0.386 (0.079)	0.526 (0.135)	322
Log wage	0.294 (0.052)	0.449 (0.095)	316
Log family income	0.483 (0.069)	0.530 (0.123)	313
Log (family income/poverty line)	0.476 (0.060)	0.563 (0.103)	313

Note: Standard-error estimates are in parentheses.

Figure 1.1

A paper by Cooper, Durlauf and Johnson, in turn, finds that there is much more persistence at the top and the bottom of income distribution than at the middle.

That the difference between 0.3 and 0.55 is in fact substantial can be seen by looking at the implications of using $\alpha = 0.55$ in (1.16). Now the long-run income distribution will be substantially more disperse than the transitory shocks. More specifically, we will have $\sigma_y^2 \approx 1.45 \cdot \sigma_\varepsilon^2$.

To deal with the second empirical issue, one needs a source of exogenous variation in incomes to implement an IV strategy. There are no perfect candidates, but some imperfect ones exist. One possibility, pursued in Acemoglu and Pischke (2001), is to exploit changes in the income distribution that have taken place over the past 30 years to get a source of exogenous variation in household income. The basic idea is that the rank of a family in the income distribution is a good proxy for parental human capital, and conditional on that rank, the income gap has widened over the past 20 years. Moreover, this has happened differentially across states. One can exploit this source of variation by estimating regression of the form

$$s_{iqjt} = \delta_q + \delta_j + \delta_t + \beta_q \ln y_{iqjt} + \varepsilon_{iqjt}, \tag{1.17}$$

where q denotes income quartile, j denotes region, and t denotes time. s_{iqjt} is education of individual i in income quartile q region j time t. With no effect of income on education, β_q's should be zero. With credit constraints, we might expect lower quartiles to have positive β's. Acemoglu and Pischke report versions of this equation using data aggregated to income quartile, region and time cells. The estimates of β are typically positive and significant, as shown in the next two tables.

However, the evidence does not indicate that the β's are higher for lower income quartiles, which suggests that there may be more to the relationship between income and education than simple credit constraints. Potential determinants of the relationship between income and education have already been discussed extensively in the literature, but we still do not have a satisfactory understanding of why parental income may affect children's educational outcomes (and to what extent it does so).

8. The Ben-Porath Model

The baseline Ben-Porath model enriches the models we have seen so far by allowing human capital investments and non-trivial labor supply decisions throughout the lifetime of the individual. It also acts as a bridge to models of investment in human capital on-the-job, which we will discuss below.

Let $s(t) \in [0,1]$ for all $t \geq 0$. Together with the Mincer equation (1.12) above, the Ben-Porath model is the basis of much of labor economics. Here it is sufficient to consider a simple version of this model where the human capital accumulation equation, (1.2), takes the form

$$\dot{h}(t) = \phi(s(t)h(t)) - \delta_h h(t), \tag{1.18}$$

Table 4
Fixed effects regressions for the probability of attending college within two years of high school controlling for income quartile region by income quartile cells, 1972–1992[a]

	Ever attending any college				Ever attending four-year college			
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log mean family income	0.218 (0.101)	0.107 (0.044)	0.102 (0.044)	0.146 (0.107)	0.212 (0.065)	0.148 (0.041)	0.142 (0.040)	0.093 (0.108)
Return to college	1.336 (0.491)	—	−0.887 (0.616)	—	0.817 (0.314)	—	−0.994 (0.556)	—
Region effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income quartile effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Income quartile × Region effects	No	No	No	Yes	No	No	Yes	Yes
Income quartile × Year effects	No	No	No	Yes	No	No	Yes	Yes
Region × Year effects	No	No	No	Yes	No	No	No	Yes

[a]Data are cell level means for 4 Census regions, 4 years, and 4 quartiles for the income of the student's family. Number of cells is 64. Dependent variable is the fraction of students enrolled in any college or in a four-year college within two years of high school graduation calculated from the NLS-72, HSB Senior and Sophomore cohorts, and the NELS. Students left high school in 1972, 1980, 1982, and 1992. Return to college is the relative wage of those with exactly 4 years of college to those with a high school degree (for workers with 1–5 years of experience) calculated from the Census for 1970, 1980, and 1990.

FIGURE 1.2

where $\delta_h > 0$ captures "depreciation of human capital," for example because new machines and techniques are being introduced, eroding the existing human capital of the worker. The individual starts with an initial value of human capital $h(0) > 0$. The function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, continuously differentiable and strictly concave. Furthermore, we simplify the analysis by assuming that this function satisfies the Inada-type conditions,

$$\lim_{x \to 0} \phi'(x) = \infty \text{ and } \lim_{x \to h(0)} \phi'(x) = 0.$$

The latter condition makes sure that we do not have to impose additional constraints to ensure $s(t) \in [0,1]$..

Let us also suppose that there is no non-human capital component of labor, so that $\omega(t) = 0$ for all t, that $T = \infty$, and that there is a flow rate of death $\nu > 0$. Finally, we assume that the wage per unit of human capital is constant at w and the interest rate is constant and equal to r. We also normalize $w = 1$ without loss of any generality.

Again using Theorem 1.1, human capital investments can be determined as a solution to the following problem

$$\max \int_0^\infty \exp\left(-\left(r+\nu\right)\right)\left(1-s\left(t\right)\right)h\left(t\right)dt$$

subject to (1.18).

This problem can then be solved by setting up the current-value Hamiltonian, which in this case takes the form

$$\mathcal{H}\left(h,s,\mu\right) = \left(1-s\left(t\right)\right)h\left(t\right) + \mu\left(t\right)\left(\phi\left(s\left(t\right)h\left(t\right)\right) - \delta_h h\left(t\right)\right),$$

where we used $\mathcal{H}$ to denote the Hamiltonian to avoid confusion with human capital. The necessary conditions for an optimal solution to this problem are

$$\begin{aligned}
\mathcal{H}_s\left(h,s,\mu\right) &= -h\left(t\right) + \mu\left(t\right)h\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right) = 0 \\
\mathcal{H}_h\left(h,s,\mu\right) &= \left(1-s\left(t\right)\right) + \mu\left(t\right)\left(s\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right) - \delta_h\right) \\
&= \left(r+\nu\right)\mu\left(t\right) - \dot{\mu}\left(t\right) \\
\lim_{t\to\infty} \exp\left(-\left(r+\nu\right)t\right)\mu\left(t\right)h\left(t\right) &= 0.
\end{aligned}$$

To solve for the optimal path of human capital investments, let us adopt the following transformation of variables:

$$x\left(t\right) \equiv s\left(t\right)h\left(t\right).$$

Instead of $s(t)$ (or $\mu(t)$) and $h(t)$, we will study the dynamics of the optimal path in $x(t)$ and $h(t)$.

The first necessary condition then implies that

$$1 = \mu\left(t\right)\phi'\left(x\left(t\right)\right), \tag{1.19}$$

while the second necessary condition can be expressed as

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - s(t)\,\phi'(x(t)) - \frac{1-s(t)}{\mu(t)}.$$

Substituting for $\mu(t)$ from (1.19), and simplifying, we obtain

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - \phi'(x(t)). \tag{1.20}$$

The steady-state (stationary) solution of this optimal control problem involves $\dot{\mu}(t) = 0$ and $\dot{h}(t) = 0$, and thus implies that

$$x^* = \phi'^{-1}(r + \nu + \delta_h), \tag{1.21}$$

where $\phi'^{-1}(\cdot)$ is the inverse function of $\phi'(\cdot)$ (which exists and is strictly decreasing since $\phi(\cdot)$ is strictly concave). This equation shows that $x^* \equiv s^* h^*$ will be higher when the interest rate is low, when the life expectancy of the individual is high, and when the rate of depreciation of human capital is low.

To determine s^* and h^* separately, we set $\dot{h}(t) = 0$ in the human capital accumulation equation (1.18), which gives

$$\begin{aligned} h^* &= \frac{\phi(x^*)}{\delta_h} \\ &= \frac{\phi\left(\phi'^{-1}(r + \nu + \delta_h)\right)}{\delta_h}. \end{aligned} \tag{1.22}$$

Since $\phi'^{-1}(\cdot)$ is strictly decreasing and $\phi(\cdot)$ is strictly increasing, this equation implies that the steady-state solution for the human capital stock is uniquely determined and is decreasing in r, ν and δ_h.

More interesting than the stationary (steady-state) solution to the optimization problem is the time path of human capital investments in this model. To derive this, differentiate (1.19) with respect to time to obtain

$$\frac{\dot{\mu}(t)}{\mu(t)} = \varepsilon_{\phi'}(x)\,\frac{\dot{x}(t)}{x(t)},$$

where

$$\varepsilon_{\phi'}(x) = -\frac{x\phi''(x)}{\phi'(x)} > 0$$

is the elasticity of the function $\phi'(\cdot)$ and is positive since $\phi'(\cdot)$ is strictly decreasing (thus $\phi''(\cdot) < 0$). Combining this equation with (1.20), we obtain

$$\frac{\dot{x}(t)}{x(t)} = \frac{1}{\varepsilon_{\phi'}(x(t))}\left(r + \nu + \delta_h - \phi'(x(t))\right). \tag{1.23}$$

Figure 1.4 plots (1.18) and (1.23) in the h-x space. The upward-sloping curve corresponds to the locus for $\dot{h}(t) = 0$, while (1.23) can only be zero at x^*, thus the locus for $\dot{x}(t) = 0$ corresponds to the horizontal line in the figure. The arrows of motion are also plotted in this phase diagram and make it clear that the steady-state solution (h^*, x^*) is globally saddle-path stable, with the stable arm coinciding with the horizontal line for $\dot{x}(t) = 0$. Starting with $h(0) \in (0, h^*)$, $s(0)$ jumps to the level necessary to ensure $s(0)h(0) = x^*$. From then on, $h(t)$ increases and $s(t)$ decreases so as to keep $s(t)h(t) = x^*$. Therefore, the pattern of human capital investments implied by the Ben-Porath model is one of high investment at the beginning of an individual's life followed by lower investments later on.

In our simplified version of the Ben-Porath model this all happens smoothly. In the original Ben-Porath model, which involves the use of other inputs in the production of human capital and finite horizons, the constraint for $s(t) \leq 1$ typically binds early on in the life of the individual, and the interval during which $s(t) = 1$ can be interpreted as full-time schooling. After full-time schooling, the individual starts working (i.e., $s(t) < 1$). But even on-the-job, the individual continues to accumulate human capital (i.e., $s(t) > 0$), which can be interpreted as spending time in training programs or allocating some of his time on the job to learning rather than production. Moreover, because the horizon is finite, if the Inada conditions were relaxed, the individual could prefer to stop investing in human capital at some point. As a result, the time path of human capital generated by the standard Ben-Porath model may be hump-shaped, with a possibly declining portion at the end. Instead, the path of human capital (and the earning potential of the individual) in the current model is always increasing as shown in Figure 1.5.

The importance of the Ben-Porath model is twofold. First, it emphasizes that schooling is not the only way in which individuals can invest in human capital and there is a continuity between schooling investments and other investments in human capital. Second, it suggests that in societies where schooling investments are high we may also expect higher levels of on-the-job investments in human capital. Thus there may be systematic mismeasurement of the amount or the quality human capital across societies.

This model also provides us with a useful way of thinking of the lifecycle of the individual, which starts with higher investments in schooling, and then there is a period of "full-time" work (where $s(t)$ is high), but this is still accompanied by investment in human capital and thus increasing earnings. The increase in earnings takes place at a slower rate as the individual ages. There is also some evidence that earnings may start falling at the very end of workers' careers, though this does not happen in the simplified version of the model presented here (how would you modify it to make sure that earnings may fall in equilibrium?).

The available evidence is consistent with the broad patterns suggested by the model. Nevertheless, this evidence comes from cross-sectional age-experience profiles, so it has to be interpreted with some caution (in particular, the decline at the very end of an individual's life cycle that is found in some studies may be due to "selection," as the higher-ability workers retire earlier).

Perhaps more worrisome for this interpretation is the fact that the increase in earnings may reflect not the accumulation of human capital due to investment, but either:

(1) simple age effects; individuals become more productive as they get older. Or
(2) simple experience effects: individuals become more productive as they get more experienced—this is independent of whether they choose to invest or not.

It is difficult to distinguish between the Ben-Porath model and the second explanation. But there is some evidence that could be useful to distinguish between age effects vs. experience effects (automatic or due to investment).

Josh Angrist's paper on Vietnam veterans basically shows that workers who served in the Vietnam War lost the experience premium associated with the years they served in the war. This is shown in the next figure.

Presuming that serving in the war has no productivity effects, this evidence suggests that much of the age-earnings profiles are due to experience not simply due to age. Nevertheless, this evidence is consistent both with direct experience effects on worker productivity, and also a Ben Porath type explanation where workers are purposefully investing in their human capital while working, and experience is proxying for these investments.

9. Selection and Wages—The One-Factor Model

Issues of selection bias arise often in the analysis of education, migration, labor supply, and sectoral choice decisions. This section illustrates the basic issues of selection using a single-index model, where each individual possesses a one-dimensional skill. Richer models, such as the famous Roy model of selection, incorporate multi-dimensional skills. While models with multi-dimensional skills make a range of additional predictions, the major implications of selection for interpreting wage differences across different groups can be derived using the single-index model.

Suppose that individuals are distinguished by an unobserved type, z, which is assumed to be distributed uniformly between 0 and 1. Individuals decide whether to obtain education, which costs c. The wage of an individual of type z when he has no education is

$$w_0(z) = z$$

and when he obtains education, it is

$$w_1(z) = \alpha_0 + \alpha_1 z, \tag{1.24}$$

where $\alpha_0 > 0$ and $\alpha_1 > 1$. α_0 is the main effect of education on earnings, which applies irrespective of ability, whereas α_1 interacts with ability. The assumption that $\alpha_1 > 1$ implies that education is complementary to ability, and will ensure that high-ability individuals are "positively selected" into education.

Individuals make their schooling choices to maximize income. It is straightforward to see that all individuals of type $z \geq z^*$ will obtain education, where

$$z^* \equiv \frac{c - \alpha_0}{\alpha_1 - 1},$$

which, to make the analysis interesting, we assume lies between 0 and 1. Figure 1.7 gives the wage distribution in this economy.

Now let us look at mean wages by education group. By standard arguments, these are

$$\bar{w}_0 = \frac{c - \alpha_0}{2(\alpha_1 - 1)}$$
$$\bar{w}_1 = \alpha_0 + \alpha_1 \frac{\alpha_1 - 1 + c - \alpha_0}{2(\alpha_1 - 1)}$$

It is clear that $\bar{w}_1 - \bar{w}_0 > \alpha_0$, so the wage gap between educated and uneducated groups is greater than the main effect of education in equation (1.24)—since $\alpha_1 - 1 > 0$. This reflects two components. First, the return to education is not α_0, but it is $\alpha_0 + \alpha_1 \cdot z$ for individual z. Therefore, for a group of mean ability $\bar{z}$, the return to education is

$$w_1(\bar{z}) - w_0(\bar{z}) = \alpha_0 + (\alpha_1 - 1)\bar{z},$$

which we can simply think of as the return to education evaluated at the mean ability of the group.

But there is one more component in $\bar{w}_1 - \bar{w}_0$, which results from the fact that the average ability of the two groups is not the same, and the earning differences resulting from this ability gap are being counted as part of the returns to education. In fact, since $\alpha_1 - 1 > 0$, high-ability individuals are selected into education increasing the wage differential. To see this, rewrite the observed wage differential as follows

$$\bar{w}_1 - \bar{w}_0 = \alpha_0 + (\alpha_1 - 1)\left[\frac{c - \alpha_0}{2(\alpha_1 - 1)}\right] + \frac{\alpha_1}{2}$$

Here, the first two terms give the return to education evaluated at the mean ability of the uneducated group. This would be the answer to the counter-factual question of how much the earnings of the uneducated group would increase if they were to obtain education. The third term is the additional effect that results from the fact that the two groups *do not* have the same *ability* level. It is therefore the selection effect. Alternatively, we could have written

$$\bar{w}_1 - \bar{w}_0 = \alpha_0 + (\alpha_1 - 1)\left[\frac{\alpha_1 - 1 + c - \alpha_0}{2(\alpha_1 - 1)}\right] + \frac{1}{2},$$

where now the first two terms give the return to education evaluated at the mean ability of the educated group, which is greater than the return to education evaluated at the mean ability level of the uneducated group. So the selection effect is somewhat smaller, but still positive.

This example illustrates how looking at observed averages, without taking selection into account, may give misleading results, and also provides a simple example of how to think of decisions in the presence of this type of heterogeneity.

It is also interesting to note that if $\alpha_1 < 1$, we would have negative selection into education, and observed returns to education would be less than the true returns. The case of $\alpha_1 < 1$ appears less plausible, but may arise if high ability individuals do not need to obtain education to perform certain tasks.

Table 5
Fixed effects regressions for the probability of attending college within two years of high school effects by income quartile region by income quartile cells, 1972–1992[a]

	Ever attending any college				Ever attending four-year college			
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log mean family income Quartile 1	0.018 (0.143)	0.154 (0.056)	0.139 (0.064)	− 0.039 (0.187)	0.010 (0.085)	0.108 (0.052)	0.064 (0.053)	− 0.016 (0.190)
Log mean family income Quartile 2	0.229 (0.258)	0.189 (0.113)	0.167 (0.117)	0.201 (0.334)	0.151 (0.153)	0.128 (0.105)	0.087 (0.101)	− 0.205 (0.339)
Log mean family income Quartile 3	0.617 (0.273)	0.161 (0.116)	0.148 (0.129)	0.328 (0.283)	0.428 (0.162)	0.174 (0.107)	0.150 (0.112)	− 0.039 (0.287)
Log mean family income Quartile 4	0.405 (0.152)	0.012 (0.071)	− 0.005 (0.072)	0.231 (0.132)	0.392 (0.092)	0.212 (0.066)	0.183 (0.063)	0.147 (0.134)
Return to college Quartile 1	0.691 (1.052)	—	− 1.049 (0.759)	—	− 0.053 (0.623)	—	− 1.577 (0.659)	—
Return to college Quartile 2	1.144 (0.938)	—	− 1.032 (0.726)	—	0.599 (0.556)	—	− 1.121 (0.630)	—
Return to college Quartile 3	0.481 (1.050)	—	− 0.963 (0.722)	—	0.171 (0.622)	—	− 1.115 (0.627)	—
Return to college Quartile 4	1.367 (0.952)	—	− 0.438 (0.723)	—	1.304 (0.564)	—	− 0.226 (0.627)	—
Region effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income quartile effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Income quartile × Region effects	No	No	No	Yes	No	Yes	Yes	Yes
Income quartile × Year effects	No	No	No	Yes	No	Yes	Yes	Yes
Region × Year effects	No	No	No	Yes	No	No	No	Yes

[a]Data are cell level means for 4 Census regions, 4 years, and 4 quartiles for the income of the student's family. Number of cells is 64. Dependent variable is the fraction of students enrolled in any college or in a four-year college within two years of high school graduation calculated from the NLS-72, HSB Senior and Sophomore cohorts, and the NELS. Students left high school in 1972, 1980, 1982, and 1992. Return to college is the relative wage of those with exactly 4 years of college to those with a high school degree (for workers with 1–5 years of experience) calculated from the Census for 1970, 1980, and 1990.

Figure 1.3

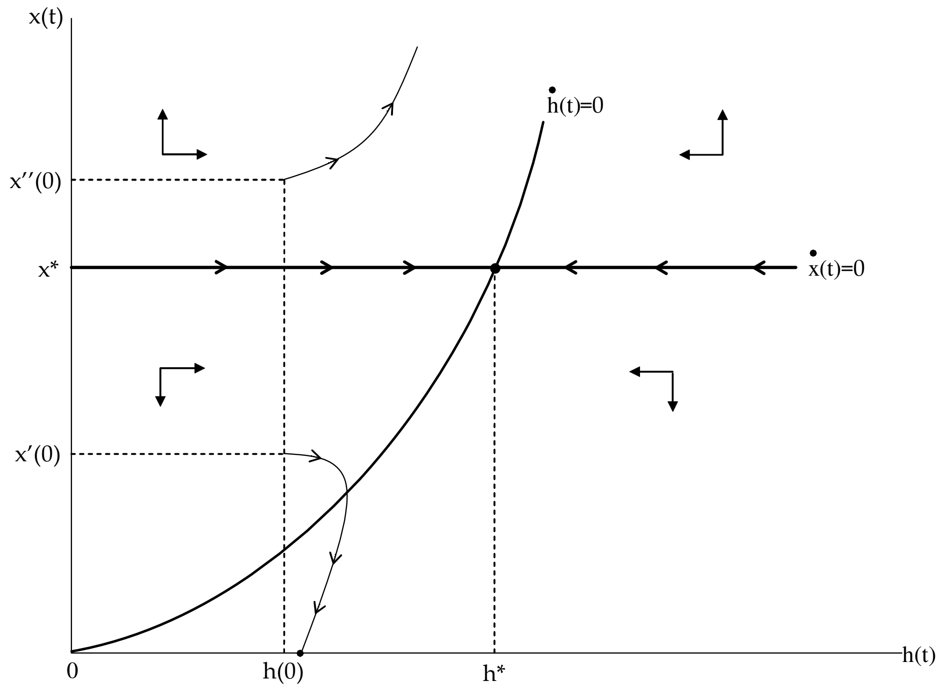


FIGURE 1.4. Steady state and equilibrium dynamics in the simplified Ben Porath model.

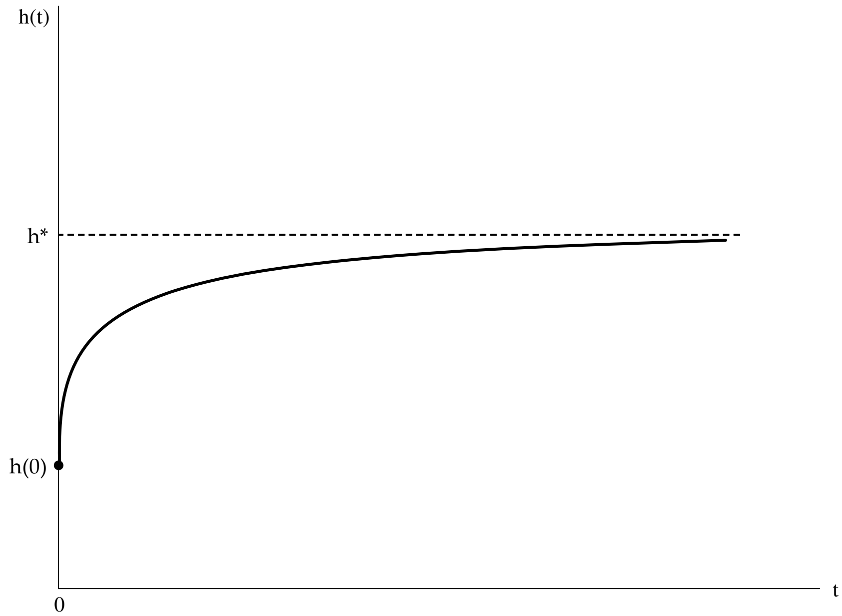
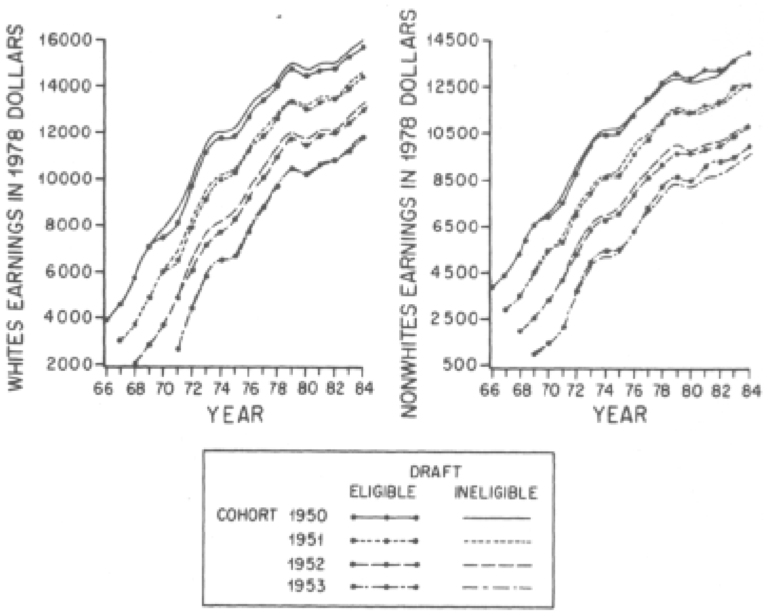


FIGURE 1.5. Time path of human capital investments in the simplified Ben Porath model.



Notes: The figure plots mean W-2 compensation in 1981–84 against probabilities of veteran status by cohort and groups of five consecutive lottery numbers for white men born 1950–53. Plotted points consist of the average residuals (over four years of earnings) from regressions on period and cohort effects. The slope of the least squares regression line drawn through the points is $-2{,}384$ with a standard error of 778, and is an estimate of α in the equation

$$\bar{y}_{ctj} = \beta_c + \delta_t + \hat{p}_{cj}\alpha + \bar{u}_{ctj}.$$

Figure 3. Earnings and the Probability of Veteran Status by Lottery Number

Figure 1.6

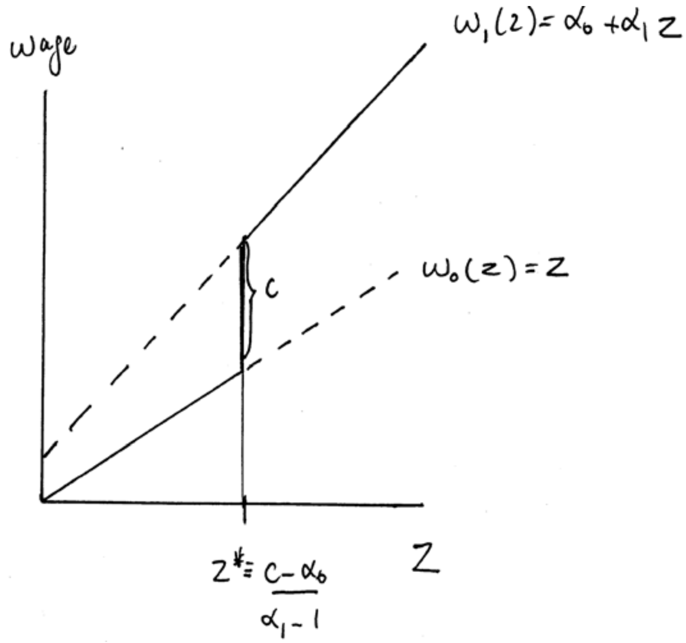


FIGURE 1.7. Selection in the One-Factor Model.

CHAPTER 2

Human Capital and Signaling

1. The Basic Model of Labor Market Signaling

The models we have discussed so far are broadly in the tradition of Becker's approach to human capital. Human capital is viewed as an input in the production process. The leading alternative is to view education purely as a signal. Consider the following simple model to illustrate the issues.

There are two types of workers, high ability and low ability. The fraction of high ability workers in the population is λ. Workers know their own ability, but employers do not observe this directly. High ability workers always produce y_H, while low ability workers produce y_L. In addition, workers can obtain education. The cost of obtaining education is c_H for high ability workers and c_L for low ability workers. The crucial assumption is that $c_L > c_H$, that is, education is more costly for low ability workers. This is often referred to as the "single-crossing" assumption, since it makes sure that in the space of education and wages, the indifference curves of high and low types intersect only once. For future reference, let us denote the decision to obtain education by $e = 1$.

For simplicity, we assume that education does not increase the productivity of either type of worker. Once workers obtain their education, there is competition among a large number of risk-neutral firms, so workers will be paid their expected productivity. More specifically, the timing of events is as follows:

- Each worker finds out their ability.
- Each worker chooses education, $e = 0$ or $e = 1$.
- A large number of firms observe the education decision of each worker (but not their ability) and compete a la Bertrand to hire these workers.

Clearly, this environment corresponds to a dynamic game of incomplete information, since individuals know their ability, but firms do not. In natural equilibrium concept in this case is the Perfect Bayesian Equilibrium. Recall that a Perfect Bayesian Equilibrium consists of a strategy profile σ (designating a strategy for each player) and a brief profile μ (designating the beliefs of each player at each information set) such that σ is *sequentially rational* for each player given μ (so that each player plays the best response in each information set given their beliefs) and μ is derived from σ using Bayes's rule whenever possible. While Perfect Bayesian Equilibria are straightforward to characterize and often reasonable, in incomplete information games where players with private information move before those without this information, there may also exist Perfect Bayesian Equilibria with certain undesirable characteristics. We may therefore wish to strengthen this notion of equilibrium (see below).

In general, there can be two types of equilibria in this game.

(1) Separating, where high and low ability workers choose different levels of schooling, and as a result, in equilibrium, employers can infer worker ability from education (which is a straightforward application of Bayesian updating).
(2) Pooling, where high and low ability workers choose the same level of education.

In addition, there can be semi-separating equilibria, where some education levels are chosen by more than one type.

1.1. A separating equilibrium. Let us start by characterizing a possible separating equilibrium, which illustrates how education can be valued, even though it has no directly productive role.

Suppose that we have

$$y_H - c_H > \ y_L > y_H - c_L \tag{2.1}$$

This is clearly possible since $c_H < c_L$. Then the following is an equilibrium: all high ability workers obtain education, and all low ability workers choose no education. Wages (conditional on education) are:

$$w(e=1) = y_H \text{ and } w(e=0) = y_L$$

Notice that these wages are conditioned on education, and not directly on ability, since ability is not observed by employers. Let us now check that all parties are playing best responses. First consider firms. Given the strategies of workers (to obtain education for high ability and not to obtain education for low ability), a worker with education has productivity y_H while a worker with no education has productivity y_L. So no firm can change its behavior and increase its profits.

What about workers? If a high ability worker deviates to no education, he will obtain $w(e=0) = y_L$, whereas he's currently getting $w(e=1) - c_H = y_H - c_H > y_L$. If a low ability worker deviates to obtaining education, the market will perceive him as a high ability worker, and pay him the higher wage $w(e=1) = y_H$. But from (2.1), we have that $y_H - c_L < y_L$, so this deviation is not profitable for a low ability worker, proving that the separating allocation is indeed an equilibrium.

In this equilibrium, education is valued simply because it is a signal about ability. Education can be a signal about ability because of the single-crossing property. This can be easily verified by considering the case in which $c_L \leq c_H$. Then we could never have condition (2.1) hold, so it would not be possible to convince high ability workers to obtain education, while deterring low ability workers from doing so.

Notice also that if the game was one of perfect information, that is, the worker type were publicly observed, there could never be education investments here. This is an extreme result, due to the assumption that education has no productivity benefits. But it illustrates the forces at work.

1.2. Pooling equilibria in signaling games. However, the separating equilibrium is not the only one. Consider the following allocation: both low and high ability workers do not obtain education, and the wage structure is

$$w(e=1) = (1-\lambda)\, y_L + \lambda y_H \text{ and } w(e=0) = (1-\lambda)\, y_L + \lambda y_H$$

It is straightforward to check that no worker has any incentive to obtain education (given that education is costly, and there are no rewards to obtaining it). Since all workers choose no education, the expected productivity of a worker with no education is $(1-\lambda)\, y_L + \lambda y_H$, so firms are playing best responses. (In Nash Equilibrium and Perfect Bayesian Equilibrium, what they do in response to a deviation by a worker who obtains education is not important, since this does not happen along the equilibrium path).

What is happening here is that the market does not view education as a good signal, so a worker who "deviates" and obtains education is viewed as an average-ability worker, not as a high-ability worker.

What we have just described is a Perfect Bayesian Equilibrium. But is it reasonable? The answer is no. This equilibrium is being supported by the belief that the worker who gets education is no better than a worker who does not. But education is more costly for low ability workers, so they should be less likely to deviate to obtaining education. There are many refinements in game theory which basically try to restrict beliefs in information sets that are not reached along the equilibrium path, ensuring that "unreasonable" beliefs, such as those that think a deviation to obtaining education is more likely from a low ability worker, are ruled out.

Perhaps the simplest is *The Intuitive Criterion* introduced by Cho and Kreps. The underlying idea is as follows. If there exists a type who will never benefit from taking a particular deviation, then the uninformed parties (here the firms) should deduce that this deviation is very unlikely to come from this type. This falls within the category of "forward induction" where rather than solving the game simply backwards, we think about what type of inferences will others derive from a deviation.

To illustrate the main idea, let us simplify the discussion by slightly strengthening condition (2.1) to

$$y_H - c_H > (1-\lambda)\, y_L + \lambda y_H \text{ and } y_L > y_H - c_L. \tag{2.2}$$

Now take the pooling equilibrium above. Consider a deviation to $e = 1$. There is no circumstance under which the low type would benefit from this deviation, since by assumption (2.2) we have $y_L > y_H - c_L$, and the most a worker could ever get is y_H, and the low ability worker is now getting $(1-\lambda)\, y_L + \lambda y_H$. Therefore, firms can deduce that the deviation to $e = 1$ must be coming from the high type, and offer him a wage of y_H. Then (2.2) also ensures that this deviation is profitable for the high types, breaking the pooling equilibrium.

The reason why this refinement is referred to as "The Intuitive Criterion" is that it can be supported by a relatively intuitive "speech" by the deviator along the following lines: "you have to deduce that I must be the high type deviating to $e = 1$, since low types would never ever consider such a deviation, whereas I would find it profitable if I could convince you that I am indeed the high type)." You should bear in mind that this speech is used simply as a loose and intuitive description of the reasoning underlying this equilibrium refinement. In practice there are no such speeches, because the possibility of making such speeches has not been modeled as part of the game. Nevertheless, this heuristic device gives the basic idea.

The overall conclusion is that as long as the separating condition is satisfied, we expect the equilibrium of this economy to involve a separating allocation, where education is valued as a signal.

2. Generalizations

It is straightforward to generalize this equilibrium concept to a situation in which education has a productive role as well as a signaling role. Then the story would be one where education is valued for more than its productive effect, because it is also associated with higher ability.

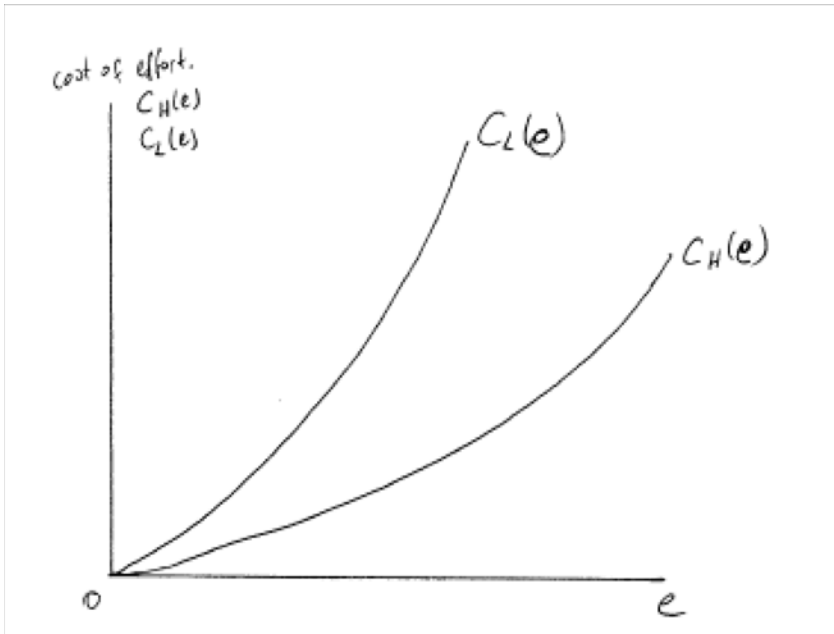


FIGURE 2.1

Let me give the basic idea here. Imagine that education is continuous $e \in [0, \infty)$. And the cost functions for the high and low types are $c_H(e)$ and $c_L(e)$, which are both strictly increasing and convex, with $c_H(0) = c_L(0) = 0$. The single crossing property is that

$$c_H'(e) < c_L'(e) \text{ for all } e \in [0, \infty),$$

that is, the marginal cost of investing in a given unit of education is always higher for the low type. Figure 3.1 shows these cost functions.

Moreover, suppose that the output of the two types as a function of their educations are $y_H(e)$ and $y_L(e)$, with

$$y_H(e) > y_L(e) \text{ for all } e.$$

Figure 2.2 shows the first-best, which would arise in the absence of incomplete information.

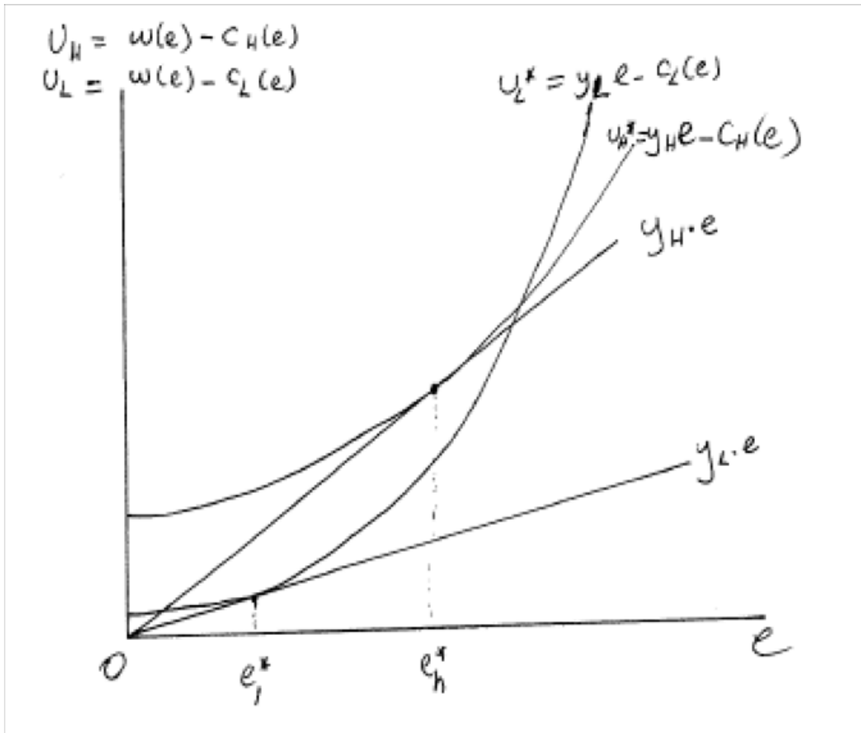


FIGURE 2.2. The first best allocation with complete information.

In particular, as the figure shows, the first best involves effort levels (e_l^*, e_h^*) such that

$$y_L'(e_l^*) = c_L'(e_l^*) \tag{2.3}$$

and

$$y_H'(e_h^*) = c_H'(e_h^*) \,. \tag{2.4}$$

With incomplete information, there are again many equilibria, some separating, some pooling and some semi-separating. But applying a stronger form of the Intuitive Criterion reasoning, we will pick the *Riley equilibrium* of this game, which is a particular separating equilibrium. It is characterized as follows. We first find the most preferred education level for the low type in the perfect information case, which coincides with the first best e_l^* determined in (2.3). Then we can write the incentive compatibility constraint for the low type, such that when the market expects low types to obtain education e_l^*, the low type does not try to mimic the high type; in other words, the low type agent should not prefer to choose the education level the market expects from the high type, e, and receive the wage associated with this level of education. This incentive compatibility constraint is straightforward to write once we note that in the wage level that low type workers will obtain is exactly $y_L(e_l^*)$ in this case, since we are looking at the separating equilibrium. Thus the incentive compatibility constraint is simply

$$y_L(e_l^*) - c_L(e_l^*) \geq w(e) - c_L(e) \text{ for all } e, \tag{2.5}$$

where $w(e)$ is the wage rate paid for a worker with education e. Since e_l^* is the first-best effort level for the low type worker, if we had $w(e) = y_L(e)$, this constraint would always be satisfied. However, since the market can not tell low and high type workers apart, by choosing a different level of education, a low type worker may be able to "mimic" and high type worker and thus we will typically have $w(e) \geq y_L(e)$ when $e \geq e_l^*$, with a strict inequality for some values of education. Therefore, the separating (Riley) equilibrium must satisfy (2.5) for the equilibrium wage function $w(e)$.

To make further progress, note that in a separating equilibrium, there will exist some level of education, say e_h, that will be chosen by high type workers. Then, Bertrand competition among firms, with the reasoning similar to that in the previous section, implies that $w(e_h) = y_H(e_h)$. Therefore, if a low type worker deviates to this level of effort, the market will take him to be a high type worker and pay him the wage $y_H(e_h)$. Now take this education level e_h to be such that the incentive compatibility constraint, (2.5), holds as an equality, that is,

$$y_L(e_l^*) - c_L(e_l^*) = y_H(e_h) - c_L(e_h). \tag{2.6}$$

Then the Riley equilibrium is such that low types choose e_l^* and obtain the wage $w(e_l^*) = y_L(e_l^*)$, and high types choose e_h and obtain the wage $w(e_h) = y_H(e_h)$. That high types are happy to do this follows immediately from the single-crossing

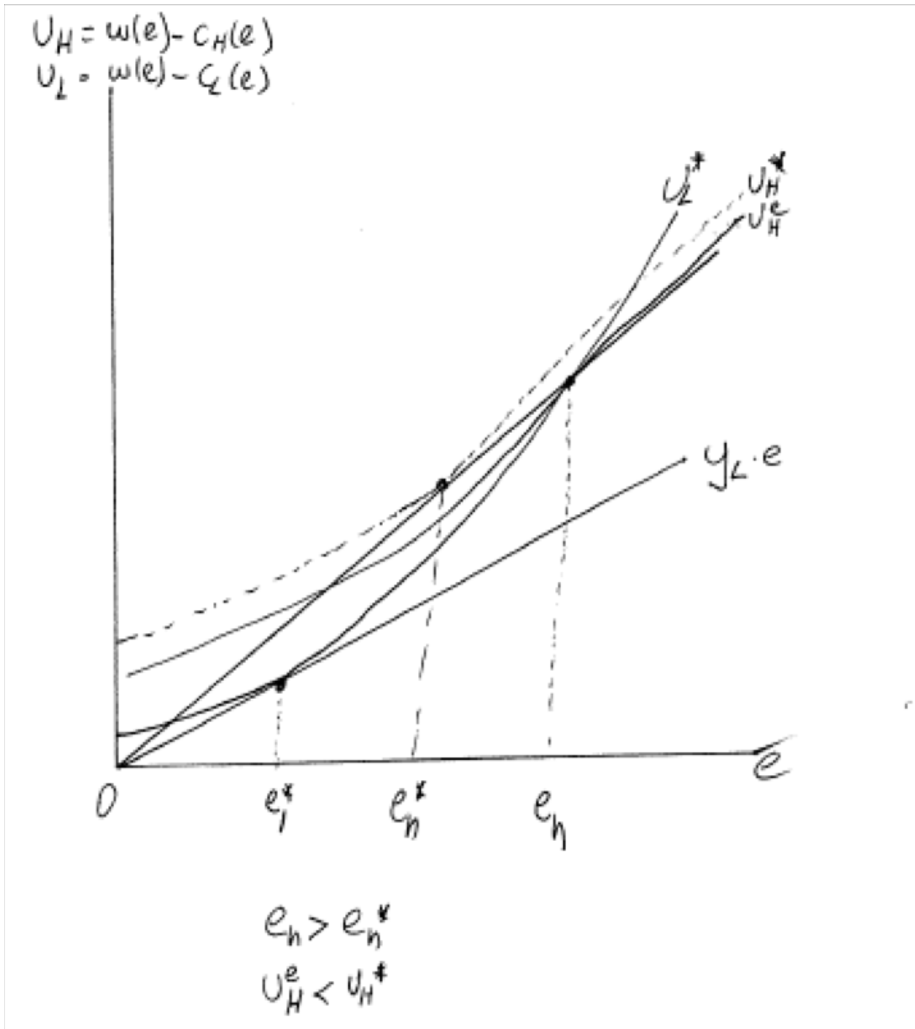


FIGURE 2.3. The Riley equilibrium.

property, since

$$\begin{aligned} y_H(e_h) - c_H(e_h) &= y_H(e_h) - c_L(e_h) - (c_H(e_h) - c_L(e_h)) \\ &> y_H(e_h) - c_L(e_h) - (c_H(e_l^*) - c_L(e_l^*)) \\ &= y_L(e_l^*) - c_L(e_l^*) - (c_H(e_l^*) - c_L(e_l^*)) \\ &= y_L(e_l^*) - c_H(e_l^*), \end{aligned}$$

where the first line is introduced by adding and subtracting $c_L(e_h)$. The second line follows from single crossing, since $c_H(e_h) - c_L(e_h) < c_H(e_l^*) - c_L(e_l^*)$ in view of the fact that $e_l^* < e_h$. The third line exploits (2.6), and the final line simply cancels the two $c_L(e_l^*)$ terms from the right hand side.

Figure 2.3 depicts this equilibrium diagrammatically (for clarity it assumes that $y_H(e)$ and $y_L(e)$ are linear in e).

Notice that in this equilibrium, high type workers invest more than they would have done in the perfect information case, in the sense that e_h characterized here is greater than the education level that high type individuals chosen with perfect information, given by e_h^* in (2.4).

3. Evidence on Labor Market Signaling

Is the signaling role of education important? There are a number of different ways of approaching this question. Unfortunately, direct evidence is difficult to find since ability differences across workers are not only unobserved by firms, but also by econometricians. Nevertheless, number of different strategies can be used to gauge the importance of signaling in the labor market. Here we will discuss a number of different attempts that investigate the importance of labor market signaling. In the next section, we will discuss empirical work that may give a sense of how important signaling considerations are in the aggregate.

Before this discussion, note the parallel between the selection stories discussed above and the signaling story. In both cases, the observed earnings differences between high and low education workers will include a component due to the fact that the abilities of the high and low education groups differ. There is one important difference, however, in that in the selection stories, the market observed ability, it was only us, the economists or the econometricians, who were unable to do so. In the signaling story, the market is also unable to observed ability, and is inferring it from education. For this reason, proper evidence in favor of the signaling story should go beyond documenting the importance of some type of "selection".

There are four different approaches to determining whether signaling is important. The first line of work looks at whether degrees matter, in particular, whether a high school degree or the fourth year of college that gets an individual a university degree matter more than other years of schooling (e.g., Kane and Rouse). This approach suffers from two serious problems. First, the final year of college (or high school) may in fact be more useful than the third-year, especially because it shows that the individual is being able to learn all the required information that makes up a college degree. Second, and more serious, there is no way of distinguishing selection and signaling as possible explanations for these patterns. It may be that those who drop out of high school are observationally different to employers, and hence receive different wages, but these differences are not observed by us in the standard data sets. This is a common problem that will come back again: the implications of unobserved heterogeneity and signaling are often similar.

Second, a creative paper by Lang and Kropp tests for signaling by looking at whether compulsory schooling laws affect schooling above the regulated age. The reasoning is that if the 11th year of schooling is a signal, and the government legislates that everybody has to have 11 years of schooling, now high ability individuals have to get 12 years of schooling to distinguish themselves. They find evidence for this, which they interpret as supportive of the signaling model. The problem is that there are other reasons for why compulsory schooling laws may have such effects. For example, an individual who does not drop out of 11th grade may then decide to complete high school. Alternatively, there can be peer group effects in that as fewer people drop out of school, it may become less socially acceptable the drop out even at later grades.

The third approach is the best. It is pursued in a very creative paper by Tyler, Murnane and Willett. They observe that passing grades in the Graduate Equivalent Degree (GED) differ by state. So an individual with the same grade in the GED exam will get a GED in one state, but not in another. If the score in the exam is an unbiased measure of human capital, and there is no signaling, these two individuals should get the same wages. In contrast, if the GED is a signal, and employers do not know where the individual took the GED exam, these two individuals should get different wages.

Using this methodology, the authors estimate that there is a 10-19 percent return to a GED signal. The attached table shows the results.

An interesting result that Tyler, Murnane and Willett find is that there are no GED returns to minorities. This is also consistent with the signaling view, since it turns out that many minorities prepare for and take the GED exam in prison. Therefore, GED would not only be a positive signal about ability, but also potentially a signal that the individual was at some point incarcerated. This latter feature makes a GED less of that positive signal for minorities.

CHAPTER 3

Externalities and Peer Effects

Many economists believe that human capital not only creates private returns, increasing the earnings of the individual who acquires it, but it also creates externalities, i.e., it increases the productivity of other agents in the economy (e.g., Jacobs, Lucas). If so, existing research on the private returns to education is only part of the picture—the social return, i.e., the private return plus the external return, may far exceed the private return. Conversely, if signaling is important, the private return overestimates the social return to schooling. Estimating the external and the social returns to schooling is a first-order question.

1. Theory

To show how and why external returns to education may arise, we will briefly discuss two models. The first is a theory of non-pecuniary external returns, meaning that external returns arise from technological linkages across agents or firms. The second is pecuniary model of external returns, thus externalities will arise from market interactions and changes in market prices resulting from the average education level of the workers.

1.1. Non-pecuniary human capital externalities. Suppose that the output (or marginal product) of a worker, i, is

$$y_i = Ah_i^{\nu},$$

where h_i is the human capital (schooling) of the worker, and A is aggregate productivity. Assume that labor markets are competitive. So individual earnings are $W_i = Ah_i^{\nu}$.

The key idea of externalities is that the exchange of ideas among workers raises productivity. This can be modeled by allowing A to depend on aggregate human capital. In particular, suppose that

$$A = BH^{\delta} \equiv \mathbb{E}\left[h_i\right]^{\delta}, \tag{3.1}$$

where H is a measure of aggregate human capital, $\mathbb{E}$ is the expectation operator, B is a constant

Individual earnings can then be written as $W_i = Ah_i^{\nu} = BH^{\delta}h_i^{v}$. Therefore, taking logs, we have:

$$\ln W_i = \ln B + \delta \ln H + \nu \ln h_i. \tag{3.2}$$

If external effects are stronger within a geographical area, as seems likely in a world where human interaction and the exchange of ideas are the main forces behind the externalities, then equation (3.2) should be estimated using measures of H at the local level. This is a theory of non-pecuniary externalities, since the external returns arise from the technological nature of equation (3.1).

1.2. Pecuniary human capital externalities. The alternative is pecuniary externalities, as first conjectured by Alfred Marshall in his *Principles of Economics*, increasing the geographic concentration of specialized inputs may increase productivity since the matching between factor inputs and industries is improved. A similar story is developed in Acemoglu (1997), where firms find it profitable to invest in new technologies only when there is a sufficient supply of trained workers to replace employees who quit. We refer to this sort of effect as a pecuniary externality since greater human capital encourages more investment by firms and raises other workers' wages via this channel.

Here, we will briefly explain a simplified version of the model in Acemoglu (1996).

Consider an economy lasting two periods, with production only in the second period, and a continuum of workers normalized to 1. Take human capital, h_i, as given. There is also a continuum of risk-neutral firms. In period 1, firms make an irreversible investment decision, k, at cost Rk. Workers and firms come together in the second period. The labor market is not competitive; instead, firms and workers are matched randomly, and each firm meets a worker. The only decision workers and firms make after matching is whether to produce together or not to produce at all (since there are no further periods). If firm f and worker i produce together, their output is

$$k_f^\alpha h_i^\nu, \tag{3.3}$$

where $\alpha < 1$, $\nu \leq 1 - \alpha$. Since it is costly for the worker-firm pair to separate and find new partners in this economy, employment relationships generate quasi-rents. Wages will therefore be determined by rent-sharing. Here, simply assume that the worker receives a share β of this output as a result of bargaining, while the firm receives the remaining $1 - \beta$ share.

An equilibrium in this economy is a set of schooling choices for workers and a set of physical capital investments for firms. Firm f maximizes the following expected profit function:

$$(1 - \beta) k_f^\alpha \mathbb{E}[h_i^\nu] - R k_f, \tag{3.4}$$

with respect to k_f. Since firms do not know which worker they will be matched with, their expected profit is an average of profits from different skill levels. The function (3.4) is strictly concave, so all firms choose the same level of capital investment, $k_f = k$, given by

$$k = \left(\frac{(1-\beta)\alpha H}{R} \right)^{1/(1-\alpha)}, \tag{3.5}$$

where

$$H \equiv \mathbb{E}[h_i^\nu]$$

is the measure of aggregate human capital. Substituting (3.5) into (3.3), and using the fact that wages are equal to a fraction β of output, the wage income of individual i is given by $W_i = \beta \left((1-\beta)\alpha H \right)^{\alpha/(1-\alpha)} R^{-\alpha/(1-\alpha)} (h_i)^\nu$. Taking logs, this is:

$$\ln W_i = c + \frac{\alpha}{1-\alpha} \ln H + \nu \ln h_i, \tag{3.6}$$

where c is a constant and $\alpha / (1 - \alpha)$ and ν are positive coefficients.

Human capital externalities arise here because firms choose their physical capital in anticipation of the average human capital of the workers they will employ in the future. Since physical and human capital are complements in this setup, a more educated labor force encourages greater investment in physical capital and to higher wages. In the absence of the need for search and matching, firms would immediately hire workers with skills appropriate to their investments, and there would be no human capital externalities.

Nonpecuniary and pecuniary theories of human capital externalities lead to similar empirical relationships since equation (3.6) is identical to equation (3.2), with $c = \ln B$ and $\delta = \alpha / (1 - \alpha)$. Again presuming that these interactions exist in local labor markets, we can estimate a version of (3.2) using differences in schooling across labor markets (cities, states, or even countries).

1.3. Signaling and negative externalities. The above models focused on positive externalities to education. However, in a world where education plays a signaling role, we might also expect significant *negative externalities*. To see this, consider the most extreme world in which education is only a signal—it does not have any productive role.

Contrast two situations: in the first, all individuals have 12 years of schooling and in the second all individuals have 16 years of schooling. Since education has no productive role, and all individuals have the same level of schooling, in both allocations they will earn exactly the same wage (equal to average productivity). Therefore, here the increase in aggregate schooling does not translate into aggregate increases in wages. But in the same world, if one individual obtains more education than the rest, there will be a private return to him, because he would signal that he is of higher ability. Therefore, in a world where signaling is important, we might also want to estimate an equation of the form (3.2), but when signaling issues are important, we would expect δ to be negative.

The basic idea here is that in this world, what determines an individual's wages is his "ranking" in the signaling distribution. When others invest more in their education, a given individual's rank in the distribution declines, hence others are creating a negative externality on this individual via their human capital investment.

2. Evidence

Ordinary Least Squares (OLS) estimation of equations like (3.2) using city or state-level data yield very significant and positive estimates of δ, indicating substantial positive human capital externalities. The leading example is the paper by Jim Rauch.

There are at least two problems with this type OLS estimates. First, it may be precisely high-wage cities or states that either attract a large number of high education workers or give strong support to education. Rauch's estimates were using a cross-section of cities. Including city or state fixed affects ameliorates this problem, but does not solve it, since states' attitudes towards education and the demand for labor may comove. The ideal approach would be to find a source of quasi-exogenous variation in average schooling across labor markets (variation unlikely to be correlated with other sources of variation in the demand for labor in the state).

Acemoglu and Angrist try to accomplish this using differences in compulsory schooling laws. The advantage is that these laws not only affect individual schooling but average schooling in a given area.

There is an additional econometric problem in estimating externalities, which remains even if we have an instrument for average schooling in the aggregate. This is that if individual schooling is measured with error (or for some other reason OLS returns to individual schooling are not the causal effect), some of this discrepancy between the OLS returns and the causal return may load on average schooling, even when average schooling is instrumented. This suggests that we may need to instrument for individual schooling as well (so as to get to the correct return to individual schooling).

More explicitly, let Y_{ijt} be the log weekly wage, than the estimating equation is

$$Y_{ijt} = X_i'\mu + \delta_j + \delta_t + \gamma_1 \overline{S}_{jt} + \gamma_{2i} s_i + u_{jt} + \varepsilon_i, \tag{3.7}$$

To illustrate the main issues, ignore time dependence, and consider the population regression of Y_i on s_i:

$$Y_{ij} = \mu_0 + \rho_0 s_i + \varepsilon_{0i}; \text{ where } \mathbb{E}[\varepsilon_{0i} s_i] \equiv 0. \tag{3.8}$$

Next consider the IV population regression using a full set of state dummies. This is equivalent to

$$Y_{ij} = \mu_1 + \rho_1 \overline{S}_j + \varepsilon_{1i}; \text{ where } \mathbb{E}[\varepsilon_{1i} \overline{S}_j] \equiv 0, \tag{3.9}$$

since the projection of individual schooling on a set of state dummies is simply average schooling in each state.

Now consider the estimation of the empirical analogue of equation (3.2):

$$Y_{ij} = \mu^* + \pi_0 s_i + \pi_1 \overline{S}_j + \xi_i; \text{ where } \mathbb{E}[\xi_i s_i] = \mathbb{E}[\xi_i \overline{S}_j] \equiv 0. \tag{3.10}$$

Then, we have

$$\begin{aligned} \pi_0 &= \rho_1 + \phi(\rho_0 - \rho_1) \\ \pi_1 &= \phi(\rho_1 - \rho_0) \end{aligned} \tag{3.11}$$

where $\phi = 1/1 - R^2 > 1$, and R^2 is the first-stage R-squared for the 2SLS estimates in (3.9). Therefore, when $\rho_1 > \rho_0$, for example because there is measurement error in individual schooling, we may find positive external returns even when there are none.

If we could instrument for both individual and average schooling, we would solve this problem. But what type of instrument?

Consider the relationship of interest:

$$Y_{ij} = \mu + \gamma_1 \overline{S}_j + \gamma_{2i} s_i + u_j + \varepsilon_i, \tag{3.12}$$

which could be estimated by OLS or instrumental variables, to obtain an estimate of γ_1 as well as an average estimate of γ_{2i}, say γ_2^*.

An alternative way of expressing this relationship is to adjust for the effect of individual schooling by directly rewriting (3.12):

$$
\begin{aligned}
Y_{ij} - \gamma_2^* s_i &\equiv \widetilde{Y}_{ij} \\
&= \mu + \gamma_1 \overline{S}_j + [u_j + \varepsilon_i + (\gamma_{2i} - \gamma_2^*) s_i].
\end{aligned} \tag{3.13}
$$

In this case, instrumental variables estimate of external returns is equivalent to the Wald formula

$$
\begin{aligned}
\gamma_1^{IV} &= \frac{\mathbb{E}[\widetilde{Y}_{ij}|z_i = 1] - \mathbb{E}[\widetilde{Y}_{ij}|z_i = 0]}{\mathbb{E}[\overline{S}_j|z_i = 1] - \mathbb{E}[\overline{S}_j|z_i = 0]} \\
&= \gamma_1 + \left[\frac{\mathbb{E}[\gamma_{2i} s_i|z_i = 1] - \mathbb{E}[\gamma_{2i} s_i|z_i = 0]}{\mathbb{E}[s_i|z_i = 1] - \mathbb{E}[s_i|z_i = 0]} - \gamma_2^* \right] \cdot \left[\frac{\mathbb{E}[s_i|z_i = 1] - \mathbb{E}[s_i|z_i = 0]}{\mathbb{E}[\overline{S}_j|z_i = 1] - \mathbb{E}[\overline{S}_j|z_i = 0]} \right].
\end{aligned}
$$

This shows that we should set

$$
\begin{aligned}
\gamma_2^* &= \frac{\mathbb{E}[\gamma_{2i} s_i|z_i = 1] - \mathbb{E}[\gamma_{2i} s_i|z_i = 0]}{\mathbb{E}[s_i|z_i = 1] - \mathbb{E}[s_i|z_i = 0]} \\
&= \frac{\mathbb{E}[(Y_{ij} - \gamma_1 \overline{S}_j)|z_i = 1] - \mathbb{E}[(Y_{ij} - \gamma_1 \overline{S}_j)|z_i = 0]}{\mathbb{E}[s_i|z_i = 1] - \mathbb{E}[s_i|z_i = 0]}
\end{aligned} \tag{3.14}
$$

This is typically not the OLS estimator of the private return, and we should be using some instrument to simultaneously estimate the private return to schooling. The ideal instrument would be one affecting exactly the same people as the compulsory schooling laws.

Quarter of birth instruments might come close to this. Since quarter of birth instruments are likely to affect the same people as compulsory schooling laws, adjusting with the quarter of birth estimate, or using quarter of birth dummies as instrument for individual schooling, is the right strategy.

So the strategy is to estimate an equation similar to (3.2) or (3.10) using compulsory schooling laws for average schooling and quarter of birth dummies for individual schooling.

The estimation results from using this strategy in Acemoglu and Angrist (2000) suggest that there are no significant external returns. The estimates are typically around 1 or 2 percent, and statistically not different from zero. They also suggest that in the aggregate signaling considerations are unlikely to be very important (at the very least, they do not dominate positive externalities).

3. School Quality

Differences in school quality could be a crucial factor in differences in human capital. Two individuals with the same years of schooling might have very different skills and very different earnings because one went to a much better school, with better teachers, instruction and resources. Differences in school quality would add to the unobserved component of human capital.

A natural conjecture is that school quality as measured by teacher-pupil ratios, spending per-pupil, length of school year, and educational qualifications of teachers would be a major determinant of human capital. If school quality matters indeed a lot, an effective way of increasing human capital might be to increase the quality of instruction in schools.

This view was however challenged by a number of economists, most notably, Hanushek. Hanushek noted that the substantial increase in spending per student and teacher-pupil ratios, as well as the increase in the qualifications of teachers, was not associated with improved student outcomes, but on the contrary with a deterioration in many measures of high school students' performance. In addition, Hanushek conducted a meta-analysis of the large number of papers in the education literature, and concluded that there was no overwhelming case for a strong effect of resources and class size on student outcomes.

Although this research has received substantial attention, a number of careful papers show that exogenous variation in class size and other resources are in fact associated with sizable improvements in student outcomes.

Most notable:

(1) Krueger analyzes the data from the Tennessee Star experiment where students were randomly allocated to classes of different sizes.

(2) Angrist and Lavy analyze the effect of class size on test scores using a unique characteristic of Israeli schools which caps class size at 40, thus creating a natural regression discontinuity as a function of the total number of students in the school.

(3) Card and Krueger look at the effects of pupil-teacher ratio, term length and relative teacher wage by comparing the earnings of individuals working in the same state but educated in different states with different school resources.

(4) Another paper by Card and Krueger looks at the effect of the "exogenously" forced narrowing of the resource gap between black and white schools in South Carolina on the gap between black and white pupils' education and subsequent earnings.

All of these papers find sizable effects of school quality on student outcomes. Moreover, a recent paper by Krueger shows that there were many questionable decisions in the meta-analysis by Hanushek, shedding doubt on the usefulness of this analysis. On the basis of these various pieces of evidence, it is safe to conclude that school quality appears to matter for human capital.

4. Peer Group Effects

Issues of school quality are also intimately linked to those of externalities. An important type of externality, different from the external returns to education discussed above, arises in the context of education is peer group effects, or generally social effects in the process of education. The fact that children growing up in different areas may choose different role models will lead to this type of externalities/peer group effects. More simply, to the extent that schooling and learning are group activities, there could be this type of peer group effects.

There are a number of theoretical issues that need to be clarified, as well as important work that needs to be done in understanding where peer group effects are coming from. Moreover, empirical investigation of peer group effects is at its infancy, and there are very difficult issues involved in estimation and interpretation. Since there is little research in understanding the nature of peer group effects, here we will simply take peer group effects as given, and briefly discuss some of its efficiency implications, especially for community structure and school quality, and then very briefly mention some work on estimating peer group affects.

4.1. Implications of peer group effects for mixing and segregation. An important question is whether the presence of peer group effects has any particular implications for the organization of schools, and in particular, whether children who provide positive externalities on other children should be put together in a separate school or classroom.

The basic issue here is equivalent to an assignment problem. The general principle in assignment problems, such as Becker's famous model of marriage, is that if inputs from the two parties are complementary, there should be assortative matching, that is the highest quality individuals should be matched together. In the context of schooling, this implies that children with better characteristics, who are likely to create more positive externalities and be better role models, should be segregated in their own schools, and children with worse characteristics, who will tend to create negative externalities will, should go to separate schools. This practically means segregation along income lines, since often children with "better characteristics" are those from better parental backgrounds, while children with worse characteristics are often from lower socioeconomic backgrounds

So much is well-known and well understood. The problem is that there is an important confusion in the literature, which involves deducing complementarity from the fact that in equilibrium we do observe segregation (e.g., rich parents sending their children to private schools with other children from rich parents, or living in suburbs and sending their children to suburban schools, while poor parents live in ghettos and children from disadvantaged backgrounds go to school with other disadvantaged children in inner cities). This reasoning is often used in discussions of Tiebout competition, together with the argument that allowing parents with different characteristics/tastes to sort into different neighborhoods will often be efficient.

The underlying idea can be given by the following simple model. Suppose that schools consist of two kids, and denote the parental background (e.g., home education or parental expenditure on non-school inputs) of kids by e, and the resulting human capitals by h. Suppose that we have

$$\begin{aligned} h_1 &= e_1^{\alpha} e_2^{1-\alpha} \\ h_2 &= e_1^{1-\alpha} e_2^{\alpha} \end{aligned} \tag{3.15}$$

where $\alpha > 1/2$. This implies that parental backgrounds are complementary, and each kid's human capital will depend mostly on his own parent's background, but also on that of the other kid in the school. For example, it may be easier to learn or be motivated when other children in the class are also motivated. This explains why we have $\partial h_1/\partial e_2 > 0$ and $\partial h_2/\partial e_1 > 0$. But an equally important feature of (3.15) is that $\partial^2 h_1/\partial e_2 \partial e_1 > 0$ and $\partial^2 h_2/\partial e_1 \partial e_2 > 0$, that is, the backgrounds of the two kids are complementary. This implies that a classmate with a good background is especially useful to another kid with a good background. We can think of this as the "bad apple" theory of classroom: one bad kid in the classroom brings down everybody.

As a digression, notice an important feature of the way we wrote (3.15) linking the outcome variables, h_1 and h_2, to *predetermined characteristics* of children e_1 and e_2, which creates a direct analogy with the human capital externalities discussed above. However, this may simply be the reduced form of that somewhat different model, for example,

$$\begin{aligned} h_1 &= H_1 (e_1, h_2) \\ h_2 &= H_2 (e_2, h_1) \end{aligned} \tag{3.16}$$

whereby each individual's human capital depends on his own background and the human capital choice of the other individual. Although in reduced form (3.15) and

(3.16) are very similar, they provide different interpretations of peer group effects, and econometrically they pose different challenges, which we will discuss below.

The complementarity has two implications:

(1) It is socially efficient, in the sense of maximizing the sum of human capitals, to have parents with good backgrounds to send their children to school with other parents with good backgrounds. This follows simply from the definition of complementarity, positive cross-partial derivative, which is clearly verified by the production functions in (3.15).

(2) It will also be an equilibrium outcome that parents will do so. To see this, suppose that we have a situation in which there are two sets of parents with background e_l and $e_h > e_l$. Suppose that there is mixing. Now the marginal willingness to pay of a parent with the high background to be in the same school with the child of another high-background parent, rather than a low-background student, is

$$e_h - e_h^{\alpha} e_l^{1-\alpha},$$

while the marginal willingness to pay of a low background parent to stay in the school with the high background parents is

$$e_l^{\alpha} e_h^{1-\alpha} - e_l.$$

The complementarity between e_h and e_l in (3.15) implies that $e_h - e_h^{\alpha} e_l^{1-\alpha} > e_l^{\alpha} e_h^{1-\alpha} - e_l$.

Therefore, the high-background parent can always outbid the low-background parent for the privilege of sending his children to school with other high-background parents. Thus with profit maximizing schools, segregation will arise as the outcome.

Next consider a production function with substitutability (negative cross-partial derivative). For example,

$$\begin{aligned} h_1 &= \phi e_1 + e_2 - \lambda e_1^{1/2} e_2^{1/2} \\ h_2 &= e_1 + \phi e_2 - \lambda e_1^{1/2} e_2^{1/2} \end{aligned} \tag{3.17}$$

where $\phi > 1$ and $\lambda > 0$ but small, so that human capital is increasing in parental background. With this production function, we again have $\partial h_1/\partial e_2 > 0$ and $\partial h_2/\partial e_1 > 0$, but now in contrast to (3.15), we now have

$$\frac{\partial^2 h_1}{\partial e_2 \partial e_1} \text{ and } \frac{\partial^2 h_2}{\partial e_1 \partial e_2} < 0.$$

This can be thought as corresponding to the "good apple" theory of the classroom, where the kids with the best characteristics and attitudes bring the rest of the class up.

In this case, because the cross-partial derivative is negative, the marginal willingness to pay of low-background parents to have their kid together with high-background parents is higher than that of high-background parents. With perfect markets, we will observe mixing, and in equilibrium schools will consist of a mixture of children from high- and low-background parents.

Now combining the outcomes of these two models, many people jump to the conclusion that since we do observe segregation of schooling in practice, parental backgrounds must be complementary, so segregation is in fact efficient. Again the conclusion is that allowing Tiebout competition and parental sorting will most likely achieve efficient outcomes.

However, this conclusion is not correct, since even if the correct production function was (3.17), segregation would arise in the presence of credit market problems. In particular, the way that mixing is supposed to occur with (3.17) is that low-background parents make a payment to high-background parents so that the latter send their children to a mixed school. To see why such payments are necessary, recall that even with (3.17) we have that the first derivatives are positive, that is

$$\frac{\partial h_1}{\partial e_2} > 0 \text{ and } \frac{\partial h_2}{\partial e_1} > 0.$$

This means that everything else being equal all children benefit from being in the same class with other children with good backgrounds. With (3.17), however, children from better backgrounds benefit *less* than children from less good backgrounds.

This implies that there has to be payments from parents of less good backgrounds to high-background parents.

Such payments are both difficult to implement in practice, and practically impossible taking into account the credit market problems facing parents from poor socioeconomic status.

This implies that, if the true production function is (3.17) but there are credit market problems, we will observe segregation in equilibrium, and the segregation will be inefficient. Therefore we cannot simply appeal to Tiebout competition, or deduce efficiency from the equilibrium patterns of sorting.

Another implication of this analysis is that in the absence of credit market problems (and with complete markets), cross-partials determine the allocation of students to schools. With credit market problems, first there of it has become important. This is a general result, with a range of implications for empirical work.

4.2. The Benabou model. A similar point is developed by Benabou even in the absence of credit market problems, but relying on other missing markets. His model has competitive labor markets, and local externalities (externalities in schooling in the local area). All agents are assumed to be ex ante homogeneous, and will ultimately end up either low skill or high skill.

Utility of agent i is assumed to be

$$U^i = w^i - c^i - r^i$$

where w is the wage, c is the cost of education, which is necessary to become both low skill or high skill, and r is rent.

The cost of education is assumed to depend on the fraction of the agents in the neighborhood, denoted by x, who become high skill. In particular, we have $c_H(x)$ and $c_L(x)$ as the costs of becoming high skill and low skill. Both costs are decreasing in x, meaning that when there are more individuals acquiring high skill, becoming high skill is cheaper (positive peer group effects). In addition, we have

$$c_H(x) > c_L(x)$$

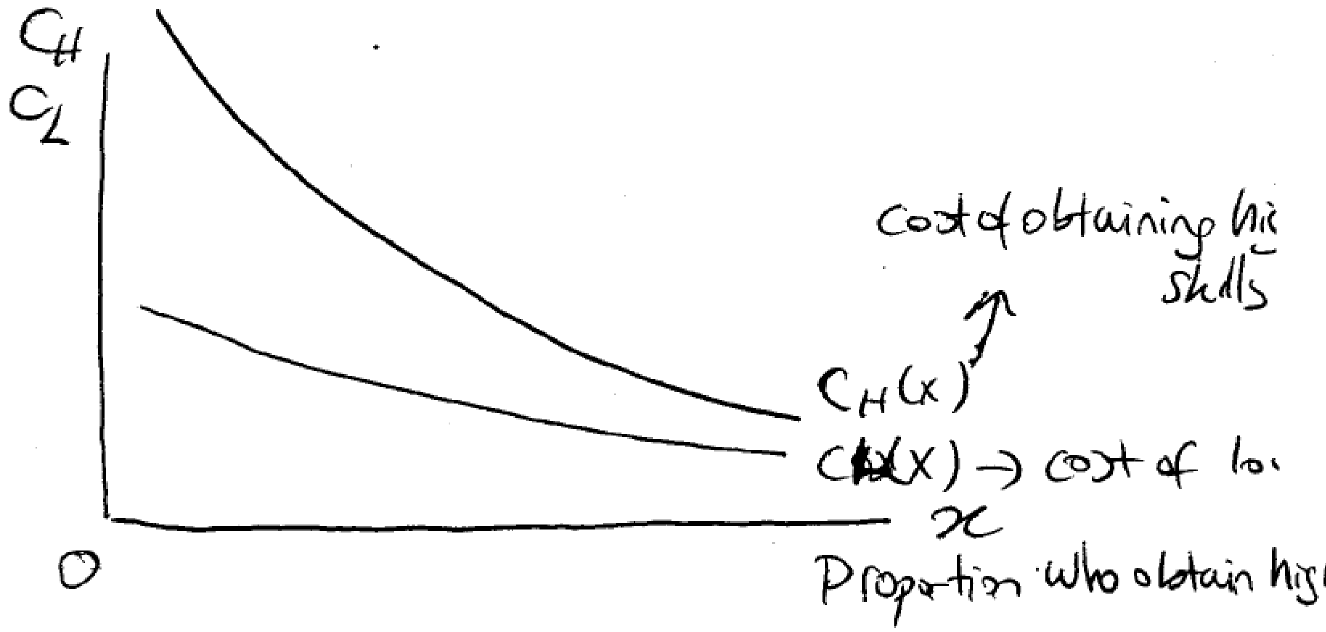


FIGURE 3.1

so that becoming high skill is always more expensive, and as shown in Figure 3.1

$$c'_H(x) < c'_L(x),$$

so that the effect of increase in the fraction of high skill individuals in the neighborhood is bigger on the cost of becoming high skill.

Since all agents are ex ante identical, in equilibrium we must have

$$U(L) = U(H)$$

that is, the utility of becoming high skill and low skill must be the same.

Assume that the labor market in the economy is global, and takes the constant returns to scale form $F(H, L)$. The important implication here is that irrespective of where the worker obtains his education, he will receive the same wage as a function of his skill level.

Also assume that there are two neighborhoods of fixed size, and individuals will compete in the housing market to locate in one neighborhood or the other.

As shown in Figures 3.2 and 3.3, there can be two types of equilibria:

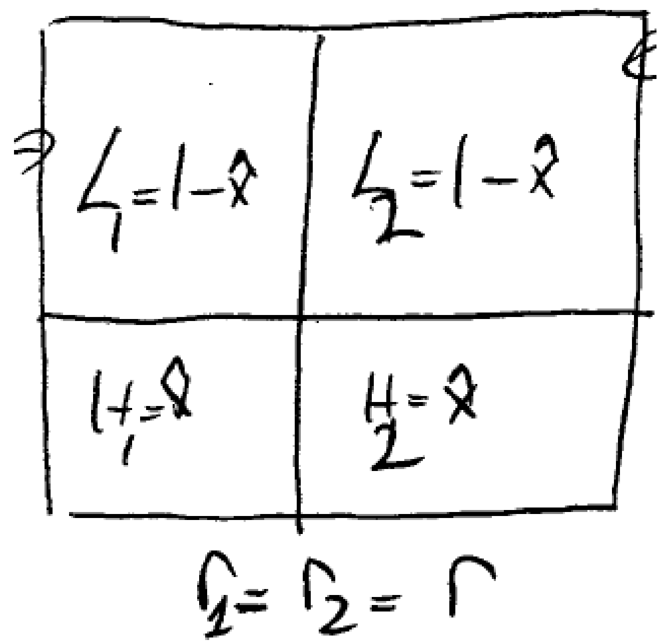


FIGURE 3.2. Integrated City Equilibrium

(1) Integrated city equilibrium, where in both neighborhoods there is a fraction $\hat{x}$ of individual obtaining high education.
(2) Segregated city equilibrium, where one of the neighborhoods is homogeneous. For example, we could have a situation where one neighborhood has $x = 1$ and the other has $\tilde{x} < 1$, or one neighborhood has $x = 0$ and the other has $\bar{x} > 0$.

The important observation here is that only segregated city equilibria are "stable". To see this consider an integrated city equilibrium, and imagine relocating a fraction ε of the high-skill individuals (that is individuals getting high skills) from neighborhood 1 to neighborhood 2. This will reduce the cost of education in neighborhood 2, both for high and low skill individuals. But by assumption, it reduces it more for high skill individuals, so all high skill individuals now will pay higher rents to be in that city, and they will outbid low-skill individuals, taking the economy toward the segregated city equilibrium.

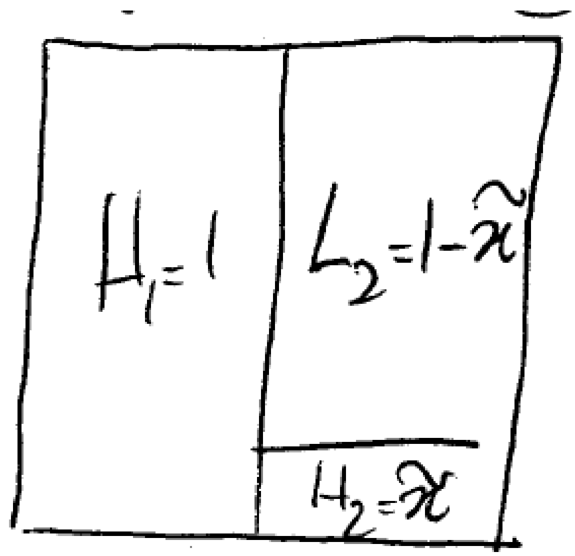


FIGURE 3.3. Segregated City Equilibrium

In contrast, the segregated city equilibrium is always stable. So we again have a situation in which segregation arises as the equilibrium outcome, and this is again because of a reasoning relying on the notion of "complementarity". As in the previous section, high-skill individuals can outbid the low-skill individuals because they benefit more from the peer group effects of high skill individuals.

But crucially there are again *missing markets* in this economy. In particular, rather than paying high skill individuals for the positive externalities that they create, as would be the case in complete markets, agents transact simply through the housing market. In the housing market, there is only one rent level, which both high and low skill individuals pay. In contrast, with complete markets, we can think of the pricing scheme for housing to be such that high skill individuals pay a lower rent (to be compensated for the positive externality that they are creating on the other individuals).

Therefore, there are missing markets, and efficiency is not guaranteed. Is the allocation with segregation efficient?

It turns out that it may or may not. To see this consider the problem of a utilitarian social planner maximizing total output minus costs of education for workers. This implies that the social planner will maximize

$$F(H,L) - H_1 c_H(x_1) - H_2 c_H(x_2) - L_1 c_L(x_1) - L_2 c_L(x_2)$$

where

$$x_1 = \frac{H_1}{L_1 + H_1} \text{ and } x_2 = \frac{H_2}{L_2 + H_2}$$

This problem can be broken into two parts: first, the planner will choose the aggregate amount of skilled individuals, and then she will choose how to actually allocate them between the two neighborhoods. The second part is simply one of cost minimization, and the solution depends on whether

$$\Phi(x) = x c_H(x) + (1-x) c_L(x)$$

is concave or convex. This function is simply the cost of giving high skills to a fraction x of the population. When it is convex, it means that it is best to choose the same level of x in both neighborhoods, and when it is concave, the social planner minimizes costs by choosing two extreme values of x in the two neighborhoods.

It turns out that this function can be convex, i.e. $\Phi''(x) > 0$. More specifically, we have:

$$\Phi''(x) = 2\left(c'_H(x) - c'_L(x)\right) + x\left(c''_H(x) - c''_L(x)\right) + c''_L(x)$$

We can have $\Phi''(x) > 0$ when the second and third terms are large. Intuitively, this can happen because although a high skill individual benefits more from being together with other high skill individuals, he is also creating a positive externality on low skill individuals when he mixes with them. This externality is not internalized, potentially leading to inefficiency.

This model gives another example of why equilibrium segregation does not imply efficient segregation.

4.3. Empirical issues and evidence. Peer group effects are generally difficult to identify. In addition, we can think of two alternative formulations where one is practically impossible to identify satisfactorily. To discuss these issues, let us go back to the previous discussion, and recall that the two "structural" formulations, (3.15) and (3.16), have very similar reduced forms, but the peer group effects work quite differently, and have different interpretations. In (3.15), it is the (predetermined) characteristics of my peers that determine my outcomes, whereas in (3.16), it is the outcomes of my peers that matter. Above we saw how to identify externalities in human capital, which is in essence similar to the structural form in (3.15). More explicitly, the equation of interest is

$$y_{ij} = \theta x_{ij} + \alpha \bar{X}_j + \varepsilon_{ij} \tag{3.18}$$

where $\bar{X}$ is average characteristic (e.g., average schooling) and y_{ij} is the outcome of the ith individual in group j. Here, for identification all we need is exogenous variation in $\bar{X}$.

The alternative is

$$y_{ij} = \theta x_{ij} + \alpha \bar{Y}_j + \varepsilon_{ij} \tag{3.19}$$

where $\bar{Y}$ is the average of the outcomes. Some reflection will reveal why the parameter α is now practically impossible to identify. Since $\bar{Y}_j$ does not vary by individual, this regression amounts to one of $\bar{Y}_j$ on itself at the group level. This is a serious econometric problem. One imperfect way to solve this problem is to replace $\bar{Y}_j$ on the right hand side by $\bar{Y}_j^{-i}$ which is the average excluding individual i. Another approach is to impose some timing structure. For example:

$$y_{ijt} = \theta x_{ijt} + \alpha \bar{Y}_{j,t-1} + \varepsilon_{ijt}$$

There are still some serious problems irrespective of the approach taken. First, the timing structure is arbitrary, and second, there is no way of distinguishing peer group effects from "common shocks".

As an example consider the paper by Sacerdote, which uses random assignment of roommates in Dartmouth. He finds that the GPAs of randomly assigned roommates are correlated, and interprets this as evidence for peer group effects. The next table summarizessome of the key results.

PEER EFFECTS WITH RANDOM ASSIGNMENT 689

TABLE II
OWN PRETREATMENT CHARACTERISTICS REGRESSED ON ROOMMATE PRETREATMENT CHARACTERISTICS
EVIDENCE OF THE RANDOM ASSIGNMENT OF ROOMMATES

	(1) SAT Math (self)	(2) SAT Verbal (self)	(3) HS Academic class index	(4) HS Rank	(5) HS Academic index
roommates' math SAT scores	−0.025 (0.028)				−0.005 (0.008)
roommates' verbal SAT scores		−0.009 (0.029)			−0.005 (0.007)
roommates' HS academic scores			0.010 (0.028)		0.055 (0.056)
roommates' HS class ranks				−0.032 (0.028)	0.031 (0.042)
roommates' HS class rank missing					−0.512 (0.838)
Dummies for housing questions	yes	yes	yes	yes	yes
F-test: All roommate background coeff = 0					$F(5, 1543) = 0.50$ $P > F = .78$
R^2	.09	.03	.04	.03	.04
N	1589	1589	1589	993	1589

Standard errors are in parentheses. In cases with more than one roommate, roommate variables are averaged.

Columns (1)–(5) are OLS. All regressions include 41 dummies representing nonempty blocks based upon responses to the housing questions.

The lack of statistical significance on the coefficients is intended to demonstrate that the assignment process resembles a ràndomized experiment. In earlier nonrandomly assigned classes (such as the classes of 1995–1996), own and roommate background are highly correlated.

FIGURE 3.4

Despite the very nice nature of the experiment, the conclusion is problematic, because Sacerdote attempts to identify (3.19) rather than (3.18). For example, to the extent that there are common shocks to both roommates (e.g., they are in a noisier dorm), this may not reflect peer group effects. Instead, the problem would not have arisen if the right-hand side regressor was some predetermined characteristic of the roommate (i.e., then we would be estimating something similar to (3.18) rather than (3.19)).

Part 2

Incentives, Agency and Efficiency Wages

A key issue in all organizations is how to give the right incentives to employees. This topic is central to contract theory and organizational economics, but it also needs to be taken into account in labor economics, especially in order to better understand the employment relationship.

Here we give a quick overview of the main issues.

CHAPTER 4

Moral Hazard: Basic Models

Moral hazard refers to a situation where individual takes a "hidden action" that affects the payoffs to his employer (the principal). We generally think of this as the level of "effort", but other actions, such as the composition of effort, the allocation of time, or even stealing, are potential examples of moral hazard-type behavior. Although effort is not observed, some of the outcomes that the principal cares about, such as output or performance, are observed.

Because the action is hidden, the principal cannot simply dictate the level of effort. She has to provide incentives through some other means. The simplest way to approach the problem is to think of the principal as providing "high-powered" incentives, and rewarding success. This will work to some degree, but will run into two sorts of problems;

(1) Limited liability
(2) Risk

More explicitly, high-powered incentives require the principal to punish the agent as well as to reward him, but limited liability (i.e., the fact that the agent cannot be paid a negative wage in many situations) implies that this is not possible. Therefore high-powered incentives come at the expense of *high average level of payments.*

The risk problem is that rewarding the agent as a function of performance conflicts with optimal risk sharing between the principal and the agent. Generally, we think of the agent as earning most of his living from this wage income, whereas the principal employs a number of similar agents, or is a corporation with diffuse ownership. In that case, we can think of the firm as risk neutral and the employment contract should not only provide incentives to the agent, but also insure him against fluctuations in performance. More generally, even if the firm is risk averse, the employment contract should involve an element of risk sharing between the firm and the worker. Risk sharing in employment contracts will often contradict with the provision of incentives.

Because the incentive-insurance tradeoff is a central problem, moral hazard problems often arise in the context of health insurance, in fact, the term moral hazard originates from this literature. In particular, the idea is that if an individual is provided with full insurance against all of the possible health expenses that he may incur (which is good from risk-sharing point of view), he may be discouraged from undertaking hazardous behavior, potentially increasing the risk of bad health outcomes.

1. The Baseline Model of Incentive-Insurance Trade off

Let us start with the one agent case, and build on the key paper by Holmström, "Moral Hazard and Observability" *Bell Journal* 1979.

Imagine a single agent is contracting with a single principal.

The agent's utility function is

$$H(w, a) = U(w) - c(a)$$

where w is the wage he receives, U is a concave (risk-averse) utility function and $a \in \mathbb{R}_+$ denotes his action, with $c(\cdot)$ an increasing and convex cost function. Basically, the agent likes more income and dislikes effort.

The agent has an "outside option," representing the minimum amount that he will accept for accepting the employment contract (for example, this outside option may be working for another firm or self-employment). These are represented by some reservation utility $\bar{H}$ such that the agent would not participate in the employment relationship unless he receives at least this level of utility. This will lead to his *participation constraint.*

The action that the agent takes affects his performance, which we simply think of as output here. Let us denote output by x, and write

$$x(a,\theta)$$

where $\theta \in \mathbb{R}$ the state of nature. In other words,

$$x : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$$

This emphasizes that output depends on effort and some other influences outside the control of the agent and the principal. There is therefore a stochastic element.

Since greater effort should correspond with good things, we assume that

$$x_a \equiv \frac{\partial x}{\partial a} > 0$$

If output were a non-stochastic function of effort, contracting on output would be equivalent to contracting on effort, and risk sharing issues would not arise. Here θ is the source of risk.

The principal cares about output minus costs, so her utility function is

$$V(x-w)$$

where V is also an increasing concave utility function. A special case of interest is where V is linear, so that the principal is risk neutral.

What is a contract here?

Let Ω be the set of observable and contractible events, so when only x is observable, $\Omega = \mathbb{R}$. When any two of x, a, and θ are observable, then $\Omega = \mathbb{R}_+ \times \mathbb{R}$ (the third one is redundant given the information concerning the first two).

A contract is a mapping

$$s : \Omega \to \mathbb{R}$$

which specifies how much the agent will be paid.

Alternatively when there is limited liability so that the agent cannot be paid a negative wage,

$$s : \Omega \to \mathbb{R}_+$$

Here let us start with the case without limit liability.

Digression: what is the difference between observable and contractible? What happens if something is observable only by the principal and the agent, but by nobody else?

What we have here is a dynamic game, so the timing is important. It is:

Timing:

(1) The principal offers a contract $s : \Omega \to \mathbb{R}$ to the agent.
(2) The agent accepts or rejects the contract. If he rejects the contract, he receives his outside utility $\overline{H}$.
(3) If the agent accepts the contract $s : \Omega \to \mathbb{R}$, then he chooses effort a.
(4) Nature draws θ, determining $x(a, \theta)$.
(5) Agent receives the payment specified by contract s.

This is a game of incomplete information and as in signaling games, we will look for a Perfect Bayesian Equilibrium. However, in this context, the concept of Perfect Bayesian Equilibrium will be strong enough.

2. Incentives without Asymmetric Information

Let us start with the case of full information. Then the problem is straightforward.

The principal chooses both the contract $s(x, a)$ (why is it a function of both x and a?), and the agents chooses a. The Perfect Bayesian Equilibrium can be characterized by backward induction. The first interesting action is at step 3, where the agent chooses the effort level given the contract and then at step 2, where the agent decides whether to accept contract s. Given what types of contracts will be accepted by the agent and what the corresponding effort level will be, at step 1 the principal chooses the contract that maximizes her utility. With analogy to oligopoly games, we can think of the principal, who moves first, as a Stackleberg leader. As usual with Stackelberg leaders, when choosing the contract the principal anticipates the action that the agent will choose. Thus, we should think of the principal as choosing the effort level as well, and the optimization condition of the agent will be a constraint for the principal. This is what we refer to as the *incentive compatibility constraint (IC).*

Thus the problem is

$$\begin{array}{lll} & \max\limits_{s(x,a),a} \mathbb{E}\left[V(x - s(x,a)\right] & \\ \text{s.t.} & \mathbb{E}\left[H(s(x,a),a)\right] \geq \overline{H} & \text{Participation Constraint (PC)} \\ \text{and} & a \in \arg\max\limits_{a'} \mathbb{E}\left[H(s(x,a'),a')\right] & \text{Incentive Constraint (IC)} \end{array}$$

where expectations are taken over the distribution of θ.

This problem has exactly the same structure as the canonical moral hazard problem, but is much simpler, because the principal is choosing $s(x,a)$. In particular, she can choose s such that $s(x,a) = -\infty$ for all $a \neq a^*$, thus effectively implementing a^*. This is because there is no moral hazard problem here given that there is no *hidden action.*

Therefore, presuming that the level of effort a^* is the optimum from the point of view of the principal, the problem collapses to

$$\max_{s(x)} \mathbb{E}\left[V(x - s(x)\right]$$

subject to

$$\mathbb{E}\left[U\left(s\left(x\right)\right)\right] \geq \overline{H} + c\left(a^*\right)$$

where we have already imposed that the agent will choose a^* and the expectation is conditional on effort level a^*. We have also dropped the incentive compatibility constraint, and rewrote the participation constraint to take into account of the equilibrium level of effort by the agent.

This is simply a risk-sharing problem, and the solution is straightforward. It can be found by setting up a simple Lagrangean:

$$\min_{\lambda} \max_{s(x)} \mathcal{L} = \mathbb{E}\left[V(x - s(x)\right] - \lambda\left[\overline{H} + c\left(a^*\right) - \mathbb{E}\left[U\left(s\left(x\right)\right)\right]\right]$$

Now this might appear as a complicated problem, because we are choosing a function $s(x)$, but this specific case is not difficult because there is no constraint on the form of the function, so the maximization can be carried out pointwise (think, for example, that x only took discrete values).

We might then be tempted to write:

$$\mathbb{E}\left[V'(x - s(x)\right] = \lambda \mathbb{E}\left[U'\left(s\left(x\right)\right)\right].$$

However, this is not quite right, and somewhat misleading. Recall that $x = x\left(a, \theta\right)$, so once we fix $a = a^*$, and conditional on x, there is no more uncertainty. In other words, the right way to think about the problem is that for a given level of a, the variation in θ induces a distribution of x, which typically we will refer to as $F\left(x \mid a\right)$ in what follows. For now, since $a = a^*$ and we can choose $s\left(x\right)$ separately for each x, there is no more uncertainty conditional on x.

Hence, the right first-order conditions are:

$$\frac{V'(x - s(x))}{U'\left(s\left(x\right)\right)} = \lambda \text{ for all } x, \tag{4.1}$$

i.e., perfect risk sharing. In all states, represented by x, the marginal value of one more dollar to the principal divided by the marginal value of one more dollar to the agent must be constant.

3. Incentives-Insurance Trade-off

Next, let us move to the real principle-agent model where Ω only includes the output performance, x, so feasible contracts are of the form $s\left(x\right)$, and are not conditioned on a. The effort is chosen by the agents to maximize his utility, and the incentive compatibility constraint will play an important role. The problem can be written in a similar form to before as

$$\begin{array}{lll} & \max\limits_{s(x),a} \mathbb{E}\left[V(x - s(x)\right] & \\ \text{s.t.} & \mathbb{E}\left[H(s(x), a)\right] \geq \overline{H} & \text{Participation Constraint (PC)} \\ \text{and} & a \in \arg\max\limits_{a'} \mathbb{E}\left[H(s(x), a')\right] & \text{Incentive Constraint (IC)} \end{array}$$

with the major difference that $s\left(x\right)$ instead of $s\left(x, a\right)$ is used.

As already hinted above, the analysis is more tractable when we suppress θ, and instead directly work with the distribution function of outcomes as a function of the effort level, a:

$$F(x \mid a)$$

A natural assumption is

$$F_a(x \mid a) < 0,$$

which is related to and implied by $x_a > 0$. Expressed differently, an increase in a leads to a first-order stochastic-dominant shift in F. Recall that we say a distribution function F first-order stochastically dominates another G, if

$$F(z) \leq G(z)$$

for all z (alternatively, the definition of first-order stochastic dominance may be strengthened by requiring the inequality to be strict for some z).

Using this way of expressing the problem, the principal's problem now becomes

$$\begin{aligned} \max_{s(x),a} \quad & \int V(x - s(x))dF(x \mid a) \\ \text{s.t.} \quad & \int [U(s(x) - c(a))]\, dF(x \mid a) \geq \overline{H} \\ a \in \arg\max_{a'} \quad & \int [U(s(x)) - c(a')]\, dF(x \mid a') \end{aligned}$$

This problem is considerably more difficult, because the second, the IC, constraint is no longer an inequality constraint, but an abstract constraint requiring the value of a function, $\int [U(s(x)) - c(a')]\, dF(x \mid a')$, to be highest when evaluated at $a' = a$.

It is very difficult to make progress on this unless we take some shortcuts. The standard shortcut is called the "*first-order approach*,"and involves replacing the second constraint with the first-order conditions of the agent. Now this would be no big step if the agent's problem

$$\max_{a'} \int [U(s(x)) - c(a')]\, dF(x \mid a')$$

was strictly concave, but we can make no such statement since this problem depends on $s(x)$, which is itself the choice variable that the principal chooses. For sufficiently non-convex s functions, the whole program will be non-concave, thus first-order conditions will not be sufficient.

Therefore, the first-order approach always comes with some risks (and one should not apply it without recognizing these risks and the potential for making mistakes, though there are many instances in which it is applied when it should not be). Nevertheless, bypassing those, and in addition assuming that F is twice continuously differentiable, so that the density function f exists, and in turn is differentiable with respect to a, the first-order condition for the agent is:

$$\int U(s(x))f_a(x \mid a)dx = c'(a).$$

Now using this, we can modify the principal's problem to

$$\min_{\lambda,\mu} \max_{s(x),a} \mathcal{L} = \int \{V(x - s(x)) \quad + \lambda \left[U(s(x)) - c(a) - \overline{H}\right] \qquad + \\ \mu \left[U(s(x))\frac{f_a(x \mid a)}{f(x \mid a)} - c'(a)\right]\right\} f(x \mid a)dx$$

Again carrying out point-wise maximization with respect to $s(x)$:

$$0 = \frac{\partial \mathcal{L}}{\partial s(x)} = -V'(x - s(x)) + \lambda U'(s(x)) + \mu U'(s(x))\frac{f_a(x \mid a)}{f(x \mid a)} \text{ for all } x$$

This implies

$$\frac{V'(x - s(x))}{U'(s(x))} = \lambda + \mu \frac{f_a(x \mid a)}{f(x \mid a)}. \tag{4.2}$$

The nice thing now is that this is identical to (4.1) if $\mu = 0$, that is, if the incentive compatibility constraint is slack. As a corollary, if $\mu \neq 0$, and the incentive compatibility constraint is binding, there will be a trade-off between insurance and incentives, and (4.2) will be different from (4.1). What sign should μ be? Since μ is the multiplier associated with anequality constraint, we cannot say this on a priori grounds. But it can be proved under some regularity conditions (in particular, when Monotone Likelihood Ratio Principle introduced below holds) that $\mu > 0$. Let us assume that this is indeed the case for now.

Note also that the solution must feature $\lambda > 0$, i.e., the participation constraint is binding. Why is this? Suppose not. Then, the principal could reduce $s(x)$ for all x by a little without violating incentive compatibility and increase her net income. Equivalently, in this problem the agent is receiving exactly what he would in his next best opportunity, and is obtaining no *rents*. (There are no rents, because there is no constraints on the level of payments; we will see below how this will change with limited liability constraints).

We can also use (4.2) to derive further insights about the trade-off between insurance and incentives.

To do this, let us assume that V' is constant, so that the principal is risk neutral.

Let us ask what it would take to make sure that we have full insurance, i.e., $V'(x - s(x))/U'(s(x))=$ constant. Since V' is constant, this is only possible if U' is constant. Suppose that the agent is risk-averse, so that U is strictly concave or U' is strictly decreasing. Therefore, full insurance (or full risk sharing) is only possible if $s(x)$ is constant. But in turn, if $s(x)$ is constant, the incentive compatibility constraint will be typically violated (unless the optimal contract asks for $a = 0$), and the agent will choose $a = 0$.

Next, consider another extreme case, where the principal simply sells the firm to the agent for a fixed amount, so $s(x) = x - s_0$. In this case, the agent's first-order condition will give a high level of effort (we can think of this as the "first-best" level of effort, though this is not literally true, since this level of effort potentially depends on s_0):

$$\int U(x - s_0) f_a(x \mid a) dx = c'(a).$$

This higheeer level of effort comes at expense of no insurance for the agent.

Instead of these two extremes, the optimal contract will be "second-best", trading off incentives and insurance.

We can interpret the solution (4.2) further. But first, note that as the optimization problem already makes it clear, as long as the IC constraint of the agent has a unique solution, once the agent signs to contract $s(x)$, there is no uncertainty about action choice a. Nevertheless, lack of full insurance means that the agent is being punished for low realizations of x. Why is this?

At some intuitive level, this is because had it not been so, ex ante the agent would have had no incentive to exert high effort. What supports high effort here is the *threat of punishment ex post.*

This interpretation suggests that there is no need for the principal to draw inferences about the effort choice a from the realizations of x. However, it turns out that the optimal way of incentivizing the agent has many similarities to an optimal signal extraction problem.

To develop this intuition, consider the following maximum likelihood estimation problem: we know the distribution of x conditional on a, we observe x, and we want to estimate a. This is a solution to the following maximization problem

$$\max_{a'} \ln f(x \mid a'),$$

for given x, which has the first-order condition

$$\frac{f_a(x \mid a')}{f(x \mid a')} = 0$$

which can be solved for $a(x)$. Let the level of effort that the principal wants to implement be $\bar{a}$, then $a(x) = \bar{a}$, this first-order condition is satisfied.

Now going back to (4.2), we can write this as:

$$\frac{V'(x - s(x))}{U'(s(x))} = \lambda + \mu \frac{f_a(x \mid \bar{a})}{f(x \mid \bar{a})}.$$

If $a(x) > \bar{a}$, then $f_a(x \mid \bar{a})/f(x \mid \bar{a}) > 0$. Since $\mu > 0$, this implies that V'/U' must be greater and therefore U' must be lower. This is in turn possible only when $s(x)$ is increasing in x. Therefore, when the realization of output is *good news* relative to what was expected, the agent is rewarded, when it is *bad news*, he is punished.

Thus in a way, the principal is acting as if she's trying to infer what the agent did, even though of course the principal *knows* the agent's action along the equilibrium path.

4. The Form of Performance Contracts

Can we say anything else on the form of $s(x)$? At a minimum, we would like to say that $s(x)$ is increasing, so that greater output leads to greater renumeration for the agent, which seems to be a feature of real world contracts for managers, workers etc.

Unfortunately, this is not true without putting more structure on technology. Consider the following example where the agent chooses between two effort levels, high and low:

$$a \in \{a_H, a_L\}$$

and the distribution function of output conditional on effort is as follows:

$$F(x \mid a_H) = \begin{cases} 4 & \text{with probability } \frac{1}{2} \\ 2 & \text{with probability } \frac{1}{2} \end{cases}$$

$$F(x \mid a_L) = \begin{cases} 3 & \text{with probability } \frac{1}{2} \\ 1 & \text{with probability } \frac{1}{2} \end{cases}$$

The agent has an arbitrary strictly concave utility function.

It is quite clear that in this case full risk-sharing can be achieved (what does this mean in terms of the multipliers in our formulation above?). In particular, full risk sharing is possible if the principal punishes the agent whenever 1 or 3 is observed. In fact, the following contract would do the trick:

$$s(2) = s(4) = \overline{H} + c\,(a_H)$$

$$s(1) = s(3) = -K$$

where K is a very large number. Thus the agent is punished severely for the outcomes 1 or 3, since these occur only when he chooses low effort. When the outcome is 2 or 4, he gets a payment consistent with his participation constraint.

Clearly this contract is not increasing in x, in particular, $s\,(3) < s\,(2)$.

You might wonder whether there is something special here because of the discrete distribution of x. This is not the case. For example, a continuous distribution with peaks at $\{2, 4\}$ for $a = a_H$ and $\{1, 3\}$ for $a = a_L$ would do the same job.

So how can we ensure that $s\,(x)$ is increasing in x?

Milgrom, *Bell Journal*, 1981, "Good News, Bad News" shows the following result:

A sufficient condition for $s(x)$ to be increasing is that higher values of x are "good news" about a

$$\text{i.e., } \frac{f_a(x \mid a)}{f(x \mid a)} \text{ is increasing in } x$$

$$\Leftrightarrow \frac{f(x \mid a_1)}{f(x \mid a_2)} \text{ is increasing in } x \text{ for } a_1 > a_2$$

This is referred to as the *Monotone Likelihood Ratio Property* (MLRP).

We can also note that this implies that

$$\int xf(x \mid a_1)dx > \int xf(x \mid a_2)dx \text{ for } a_1 > a_2,$$

meaning this condition is sufficient (but not necessary) for the expected value of x to increase with the level of effort.

Given the characterization above in terms of inferring $a(x)$ from x, we need $a(x)$ to be increasing, so that higher output levels correspond to better news about the level of effort that the agent must have exerted.

This is clearly not the case with our example above, where 3 relative to 2 is bad news about the agent having exerted a high level of effort.

When we assume MLRP, the result that we can also show that the multiplier μ must be positive. To see this, note that if MLRP holds and $\mu < 0$, then with the same argument as above (4.2) implies that $s(x)$ must be decreasing in x everywhere (since $f_a(x \mid a)/f(x \mid a)$ is increasing, V'/U' would be decreasing in x). However, if $s(x)$ is decreasing everywhere, the agent would necessarily choose the lowest effort level and the incentive compatibility constraint would then be slack, and thus μ must be equal to zero,leading to a contradiction with the hypothesis that $\mu < 0$. This establishes that when MLRP holds, we must have $\mu > 0$.

5. The Use of Information: Sufficient Statistics

Finally, another important result that follows from this framework is that of a *sufficient statistic* result. Imagine that in addition to x, the principal observes another signal of the agent's effort, y, in the sense that y is a random variable with distribution $G(y \mid a)$. The principal does not care about y per se, and still wants to $\max \mathbb{E}(V(x-s))$.

The key question is whether the principal should offer a contract $s(x, y)$ which depends (non-trivially) on the signal y as well as the output x?

The answer is: *yes*, if y helps reduce noise or yields extra information on a, and *no* if x is a sufficient statistic for (x, y) in the estimation of a. Recall that a statistic T is a sufficient statistic for some family of random variables $\mathcal{F}$ in estimating a parameter $\theta \in \Theta$ if and only if the marginal distribution of θ conditional on T and $\mathcal{F}$ coincide, that is,

$$f(\theta \mid T) = f(\theta \mid \mathcal{F}) \text{ for all } \theta \in \Theta.$$

To develop this point more formally, let us look at the first-order conditions for choosing $s(x, y)$. With direct analogy to before, the first-order conditions imply:

$$\frac{V'(x - s(x, y))}{U'(s(x, y))} = \lambda + \mu \frac{f_a(x, y \mid a)}{f(x, y \mid a)}$$

The problem we are interested in can be posed as whether $s(x, y) = S(x)$ for all x and y, where $s(x, y)$ is the solution to the maximization problem.

Equivalently,

$$s(x, y) = S(x) \text{ for all } x \text{ and } y \text{ if and only if } \frac{f_a(x, y \mid a)}{f(x, y \mid a)} = k(x \mid a) \text{ for all } x \text{ and } y$$

for some function k.

What does this condition mean? The condition

$$\frac{f_a(x, y \mid a)}{f(x, y \mid a)} = k(x \mid a) \; \forall \, x, y$$

is equivalent to $f(x, y \mid a) = g(x, y)h(x \mid a)$ (simply differentiate both sides with respect to a to verify this claim). This condition, in turn, means that *conditional on* x, y has no additional information on a, or using Bayes' rule

$$f(a \mid x, y) = f(a \mid x)$$

that is, x is a *sufficient statistic* for (x, y) with respect to inferences about a.

The implication is the important suggested result: the optimal contract conditional on x and y, $s(x, y)$, will not use y if and only if x is a sufficient statistic for (x, y) with respect to a.

CHAPTER 5

Moral Hazard with Limited Liability, Multitasking, Career Concerns, and Applications

1. Limited Liability

Let us modify the baseline moral hazard model by adding a *limited liability constraint*, so that $s(x) \geq 0$.

The problem becomes:

$$\max_{s(x),a} \int V(x - s(x))dF(x \mid a)$$

subject to

$$\begin{aligned} & \int [U(s(x) - c(a))] \, dF(x \mid a) \geq \overline{H} \\ a \quad \in \quad & \arg\max_{a'} \int [U(s(x)) - c(a')] \, dF(x \mid a') \\ & s(x) \geq 0 \text{ for all } x \end{aligned}$$

Again taking the first-order approach, and assigning a multiplier $\eta(x)$ to the last set of constraints, the first but her conditions become:

$$V'(x - s(x)) = \left[\lambda + \mu \frac{f_a(x \mid a)}{f(x \mid a)}\right] U'(s(x)) + \eta(x).$$

If $s(x)$ was going to be positive for all x in any case, the multiplier for the last set of constraints, $\eta(x)$, would be equal to zero, and the problem would have an identical solution to before.

However, if, previously, $s(x) < 0$ for some x, the structure of the solution has to change. In particular, to obtain the intuition, suppose that we shift up the entire function $s(x)$ to $\tilde{s}(x)$ so that $\tilde{s}(x) \geq 0$. Since the participation constraint was binding at $s(x)$, it must be slack at $\tilde{s}(x)$. Clearly this will not be optimal and in fact because of income effects, this shifted-up schedule may no longer lead to the same optimal choice of effort for the agent. In particular, as we increase the level of payments at low realizations of x, the entire payment schedule has to change in a more complex way. Nevertheless, this "shifting-up" intuition makes it clear that the participation constraint will no longer be binding, thus $\lambda = 0$.

This informally is the basis of the intuition that without limited liability constraints, there are no rents; but with limited liability there will be rents, making the agent's participation constraint slack.

Let us now illustrate this with a simple example. Suppose that effort takes two values $a \in \{a_L, a_H\}$. Assume that output also takes only two values: $x \in \{0, 1\}$, moreover,

$$\begin{aligned} F(x \mid a_H) &= \quad 1 \quad \text{with probability } 1 \\ F(x \mid a_L) &= \begin{cases} 1 & \text{with probability } q \\ 0 & \text{with probability } 1 - q \end{cases} \end{aligned}$$

Normalize $\bar{H}$ and $c(a_L)$ to zero, and assume $c(a_H) = c_H < 1 - q$.

Finally, to make things even simpler, assume that both the agent and the principal are risk neutral.

Let us first look at the problem without the limited liability constraint. The assumption that $c(a_H) = c_H < 1 - q$ implies that high effort is optimal, so in an ideal world this would be the effort level.

Let us first start by assuming that the principal would like to implement this. In this case, the problem of the principal can be written as

$$\min_{s(0), s(1)} s(1)$$

subject to

$$\begin{aligned} s(1) - c_H &\geq qs(1) + (1 - q) s(0) \\ s(1) - c_H &\geq 0 \end{aligned}$$

where $s(0)$ and $s(1)$ are the payments to the agent conditional on the outcome (Why are these the only two control variables?)

The first constraint is the incentive compatibility constraint; it requires that the agent prefers to exert high effort and to receive the high payment rather than taking the gamble between high and low payment, while also saving the cost of effort (this statement is written presuming that $s(0) < s(1)$, which will be the case).

The second constraint is the participation constraint, requiring that the along-the-equilibrium-path payment to the agent exceed his outside option, 0.

This problem does not impose a limited liability constraint yet.

The principal simply minimizes the cost of hiring the agent, since conditional on implementing the high effort, there is no other interesting choice for her.

The solution is straightforward, and involves the participation constraint holding as equality, thus

$$s(1) = c_H$$

Then the incentive compatibility constraint implies that

$$s(0) \leq -\frac{q}{1-q} c_H,$$

so that the agent receives a harsh enough punishment for generating the wrong level of output. It can also be verified that in this case the principal indeed prefers to implement the high level of effort.

Clearly, $s(0)$ needs to be negative, so the solution will not be possible when we impose limited liability.

Let us now look at the problem with the limited liability constraint. Again presuming that the high level of effort will be implemented, the maximization problem boils down to:

$$\min_{s(0),s(1)} s(1)$$

subject to

$$\begin{aligned} s(1) - c_H &\geq qs(1) + (1-q)s(0) \\ s(1) - c_H &\geq 0 \\ s(0) &\geq 0 \end{aligned}$$

where we could have also imposed $s(1) \geq 0$, but did not, because this constraint will clearly be slack (why?).

It is straightforward to verify that solution to this problem will be

$$\begin{aligned} s(0) &= 0 \\ s(1) &= \frac{c_H}{1-q} \end{aligned}$$

Thus now, when successful, the agent is paid more than the case without the limited liability constraint, and as a consequence, the participation constraint is slack. A different way of expressing this is that now the agent receives *a rent* from the employment relationship. This rent can be easily calculated to be equal to

$$\text{rent} = \frac{q}{1-q} c_H$$

As a result, with limited liability, we have the issue of rents in addition to the issue of insurance.

We can also see that the presence of rents may actually distort the choice of effort. To develop this point further, let us calculate the return to the principal with high effort. It is clearly

$$\text{Return}_H = 1 - \frac{1}{1-q} c_H$$

In contrast, if he chooses the low effort, he can pay the agent $s(0) = s(1) = 0$, thus making:

$$\text{Return}_L = q$$

which can be greater than Return_H. In contrast, without rents for the agent, the return to the principal from implementing high effort would have been $1 - c_H$, which is greater than q by assumption. This implies that even though high effort might be "socially optimal" in the sense of increasing net output (net surplus), the principal may choose low effort in order to reduce the rents that the agent receives (and thus distort the structure of production and effort).

The limited liability constraint and the associated rents will play a very important role below when we discuss efficiency wages.

2. Linear Contracts

One problem with the baseline model developed above is that, despite a number of useful insights, it is quite difficult to work with. Moreover, the exact shape of the density functions can lead to very different forms of contracts, some with very nonlinear features.

One approach in the literature has been to look for "robust" contracts that are both intuitively simpler and easier to work with to derive some first-order predictions. But why should optimal contracts be "robust"? And, how do we model "robust" contracts?

A potentially promising answer to this question is developed in an important paper by Holmstrom and Milgrom. They established the optimality of linear contracts under certain conditions, which is interesting both because linear contracts can be viewed at as more robust than highly nonlinear contracts, and also because the intuition of their result stems from robustness considerations.

Providing a detailed exposition of Holmstrom and Milgrom's model would take us too far afield from our main focus. Nevertheless, it is useful to outline the environment and the main intuition. Holmstrom and Milgrom consider a *dynamic principal-agent* problem in continuous time. The interaction between the principal and the agent take place over an interval normalized to $[0, 1]$. The agent chooses an effort level $a_t \in A$ at each instant after observing the relaxation of output up to that instant. More formally, the output process is given by the continuous time random walk, that is, the following *Brownian motion process*:

$$dx_t = a_t dt + \sigma dW_t$$

where W is a standard Brownian motion (Wiener process). This implies that its increments are independent and normally distributed, that is, $W_{t+\tau} - W_t$ for any t and τ is distributed normally with variance equal to τ. Let $X^t = (x_\tau; 0 \leq \tau < t)$ be the entire history of the realization of the increments of output x up until time t (or alternatively a "sample path" of the random variable x). The assumption that the individual chooses a_t after observing past realizations implies that a_t can be represented by a mapping $a_t : X^t \to A$. Similarly, the principal also observes the realizations of the increments (though obviously not the effort levels and the realizations of W_t), so a contract for the agent is given by a mapping $s_t : X^t \to \mathbb{R}$, specifying what the individual will be paid at time t is a function of the entire realization of output levels up to that point.

Holmstrom and Milgrom assume that the utility function of the agent be

$$u\left(C_1 - \int_0^1 a_t d_t\right)$$

where C_1 is the agent's consumption at time $t = 1$. This utility function makes two special assumptions: first, the individual only derives utility from consumption at the end (at time $t = 1$) and second, the concave utility function applies to consumption minus the total cost of effort between 0 and 1. In addition, Holmstrom and Milgrom assume that u takes the special constant absolute risk aversion, CARA, form

$$u(z) = -\exp(-rz) \tag{5.1}$$

with the degree of absolute risk aversion equal to r, and that the principal is risk neutral, so that she only cares about her net revenue at time $t = 1$, given by $x_1 - C_1$ (since consumption of the agent at time $t = 1$ is equal to total payments from the principal to the agent).

The key result of Holmstrom and Milgrom is that in this model, the optimal contract is linear in final (cumulative) output x_1. In particular, it does not depend on the exact sample path leading to this cumulative output. Moreover, in response to this contract the optimal behavior of the agent is to choose a constant level of effort, which is also independent of the history of past realizations of the stochastic shock (can you see why the utility function (5.1) is important here?).

The loose intuition is that with any nonlinear contract there will exist an event, i.e., a sample path, after which the incentives of the agent will be distorted, whereas the linear contract achieves a degree of "robustness". A more formal intuition is that we can think of a discrete approximation to the Brownian motion, which will be a binomial process specifying success or failure for the agent at each instant. The agent should be rewarded for success and punished for failure, and this will amount to the individual being renumerated according to total cumulative output. Moreover, generally this remuneration should depend on the wealth level of the agent, but with CARA, the wealth level does not matter, so the reward is constant. A linear reward schedule is the limit of this process corresponding to the continuous time limit of the binomial process, which is the Brownian motion.

Now motivated by this result, many applied papers look at the following *static* problem:

(1) The principal chooses a linear contract, of the form $s = \alpha + \beta x$ (note that this implies there is no limited liability; and we have also switched from S to s to simplify notation).
(2) The agents chooses $a \in A \equiv [0, \infty]$.
(3) $x = a + \varepsilon$ where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$

In addition, the principal is risk neutral, while the utility function of the agent is

$$U(s, a) = -\exp(-r(s - c(a)))$$

with $c(a) = ca^2/2$ corresponding to the cost of effort for some $c > 0$.

The argument is that a linear contract is approximately optimal here.

It turns out that the results of this framework are very intuitive and consistent with the baseline model. However, it is important to emphasize that a linear contract is not optimal in this case (it is only optimal in the Holmstrom-Milgrom model with continuous time and the other assumptions; in fact, it is a well-known result in agency theory that a static problem with a normally distributed outcomes has sufficiently unlikely events that the first-best level of effort, which here is $a^{fb} = 1/c$, can be approximated by highly nonlinear contracts, thus the linear contracts studied here are very different from the optimal contracts that would arise if the actual model has been the static model with normally distributed shocks).

Let us derive the optimal contract in this case.

The first-order approach works in this case. The maximization problem of the agent is

$$\begin{aligned} & \max_a \mathbb{E}\left\{-\exp\left(-r\left(s\left(a\right)-c\left(a\right)\right)\right)\right\} \\ = & \max_a \left\{-\exp\left(-r\mathbb{E}s\left(a\right)+\frac{r^2}{2}\mathrm{Var}\left(s\left(a\right)\right)-rc\left(a\right)\right)\right\} \end{aligned}$$

where the equality between the two expressions follows from the normality of s, (s is a linear in x, and x is normally distributed), so this is equivalent to

$$\max_a \left\{\mathbb{E}s\left(a\right)-\frac{r}{2}\mathrm{Var}\left(s\left(a\right)\right)-\frac{c}{2}a^2\right\}$$

Now substituting for the contract, the problem is:

$$\max_a \beta a-\frac{c}{2}a^2-\frac{r}{2}\beta^2\sigma^2$$

so the first-order condition for the agent's optimal effort choice is:

$$a=\frac{\beta}{c}$$

The principal will then maximize

$$\max_{a,\alpha,\beta} \mathbb{E}\left(\left(1-\beta\right)\left(a+\varepsilon\right)-\alpha\right)$$

subject to

$$\begin{aligned} a & = \frac{\beta}{c} \\ \alpha+\frac{\beta^2}{2}\left(\frac{1}{c}-r\sigma^2\right) & \geq \bar{h} \end{aligned}$$

where the second inequality is the participation constraint, with the definition $\bar{h}=-\ln\left(-\bar{H}\right)$, where $\bar{H}$ is the reservation utility of the agent, and requires the expected utility of the agent under the contract to be greater than $\bar{H}$.

The solution to this problem is

$$\beta^*=\frac{1}{1+rc\sigma^2} \tag{5.2}$$

and

$$\alpha^*=\bar{h}-\frac{1-rc\sigma^2}{2c^2\left(1+rc\sigma^2\right)^2},$$

and because negative salaries are allowed, the participation constraint is binding.

In other words, the more risk-averse is the agent, i.e., the greater is r, the more costly is effort, i.e., the greater is c, and the more uncertainty there is, i.e., the greater is σ^2, the lower powered are the agent's incentives.

The equilibrium level of effort is

$$a^* = \frac{1}{c\left(1 + rc\sigma^2\right)}$$

This is always lower than the first-best level of effort which is $a^{fb} = 1/c$.

We can see that as $r \to 0$ and individual becomes more and more risk neutral, the equilibrium approaches this first-best level of effort. Similarly, the first-best applies as $\sigma^2 \to 0$, which corresponds to the case where risk disappears (and thus the model has a problem becomes mute).

Let us now derives some of the other results of the baseline model. Suppose that there is another signal of the effort

$$z = a + \eta,$$

where η is $\mathcal{N}\left(0, \sigma_\eta^2\right)$ and is independent of ε. Now, let us restrict attention to linear contracts of the form

$$s = \alpha + \beta_x x + \beta_z z.$$

Note that this contract can also be interpreted alternatively as $s = \alpha + \mu w$ where $w = w_1 x + w_2 z$ is a sufficient statistic derived from the two random variables x and z. This already highlights that the sufficient statistic principle is still at work here.

Now with this type of contract, the first-order condition of the agent is

$$a = \frac{\beta_x + \beta_z}{c}$$

and the optimal contract can be obtained as:

$$\beta_x = \frac{\sigma_\eta^2}{\sigma^2 + \sigma_\eta^2 + rc\left(\sigma^2 \sigma_\eta^2\right)}$$

and

$$\beta_z = \frac{\sigma^2}{\sigma^2 + \sigma_\eta^2 + rc\left(\sigma^2 \sigma_\eta^2\right)}$$

These expressions show that generally x is not a sufficient statistic for (x, z), and the principal will use information about z as well to determine the compensation of the agent.

The exception is when $\sigma_\eta^2 \to \infty$ so that there is almost no information in z regarding the effort chosen by the agent. In this case, $\beta_z \to 0$ and $\beta_x \to \beta^*$ as given by (5.2), so in this case x becomes a sufficient statistic.

3. Evidence

The evidence on the basic principal-agent model is mixed. A series of papers, notably those by Ed Lazear using data from a large auto glass installer, present convincing evidence that in a variety of settings high incentives lead to more effort. For example, Lazear's evidence shows that when this particular company went from fixed salaries to piece rates productivity rose by 35% because of greater effort by the employees (the increase in average wages was 12%), but part of this response might be due to selection, as the composition of employees might have changed.

Similar evidence is reported in other papers. For example, Kahn and Sherer, using the personnel files of a large company, show that employees (white-collar office workers) whose pay depends more on the subjective evaluations obtain better evaluations and are more productive.

More starkly, and perhaps more interestingly, a number of papers using Chinese data, in particular work by John McMillan, show that the responsibility system in Chinese agriculture, allowing local communes to retain a share of their profits led to substantial increases in productivity. Separate work by Ted Groves finds similar effects from the Chinese industry. This evidence is quite conclusive about the effect of incentives on effort and productivity. However, the principal-agent approach to contracting and to the incentive structure not only requires that effort and performance are responsive to incentives, but also that these incentives are designed optimally, and that the types of theories developed so far capture the salient features of these optimal contracts. The evidence in favor of this latter, more stringent evaluation of the principal-agent theory is weaker.

To start with, even though in some stock examples such as in Chinese agriculture or industry, higher-powered incentives (meaning greater rewards for success) lead to better outcomes, in other contexts more high-powered incentives seem to lead to counter-productive incentives. One example of this is the evidence in the paper by Ernst Fehr and Simon Gachter showing that incentive contracts might destroy voluntary cooperation.

More standard examples are situations in which high-powered incentives lead to distortions that were not anticipated by the principals. A well-known case is the consequences of Soviet incentive schemes specifying "performance" by number of nails or the weight of the materials used, leading to totally unusable products.

Evidence closer to home also indicates similar issues. A number of papers have documented that agents with high-powered incentives try to "game" these incentives (potentially creating costs for the principals). A telling example is work by Paul Oyer, and work by Pascal Courty and Gerard Marschke, which look at performance contract that are nonlinear functions of outcomes, and show that there is considerable gaming going on. For example, managers that get bonuses for reaching a particular target by a certain date put a lot of effort before this date, and much less during other times. This would be costly if a more even distribution of effort were optimal for the firm.

More generally, the greatest challenge to the principal-agent approach is that it does not perform well in terms of its predictions regarding the types of contracts that should be offered (how these contracts should very across environments). First, as discussed at length by Prendergast, there is little association between riskiness and noisiness of tasks and the types of contracts when we look at a cross section of jobs.

Second, and perhaps more starkly, in many professions performance contracts are largely absent. There is a debate as to whether this is efficient, for example, as in teaching and bureaucracy, but many believe that such contracts are absent precisely because their use would lead to distorted incentives in other spheres—as in the models of multitasking we discuss next. A widespread view related to this is that the basic moral hazard models are not useful in thinking about bureaucracies, where there are many countervailing effects related to multitasking and "career concerns," creating incentives for other types of behavior, and consequently the power of incentives are often weak in such organizations.

4. Multitasking

We now discuss incentive models in which agents undertake more than one task or more than one agent interact with the principal or perform similar tasks. These models are useful both to extend the reach of the agency theory, and also to generate some insights on why we may not see very high-powered incentives in most occupations.

Multitasking is the broad name given by Holmstrom and Milgrom to situations in which an agent has to work in more than one tasks. Multitasking is generally associated with problems of giving incentives to the agent in one sphere without excessively distorting his other incentives. In other words, multitasking is about balancing the distortions created indifference tasks undertaken by a single agent.

Let us now modify the above linear model so that there are two efforts that the individual chooses, a_1 and a_2, with a cost function $c(a_1, a_2)$ which is increasing and convex as usual.

These efforts lead to two outcomes:

$$x_1 = a_1 + \varepsilon_1$$

and

$$x_2 = a_2 + \varepsilon_2,$$

where ε_1 and ε_2 could be correlated. The principal cares about both of these inputs with potentially different weights, so her return is

$$\phi_1 x_1 + \phi_2 x_2 - s$$

where s is the salary paid to the agent.

What is different from the previous setup is that only x_1 is observed, while x_2 is unobserved.

A simple example is a home contractor where x_1 is an inverse measure of how long it takes to finish the contracted work, while x_2 is the quality of the job, which is not observed until much later, and consequently, payments cannot be conditioned on this.

Another example would be the behavior of employees in the public sector, where quality of the service provided to citizens is often difficult to contract on.

So what is the solution to this problem?

Again let us take a linear contract of the form

$$s(x_1) = \alpha + \beta x_1$$

since x_1 is the only observable output.

The first-order condition of the agent now gives:

$$\begin{aligned} \beta &= \frac{\partial c(a_1, a_2)}{\partial a_1} \\ 0 &= \frac{\partial c(a_1, a_2)}{\partial a_2} \end{aligned} \tag{5.3}$$

So if

$$\frac{\partial c(a_1, a_2)}{\partial a_2} > 0$$

whenever $a_2 > 0$, then the agent will choose $a_2 = 0$, and there is no way of inducing him to choose $a_2 > 0$.

However, suppose that

$$\frac{\partial x(a_1, a_2 = 0)}{\partial a_2} < 0,$$

so without incentives the agent will exert some positive effort in the second task. In this case, in fact providing incentives in task 1 can undermine the incentives in task 2. This will be the case when the two efforts are substitutes, i.e., $\partial^2 c(a_1, a_2)/\partial a_1 \partial a_2 > 0$, so that exerting more effort in one task increase the cost of effort in the other task.

Now, stronger incentives for task 1 increase effort a_1, reducing effort a_2 because of the substitutability between the two efforts.

To see this more formally, imagine that the equations in (5.3) have an interior solution (why is an interior solution important?), and differentiate these two first-order conditions with respect to β. Using the fact that these two first-order conditions correspond to a maximum (i.e., the second order conditions are satisfied), we can use the Implicit Function Theorem on (5.3), immediately see that

$$\frac{\partial a_1}{\partial \beta} > 0.$$

(It is useful for you to derive this yourself). This has the natural interpretation that high-powered incentives lead to stronger incentives as the evidence discussed above suggests.

However, we also have that if $\partial^2 c\left(a_1, a_2\right) / \partial a_1 \partial a_2 > 0$, then

$$\frac{\partial a_2}{\partial \beta} < 0,$$

thus high-powered incentives in one task adversely affect the other task.

Now it is intuitive that if the second task is sufficiently important for the principal, then she will "shy away" from high-powered incentives; if you are afraid that the contractor will sacrifice quality for speed, you are unlikely to offer a contract that puts a high reward on speed.

More formally, with a similar analysis to before, it can be shown that in this case

$$\beta^{**} = \frac{\phi_1 - \phi_2 \left(\partial^2 c\left(a_1, a_2\right) / \partial a_1 \partial a_2\right) / \left(\partial^2 c\left(a_1, a_2\right) / \partial a_2^2\right)}{1 + r\sigma_1^2 \left(\partial^2 c\left(a_1, a_2\right) / \partial a_1^2 - \left(\partial^2 c\left(a_1, a_2\right) / \partial a_1 \partial a_2\right)^2 / \partial^2 c\left(a_1, a_2\right) / \partial a_2^2\right)}$$

Therefore, the optimal linear contract from the point of view of the principal has sensitivity β^{**} to performance, and β^{**} is declining in ϕ_2 (the importance of the second task) and in $-\partial^2 c\left(a_1, a_2\right) / \partial a_1 \partial a_2$ (degree of substitutability between the efforts of the two tasks).

This equation is the basis of many of the claims based on the multitask model. In particular, think of the loose claim that "in the presence of multitasking high-powered incentives may be some optimal". Basically, this equation shows why.

This model has enormous potential to explain why many organizations are unwilling to go to high-powered incentives.

These ideas are developed in Holmstrom and Milgrom's 1991 paper "Multitask Principal-Agent Analyses" in the *Journal of Law Economics and Organization.* This is a fantastic paper, and you should read it.

They also show how the multitask idea explains why you want to put restrictions on the outside activities of workers or managers, and how it gives you a new perspective on thinking of how different tasks should be organized into various jobs.

5. Relative Performance Evaluation

This framework also naturally leads to relative performance evaluation when there are many agents working on similar tasks.

Let us go back to the one task linear model where

$$U(s,a) = -\exp(-r(s-c(a)))$$

with $c(a) = ca^2/2$, and the principal cares about

$$x = a + \varepsilon$$

The only difference now is that there is another worker (perhaps working for some other principal), whose performance is given by

$$\tilde{x} = \tilde{a} + \tilde{\varepsilon},$$

where $\tilde{}$ denotes the other worker. The random shocks ε and $\tilde{\varepsilon}$ are both normally distributed. They can also be correlated, which will play an important role in relative performance evaluation.

Assume that $\tilde{x}$ is publically observed. In equilibrium, everybody will guess the level of effort that this other worker exerts given his contract, so $\tilde{x}$, along the equilibrium path, will reveal $\tilde{\varepsilon}$. (This is a very important comment; it may be obvious to you, or it may not be; in either case think about it, and this will play a very important role in what follows).

Now if ε and $\tilde{\varepsilon}$ are uncorrelated, the equilibrium derived above applies.

But suppose that these two agents are in the same line of business, thus are affected by *common shocks*. Then we may assume, for example, that

$$\text{Var}(\varepsilon) = \text{Var}(\tilde{\varepsilon}) = \sigma^2$$

$$\text{Corr}\,(\varepsilon, \tilde{\varepsilon}) = \rho,$$

which determines whether both agents are simultaneously lucky or not.

In this case, using the same argument as before, it can be shown that the optimal (linear) contract for our agent will take the form

$$s = \alpha + \beta x - \tilde{\beta}\tilde{x}$$

with

$$\beta = \frac{1}{1 + rc\sigma^2\,(1-\rho)}$$

and

$$\tilde{\beta} = \frac{\rho}{1 + rc\sigma^2\,(1-\rho)}$$

Let us now consider the case where $\rho > 0$ so that performance between the two agents is positively correlated. In this case, the agent's payment is more sensitive to his own performance (β is now larger), but he will be punished for the successful performance of the other agent (and the extent of of this depends on the degree of correlation between the two performances, ρ). This is clearly a form of *relative performance evaluation*, where the agent is judged not according to some absolute standard but with respect to a relative standard set by others in the same field.

Can you see what would happen if $\rho < 0$?

6. Tournaments

Something akin to relative performance evaluation, some form of a "yardstick competition", where employees are compared to each other, often occurs inside firms. For example, the employee who is most successful gets promoted. Also related is the very common "up-or-out contracts" where after a while employees are either promoted or fired (e.g., tenure in academic systems). The parallel between these contracts and relative performance evaluation comes from the fact that it is typically impossible to promote all low-level workers, so there is an implicit element of yardstick competition in up-or-out contracts.

This situation is sometimes referred to as "tournaments".

The analysis is developed in the famous paper by Ed Lazear and Sherwin Rosen, *JPE* 1981.

They analyze the problem of a firm employing two workers in a similar task, one producing x_1, the other producing x_2.

We know from the above analysis that the optimal contract that the principal can offer to these guys should make their renumeration a function of both x_1 and x_2.

Instead, Lazear and Rosen look at a non-optimal but intuitive contract where the agents' renumerations are a function of their "rank", exactly as in sports tournament, where the highest prize goes to the winner, etc.

More concretely, let us assume that both the principal and the agents are risk-neutral, and the output of each agent is given by

$$x_i = a_i + \theta_i$$

where a_i is effort and θ_i is a stochastic term.

Both agents have the same cost function for effort given by $c(a)$, which, as before, is increasing and convex as usual. Let us denote the reservation utility of both agents by $\bar{H}$ as before.

Clearly the first best will solve

$$\max_{a_i} \; x_i - c(a_i),$$

so will satisfy

$$c'\left(a^{fb}\right) = 1.$$

(Recall that there is no effort interaction, only interactions through stochastic elements).

Let us simplify the problem and look at the extreme case where θ_1 and θ_2 are independent, so what we are dealing with is not standard "relative performance evaluation". Thus let us assume that they are both drawn independently from a continuous distribution $F(\theta)$, with density $f(\theta)$. This is in very useful benchmark, especially from our analysis of the baseline Holmstrom model above we know quite a few things about the optimal contract in this case (what do we know?).

The principal is restricted to the following contract

$$w_i(x_1, x_2) = \begin{cases} \overline{w} & \text{if} \quad x_i > x_j \\ \underline{w} & \text{if} \quad x_i < x_j \\ \frac{1}{2}(\overline{w} + \underline{w}) & \text{if} \quad x_i = x_j \end{cases}$$

In other words, the principal only chooses two levels of payments, $\overline{w}$ for the more successful agent and $\underline{w}$ for the less successful agent.

There is a difference here from what we have studied so far, since now conditional on the contract offered by the principal, the two agents will be playing a game, since their effort choices will affect the other agent's payoff.

More specifically, the timing of moves is now given by

(1) The principal chooses $\overline{w}, \underline{w}$.
(2) Agents simultaneously choose a_1, a_2

Formally, this again corresponds to a dynamic game where the principal is like a Stackleberg leader. Since we have a dynamic game, we should look for the subgame perfect Nash equilibrium, that means *backward induction*. That is, we need to analyze the Nash equilibrium in the subgames between the agents, and then the optimal contract choice of the principal.

In other words, we first take each subgame characterized by a different choice of the contract $w(x_1, x_2)$, and find the Nash equilibrium $a_1^*(\overline{w}, \underline{w})$, $a_2^*(\overline{w}, \underline{w})$ of the two agents (why do different contracts corresponds to different subgames?). Then the principal will maximize expected profits by choosing $\overline{w}, \underline{w}$ given agents' reaction functions, $a_1^*(\overline{w}, \underline{w})$, $a_2^*(\overline{w}, \underline{w})$.

The key object will be the probability that one worker performs better than the other as a function of their efforts. Define

$$P_i(a_i, a_j) \equiv \text{Prob}\{x_i > x_j \mid a_i, a_j\}$$

Clearly

$$x_i = a_i + \theta_i \quad > \quad a_j + \theta_j = x_j$$

if and only if

$$\theta_i \quad > \quad a_j - a_i + \theta_j$$

Using this, we can derive:

$$\begin{aligned} P_i(a_i, a_j) &= \text{Prob}\{\theta_i > a_j - a_i + \theta_j \mid a_i, a_j\} \\ &= \int \text{Prob}\{\theta_i > a_j - a_i + \theta_j \mid \theta_j, a_i, a_j\} f(\theta_j) d\theta_j \\ &= \int [1 - F(a_j - a_i + \theta_j)] f(\theta_j) d\theta_j \end{aligned}$$

(using indefinite integrals to denote integration over the whole support).

Nash equilibrium in the subgame given the wage function $w(x_1, x_2)$, or simply $(\overline{w}, \underline{w})$, is defined as a pair of effort choices (a_1^*, a_2^*) such that

$$a_i^* \in \max_{a_i} P_i(a_i, a_j^*)\overline{w} + \left[1 - P_i(a_i, a_j^*)\right] \underline{w} - c(a_i)$$

The first-order condition for the Nash equilibrium for each agent is therefore given by

$$(\overline{w} - \underline{w})\frac{\partial P_i(a_i, a_j^*)}{\partial a_i} - c'(a_i) = 0$$

This equation is very intuitive: each agent will exert effort up to the point where the marginal gain, which is equal to the prize for success times the increase in the probability of success, is equal to marginal cost of exerting effort.

The solution to this first-order condition is the best response of agent i to $(\overline{w}, \underline{w}, a_j^*)$.

Since agents are risk-neutral, is possible to implement the first best here. Clearly, the first best involves both agents choosing the same level of effort, a_{fb} as shown above. Let us then look for a symmetric equilibrium implementing the first best:

$$a_i^* = a_j^* = a_{fb}$$

Now using the first-order condition of the agent, this implies that a_{fb} has to be a solution to:

$$(\overline{w} - \underline{w})\frac{\partial P_i(a_i, a_{fb})}{\partial a_i} - c'(a_{fb}) = 0$$

Since the first best effort level is defined by $c'(a_{fb}) = 1$, this is equivalent to

$$(\overline{w} - \underline{w})\frac{\partial P_i(a_i, a_{fb})}{\partial a_i} = 1.$$

Note that

$$\frac{\partial P_i(a_i, a_j^*)}{\partial a_i} = \int f(a_j^* - a_i + \theta_j) f(\theta_j) d\theta_j$$

Now using symmetry, i.e., $a_1 = a_2$, this equation becomes

$$\left.\frac{\partial P_i(a_i, a_{fb})}{\partial a_i}\right|_{a_i = a_{fb}} = \int f(\theta_j)^2 d\theta_j$$

or

$$(\overline{w} - \underline{w}) \int f(\theta_j)^2 d\theta_j = 1$$

This characterizes the constraint that the principal will face in choosing the optimal contract.

Expressed differently, the principal must set

$$\begin{aligned}
\overline{w} - \underline{w} &= \left[\int f(\theta_j)^2 d\theta_j\right]^{-1} \\
\overline{w} &= \underline{w} + \Delta,
\end{aligned}$$

where

$$\Delta \equiv \left[\int f(\theta_j)^2 d\theta_j\right]^{-1}.$$

This expression implies that in order to induce the agents to play this symmetric equilibrium, which will in turn lead to the first best, the principal has to induce the wage gap of at least Δ between the more and the less successful agent. The constraint facing the principal to ensure that both agents exert effort is the equivalent of the incentive compatibility constraint of the standard moral hazard models translated into this tournament context (Can you interpret this incentive compatibility constraint in greater detail?).

The principal also needs to satisfy the participation constraint; this will tie down the values of $\overline{w}$ and $\underline{w}$.

$$P_i(a_{fb}, a_{fb})\overline{w} + [1 - P_i(a_{fb}, a_{fb})]\,\underline{w} - c(a_{fb}) \geq \overline{H}$$

Symmetry ensures that $P_i(a_{fb}, a_{fb}) = 1/2$, that is,

$$\frac{1}{2}(\overline{w} + \underline{w}) \geq \overline{H} + c(a_{fb})$$

The principal's problem therefore boils down to:

$$\begin{array}{lcr} \min\limits_{\overline{w},\underline{w}} & \overline{w} + \underline{w} & \\ \text{s.t} & \overline{w} = \underline{w} + \Delta & \text{(IC)} \\ & \overline{w} + \underline{w} \geq 2(\overline{H} + c(a_{fb})) & \text{(PC)} \end{array}$$

This has the solutions

$$\overline{w} = \overline{H} + c(a_{fb}) + \frac{\Delta}{2}$$

$$\underline{w} = \overline{H} + c(a_{fb}) - \frac{\Delta}{2}$$

Therefore, in this case by using a simple tournament, the principal can induce an equilibrium in which both agents choose first-best effort.

Now you may wonder how this compares with what we have done so far?

Clearly the tournament is not an "optimal contract" in general (we have restricted the functional form of the contract severely). But it is implementing the first best. In fact, given our assumption that θ_1 and θ_2 are independent, the full analysis above based on Holmstrom's paper shows that the optimal contract for agent 1 *should be independent of* x_2 and vice versa. So what is happening?

The answer is that the environment here is simple enough that both the optimal contract a la Holmstrom, and the non-optimal contract a la Lazear-Rosen reach the first best. This would generally not be the case. However, tournaments still may be attractive because they are simple and one might hope they might be "more robust" to variations in the technology or information structure (even though, again, the exact meaning of "robust" is not entirely clear here).

Having said that, while "up-or-out" contracts are quite common, tournament are not as common within organizations. This may be because designing tournament among employees might create an adversarial environment, leading workers to sabotage their coworkers (Lazear) or some type of collusion among the agents (Mookherjee).

There is little careful empirical work on these issues, so a lot of potential here.

7. Application: CEO Pay

A major application of the ideas of agency theory is to the behavior and renumeration of managers.

Consistent with theory, it seems that agency considerations are important, but there are also many other issues to consider.

One easy way of thinking about the problem is to equate the principal with the shareholders of the firm (ignoring free rider problem among the shareholders), the agent with the CEO (ignoring other managers that also take important decisions). In practice, the relationship between the CEO and managers, and between shareholders and debtholders could be quite important in thinking of the right model.

Ignoring these more complex issues, we can map the CEO example into our model as:

$$\begin{array}{rcc} \text{“output” } x & = & \text{change in shareholder wealth} \\ \text{“wages” } s(x) & = & \text{salary, bonus, benefits, pensions, stock options,...} \end{array}$$

The basic question in the empirical literature is whether executive compensation is sensitive to the firm's stock market performance (i.e. whether $s'(x) > 0$) and the answer is a clear yes. CEOs that are more productive for their shareholders are paid much more (some much much much more), and are less likely to be fired.

Nevertheless, an influential paper by Jensen and Murphy in 1990 argued that $s'(x)$ is too low. They blamed regulations and social norms for it. They argued that if $s'(x)$ could be increased, performance would improve substantially.

Clearly $s'(x)$ increased over the 1990s, and there is a debate as to whether it is too high or too low, but it's clear that $s'(x)$ is positive. Again, there is little careful empirical work here.

The recent scandals also illustrate that high values of $s'(x)$ create problems similar to employees gaming the rewards discussed above. This is related to the side effects of high-powered incentives discussed above.

CEO pay data also give us an opportunity to investigate whether Relative Performance Evaluation is used in practice. We may naturally expect that there should be a lot of relative performance evaluation, since there are many common factors determining shareholder value for all the firms in a particular industry. Here the evidence is more mixed.

An early paper by Gibbons and Murphy, in *ILRR* 1990 finds some evidence for relative performance evaluation. However, it seems that the CEOs are being compared to the market rather than their own industry. More recent research, for example, Bertrand and Mullainathan find that there is not enough relative performance evaluation in general. In particular, they find that CEOs are rewarded for common shocks affecting their industry and they interpret this finding as evidence against standard moral hazard models. While this result is interesting, the interpretation may not be fully warranted. In particular, a more careful theoretical framework would be useful in interpreting such results. For example, almost all empirical work interprets the data is being generated from and model in which output is given by

$$x = a + \eta + \varepsilon$$

where η is a common shocks affecting not only one firm but all others in the industry. But this additive structure is clearly special. A more general model would be

$$x = g\left(a, \eta, \varepsilon\right),$$

which allows for effort to be more valuable in periods in which the common shock as high. In this case optimal contracts may compensate CEOs when there are positive, shocks not because they are suboptimal, but because this is necessary for optimal incentive provision. Even though this may appear not as simple as the additive structure, the fact that it is theoretically possible implies that is somewhat more careful combination of theoretical and empirical work may be fruitful. This is an area for future work.

8. The Basic Model of Career Concerns

In addition to multitasking another issue important in many settings, especially in the public sector or for politicians, but equally for managers, is that they are not simply remunerated for the current performance with wages, but their future prospects for promotion and employment depend on their current performance. This is referred to as "career concerns" following the seminal paper by Holmstrom "Managerial Incentive Schemes-A Dynamic Perspective" in *Essays in Economics and Management in the Honor of Lars Wahlbeck* 1982.

The issues here are very important theoretically, and also have practical importance. Eugene Fama in a paper in 1980 suggested that competition in market for managers might be sufficient to give them sufficient incentives without agency contracts. Perhaps more important, and anticipating the incomplete contracts, which we will discuss soon, it may be the case that the performance of the agent is "observable" so that the market knows about it and then decide whether to hire the agent or not accordingly, but is not easy to contract upon. This will naturally lead to career concern type models.

The original Holmstrom model is infinite horizon, and we will see an infinite horizon model next, but let us start with a 2-period model. This class of models are sometimes referred to as "signal jamming" models (e.g., by Fudenberg and Tirole) for reasons that will become clear soon.

Output produced is equal to

$$x_t = \underbrace{\eta}_{\text{ability}} + \underbrace{a_t}_{\text{effort}} + \underbrace{\varepsilon_t}_{\text{noise}} \quad t = 1, 2$$

which is only different from what we have seen so far because of the presence of the ability term η.

We go to the extreme case where there are no performance contracts.

Moreover, assume that

$$\varepsilon_t \sim \mathcal{N}(0, 1/h_\varepsilon)$$

where h is referred to as "precision".

Also, the prior on η has a normal distribution with mean m_0, i.e.,

$$\eta \sim \mathcal{N}(m_0, 1/h_0)$$

and $\eta, \varepsilon_1, \varepsilon_2$ are independent.

As before, $a_t \in [0, \infty)$. Even without a_t, a dynamic model of this sort has a lot of interesting features (for example, this is analyzed in the dynamic wage contract model of Harris and Holmstrom).

Differently from the basic moral hazard model this is an equilibrium model, in the sense that there are other firms out there who can hire this agent. This is the source of the career concerns. Loosely speaking, a higher perception of the market about the ability of the agent, η, will translate into higher wages.

The name signal jamming now makes sense; it originates from the fact that under certain circumstances the agent might have an interest in working harder in order to improve the perception of the market about his ability.

Given these issues, let us be more specific about the information structure. This is as follows:

- the firm, the worker, and the market all share prior belief about η (thus there is no asymmetric information and adverse selection; is this important?).
- they all observe x_t each period.
- only worker sees a_t (moral hazard/hidden action).

In equilibrium firm and market correctly conjecture a_t. This is important from a technical point of view, because along-the-equilibrium path despite the fact that there is hidden action, information will stay symmetric.

The model of the labor market is simple. It is competitive, which does not introduce any difficult technicalities since all firms have symmetric information, and the other important assumption is that contracts cannot be contingent on output and wages are paid at the beginning of each period.

In particular, competition in the labor market implies that the wage of the worker at a time t is equal to the mathematical expectation of the output he will produce given the history of its outputs

$$w_t(x^{t-1}) = \mathbb{E}(x_t \mid x^{t-1})$$

where $x^{t-1} = \{x_1, ..., x_{t-1}\}$ is the history of his output realizations.

Of course, we can write this as

$$\begin{aligned} w_t(x^{t-1}) &= \mathbb{E}(x_t \mid x^{t-1}) \\ &= \mathbb{E}(\eta \mid x^{t-1}) + a_t(x^{t-1}) \end{aligned}$$

where $a_t(x^{t-1})$ is the effort that the agent will exert given history x^{t-1}, which is perfectly anticipated by the market along the equilibrium path.

Preferences are as before. In particular, the instantaneous utility function of the agent is

$$u(w_t, a_t) = w_t - c(a_t)$$

But we live in a dynamic world, so the agent maximizes:

$$U(w, a) = \sum_{t=1}^{T} \beta^{t-1} \left[w_t - c(a_t)\right]$$

where β is the agent's discount factor and T is the length of the horizon, which equals to 2 here (later we will discuss the case where $T = \infty$).

We do not need to take a position on where β comes from (the market or just discounting). It suffices that $\beta \leq 1$.

We also have the standard assumptions on the cost function for effort

$$\begin{aligned} &c' > 0, \qquad c'' > 0 \\ &c'(0) = 0 \end{aligned}$$

to guarantee a unique interior solution. This first best level of effort a^{fb} solves $c'(a^{fb}) = 1$.

Recall that all players, including the agent himself, have prior on $\eta \sim \mathcal{N}(m_0, 1/h_0)$

So the world can be summarized as:

$$\text{period 1: } \begin{cases} \text{wage } w_1 \\ \text{effort } a_1 \text{ chosen by the agent (unobserved)} \\ \text{output is realized } x_1 = \eta + a_1 + \varepsilon_1 \end{cases}$$

$$\text{period 2: } \begin{cases} \text{wage } w_2(x_1) \\ \text{effort } a_2 \text{ chosen} \\ \text{output is realized } x_2 = \eta + a_2 + \varepsilon_2 \end{cases}$$

The appropriate equilibrium concept is again Perfect Bayesian Equilibrium, but for our purposes what matters is that there will be backward induction again, and all beliefs will be pinned down by application of Bayes' rule. So let us start from the second period.

Backward induction immediately makes it clear that $a_2^* = 0$ irrespective of what happens in the first period, i.e., the agent will exert no effort in the last period because the wage does not depend on second period output, and the world ends after that, certifications don't matter.

Given this, we can write:

$$\begin{aligned} w_2(x_1) &= \mathbb{E}(\eta \mid x_1) + a_2(x_1) \\ &= \mathbb{E}(\eta \mid x_1) \end{aligned}$$

Then the problem of the market is the estimation of η given information $x_1 = \eta + a_1 + \varepsilon_1$. The only difficulty is that x_1 depends on first period effort.

In a Perfect Bayesian Equilibrium, the market will anticipate the level of effort a_1, and given the beliefs, agents will in fact play exactly this level. Let the conjecture of the market be $\bar{a}_1$.

Define

$$z_1 \equiv x_1 - \bar{a}_1 = \eta + \varepsilon_1$$

as the deviation of observed output from this conjecture.

Once we have z_1, life is straightforward because everything is normal. In particular, standard normal updating formula implies that

$$\eta \mid z_1 \sim \mathcal{N}\left(\frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon}, h_0 + h_\varepsilon\right)$$

The interpretation of this equation is straightforward, especially with the analogy to linear regression. Intuitively, we start with prior m_0, and update η according to the information contained in z_1. How much weight we give to this new information depends on its precision relative to the precision of the prior. The greater its h_ε relative to h_0, the more the new information matters. Finally, the variance of this posterior will be less than the variance of both the prior and the new information, since these two bits of information are being combined (hence its precision is greater).

Therefore, we have

$$\mathbb{E}(\eta \mid z_1) = \frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon}$$

or going back to the original notation:

$$\mathbb{E}(\eta \mid x_1) = \frac{h_0 m_0 + h_\varepsilon (x_1 - \bar{a}_1)}{h_0 + h_\varepsilon}$$

Consequently

$$w_2(x_1) = \frac{h_0 m_0 + h_\varepsilon (x_1 - \bar{a}_1)}{h_0 + h_\varepsilon}$$

So to complete the characterization of equilibrium we have to find the level of a_1 that the agent will choose as a function of $\bar{a}_1$, and make sure that this is indeed equal to $\bar{a}_1$, that is, this will ensure that this is a fixed point, as required by our concept of Perfect Bayesian Equilibrium.

Let us first write the optimization problem of the agent. This is

$$\max_{a_1} [w_1 - c(a_1)] + \beta [\mathbb{E}\{w_2(x_1) \mid \bar{a}_1\}]$$

where we have used the fact that $a_2 = 0$. Substituting from above and dropping w_1 which is just a constant, this is equivalent to:

$$\max_{a_1} \beta\, \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon (x_1 - \bar{a}_1)}{h_0 + h_\varepsilon} \middle| \bar{a}_1 \right\} - c(a_1)$$

Recall that both η and ε_1 are uncertain, even to the agent.

Therefore

$$\max_{a_1} \beta \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon (\eta + \varepsilon_1 + a_1 - \bar{a}_1)}{h_0 + h_\varepsilon} \middle| a_1 \right\} - c(a_1)$$

and finally making use of the fact that a_1 is not stochastic (the agent is choosing it, so he knows what it is!), the problem is

$$\max_{a_1} \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} a_1 - c(a_1) + \beta\, \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon(\eta + \varepsilon_1 - \bar{a}_1)}{h_0 + h_\varepsilon} \right\}$$

Now carrying out the maximization problem, we obtain the first-order condition:

$$c'(a_1^*) = \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} < 1 = c'(a_{fb})$$

so that the agent exerts *less* than first best effort in period one. This is because there are two "leakages" (increases in output that the agent does not capture): first, the payoff from higher effort only occurs next period, therefore its value is discounted to β. Secondly, the agent only gets credit for a fraction $h_\varepsilon/(h_0 + h_\varepsilon)$ of her effort, the part that is attributed to ability.

Holmstrom shows that as long as $\beta < 1$, equilibrium effort will also be less than the first-best in a stationary infinite horizon model, but as we will see next, with finite horizon or non-stationary environments, "over-effort" is a possibility.

The characterization of the equilibrium is completed by imposing $\bar{a}_1 = a_1^*$, which enables us to compute w_1. Recall that

$$\begin{aligned} w_1 &= \mathbb{E}(y_1 \mid \text{ prior}) \\ &= \mathbb{E}(\eta) + \bar{a}_1 \\ &= m_0 + a_1^* \end{aligned}$$

The model has straightforward comparative statics. In particular, we have:

$$\frac{\partial a_1^*}{\partial \beta} > 0$$

$$\frac{\partial a_1^*}{\partial h_\varepsilon} > 0$$

$$\frac{\partial a_1^*}{\partial h_0} < 0$$

These are all intuitive. Greater β means that the agent discounts the future less, so exerts more effort because the first source of leakage is reduced.

More interestingly, a greater h_ε implies that there is less variability in the random component of performance. This, from the normal updating formula, implies that any given increase in performance is more likely to be attributed to ability, so the agent is more tempted to jam the signal by exerting more effort. Naturally, in equilibrium, nobody is fooled, but equilibrium is only consistent with a higher level of equilibrium effort.

The intuition for the negative effect of h_0 is similar. When there is more variability in ability, career concerns are stronger.

This model gives a number of insights about what type of professions might have good incentives coming from career concerns. For example, if we think that ability matters a lot and shows a lot of variability in politics, the model would suggest that career concerns should be important for politicians.

9. Career Concerns Over Multiple Periods

Let us briefly emphasize one implication of having multiple periods in this setting. There will be more learning earlier on than later.

To illustrate this, let us look at the same model with three periods. This model can be summarized by the following matrix

$$\begin{array}{ll} w_1 & a_1^* \\ w_2(x_1) & a_2^* \\ w_3(x_1, x_2) & a_3^* \end{array}$$

With similar analysis to before, the first-order conditions for the agent are

$$c'(a_1^*) = \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} + \beta^2 \frac{h_\varepsilon}{h_0 + 2h_\varepsilon}$$

$$c'(a_2^*) = \beta \frac{h_\varepsilon}{h_0 + 2h_\varepsilon}$$

This immediately implies that

$$a_1^* > a_2^* > a_3^* = 0.$$

More generally, in the T period model, the relevant first-order condition is

$$c'(a_t^*) = \sum_{\tau=t}^{T-1} \beta^{\tau-t+1} \frac{h_\varepsilon}{h_0 + \tau h_\varepsilon}.$$

Holmstrom shows that in this case, with T sufficiently large, there exists a period $\overline{\tau}$ such that

$$a^*_{t<\overline{\tau}} \geq a_{fb} \geq a^*_{t>\overline{\tau}}.$$

In other words, workers work too hard when young and not hard enough when old—think of the working hours of assistant professors versus tenured faculty). Importantly and interestingly, these effort levels depend on the horizon (time periods), but not on past realizations.

Remarkably, similar results hold when ability is not constant, but evolves over time (as long as it follows a normal process). For example, we could have

$$\eta_t = \eta_{t-1} + \delta_t$$

with

$$\begin{array}{rcl} \eta_0 & \sim & \mathcal{N}(m_0, 1/h_0) \\ \delta_t & \sim & \mathcal{N}(0, 1/h_\delta)\ \forall\, t \end{array}$$

In this case, it can be shown that the updating process is stable, so that the process and therefore the effort level converge, and in particular as $t \to \infty$, we have

$$a_t \to \overline{a}$$

but as long as $\beta < 1$, $\overline{a} < a_{fb}$.

10. Career Concerns and Multitasking: Application to Teaching

Acemoglu, Kremer and Mian investigate a dynamic model of incentives with career concerns and multitasking, motivated by the example of teachers, and use this model to discuss which tasks should be organized in markets, firms or governments.

Here is a quick overview, which will be useful in getting us to work more with infinite-horizon career concerns models.

Consider an infinite horizon economy with n infinitely lived teachers, and $n' > n$ parents in every period, each with one child to be educated. $K = 1, 2, ...$ children can be taught jointly by K teachers. Each teacher, i, is endowed with a teaching ability a_t^i at the beginning of period t. The level of a_t^i is unknown, but both teacher i and parents share the same belief about the distribution of a_t^i. The common belief about teacher i's ability at time t is given by a normal distribution:

$$a_t^i \backsim \mathcal{N}(m_t^i, v_t),$$

(where, note that, following the article we will now use variances, rather than precision, so v_t is the variance, where as precision would have been $1/v_t$).

Ability evolves over time according to the stochastic process given by:

$$a_{t+1}^i = a_t^i + \varepsilon_t^i, \tag{5.4}$$

where ε_t^i is i.i.d. with

$$\varepsilon \backsim \mathcal{N}(0, \sigma_\varepsilon^2).$$

A teacher can exert two types of effort, "good" and "bad", denoted by g_t^i and b_t^i respectively.

The human capital, h_t^j of child j is given by:

$$h_t^j = \overline{a}_t^j + \overline{f(g_t)}^j \tag{5.5}$$

where $\overline{a}_t^j = \frac{1}{K^j} \sum\limits_{i \in \mathbb{K}^j} a_t^i$ and $\overline{f(g_t)}^j = \frac{1}{K^j} \sum\limits_{i \in \mathbb{K}^j} f(g_t^i)$ with $\mathbb{K}^j$ is the set of teachers teaching child j, and K^j as the number of teachers in the set $\mathbb{K}^j$. In addition, $f(g)$ is increasing and strictly concave in g, with $f(0) = 0$, and $h_t^j = 0$ if the child is not taught by a teacher.

Let us start with the case where each child is taught by a single teacher, in which case (5.5) specializes to

$$h_t^i = a_t^i + f(g_t^i), \tag{5.6}$$

where, in this case, we can index the child taught by teacher i by i.

Parents only care about the level of human capital provided to their children. The expected utility of a parent at time t is given by:

$$U_t^P = \mathbb{E}_t \left[h_t \right] - w_t,$$

where $\mathbb{E}_t[\cdot]$ denotes expectations with respect to publicly available information at the beginning of time t and w is the wage paid to the teacher.

The expected utility of a teacher i at time t is given by:

$$U_t^i = \mathbb{E}_t\left[\sum_{\tau=0}^{\infty} \delta^\tau \left(w_{t+\tau}^i - g_{t+\tau}^i - b_{t+\tau}^i\right)\right],$$

where $w_{t+\tau}^i$ denotes the wage of the teacher at time $t+\tau$, and $\delta < 1$ is the discount rate.

The level of h_t^i provided by a teacher is not observable to parents. Instead, parents have to rely on an imperfect signal of h, given by the test scores, s. The test score of child j in the general case is given by:

$$s_t^j = h_t^j + \gamma \overline{f(b_t)}^j + \overline{\theta}_t^j + \eta_t, \tag{5.7}$$

where $\gamma \geq 0$, θ_t^i is an i.i.d. student-level shock distributed as $\mathcal{N}(0, \sigma_\theta^2)$, for example, the ability of the students to learn, and η_t is a common shock that *every* teacher receives in period t. For example, if all students are given the same test, η_t can be thought of as the overall difficulty of the test, or any other cohort-specific difference in ability or the curriculum. η_t is distributed i.i.d. and $N(0, \sigma_\eta^2)$. In addition, $\overline{f(b_t)}^j$ and $\overline{\theta}_t^j$ are defined analogously as averages over the set of teachers in $\mathbb{K}^j$.

In the special case where each child is taught by a single teacher, we have:

$$s_t^i = h_t^i + \gamma f(b_t^i) + \theta_t^i + \eta_t. \tag{5.8}$$

Naturally the variance σ_θ^2 measures the quality of signal s_t^i, but the variance of the common shock, σ_η^2, also affects the informativeness of the signal.

The timing of events is similar to the baseline career concern model. In the beginning of every period t, parents form priors, m_t^i, on the abilities of teachers based on the histories of test scores of the teachers. They then offer a wage w_t^i based on the expected ability of the teacher working with their child. The teacher then decides on the levels of good and bad effort, and h and s are realized at the end of period t. Ability a_t^i is then updated according to the stochastic process (5.4). The process then repeats itself in period $t+1$.

Again we are interested in Perfect Bayesian Equilibria, where all teachers choose $\left\{g^i_{t+\tau}, b^i_{t+\tau}\right\}_{\tau=0,1,..}$ optimally given their rewards, and the beliefs about teacher ability are given by Bayesian updating.

Let us also simplify the analysis by focusing on the stationary equilibrium where the variance of each teacher's ability is constant, i.e. $v_t = v_{t+1} = v$, and there are many teachers, so $n \to \infty$.

We have to start by deriving the equations for the evolution of beliefs.

Parents' belief about teacher i at the beginning of period t can be summarized as, $a^i_t \backsim \mathcal{N}(m^i_t, v_t)$.

Let $S_t = [s^1_t \ ...... \ s^n_t]^T$ denote the vector of n test scores that the agents observe during period t when each child is taught by a single teacher.

As in the analysis above, parents back out the part of S_t which only reflects the ability levels of the teachers, plus the noise. Let $Z_t = [z^1_t \ ...... \ z^n_t]^T$ denote this backed out signal, where

$$\begin{aligned} z^i_t &= s^i_t - f(g^i_t) - \gamma f(b^i_t) \\ &= a^i_t + \theta^i_t + \eta_t \end{aligned}$$

Let a^i_{t+1} be the updated prior on teacher i's ability conditional on observing Z_t. Then the normality of the error terms and the additive structure in equation (5.8) imply that

$$a^i_{t+1} \backsim \mathcal{N}(m^i_{t+1}, v_{t+1})$$

where m^i_{t+1} and v_{t+1} denote the mean and the variance of the posterior distribution. Using the normal updating formula, setting $v_{t+1} = v_t = v$, it can be derived that:

$$m^i_{t+1} = m^i_t + \beta(z^i_t - m^i_t) - \overline{\beta}(\overline{z}^{-i}_t - \overline{m}^{-i}_t), \tag{5.9}$$

where

$$\beta = \overline{\beta} = \frac{1 + \sqrt{1 + 4\left(\frac{\sigma^2_\theta}{\sigma^2_\varepsilon}\right)}}{1 + 2\left(\frac{\sigma^2_\theta}{\sigma^2_\varepsilon}\right) + \sqrt{1 + 4\left(\frac{\sigma^2_\theta}{\sigma^2_\varepsilon}\right)}}, \tag{5.10}$$

z_t^i is the ith element of the vector Z_t, and refers to the signal from teacher i, while $\overline{z}_t^{-i}$ is the average test score excluding teacher i. Since $n \to \infty$, we have $(\overline{z}_t^{-i} - \overline{m}_t^{-i}) \to \eta_t$, so the common shock is revealed and filtered out.

This expression indicates that we can think of the parameter β as in "career concerns" parameter, in the sense that it indicates how much a given increase in test scores of children feeds into an improved perception of the ability of the teacher.

Note also that there is a natural form of relative performance evaluation here because of the common shock η_t—by comparing two different teachers (schools), the common shock η_t can be perfectly filtered out.

Let us next look at the first and second-best by considering the social welfare function:

$$U_t^W = \sum_{\tau=0}^{\infty} \delta^\tau (\overline{A} + f(g_{t+\tau}) - g_{t+\tau} - b_{t+\tau}) \tag{5.11}$$

where $\overline{A}$ is the average ability of teachers in the population, which is constant when $n \to \infty$, and $g_{t+\tau}$ and $b_{t+\tau}$ are the good and bad effort levels chosen by all teachers.

Naturally we have:

First Best: Maximizing (5.11) gives us the first-best. In the first-best, there is no bad effort, $b_t = 0$, and the level of good effort, g^{FB}, is given by $f'(g^{FB}) = 1$.

Second-Best: Since teacher effort and the level of human capital are not directly observable, a more useful benchmark is given by solving for the optimal mechanism given these informational constraints.

Let $\Omega_t^i = [m_0^i \ s_0^i \ s_1^i \ s_2^i \ ...... \ s_{t-1}^i]$ be the information set containing the vector of test scores for teacher i at the beginning of period t when all children are taught by a single teacher.

Let $w_t^i(\Omega_t^i)$ be the wage paid to teacher i in period t. Then the constrained maximization problem to determine the second-best allocation can be written as:

$$\max_{\left\{w_{t+\tau}^i(\Omega_{t+\tau}^i)\right\}_{\tau=0,1,..}} U_t^W \text{ subject to}$$

$$\{g_{t+\tau}\ ,\ b_{t+\tau}\}_{\tau=0,1,..} \in \arg \max_{\left\{g'_{t+\tau}\ ,\ b'_{t+\tau}\right\}_{\tau=0,1,..}} \mathbb{E}_t \left[\sum_{\tau=0}^{\infty} \delta^\tau (w_{t+\tau}(\Omega_{t+\tau}^i) - g'_{t+\tau} - b'_{t+\tau})\right].$$

While the exact solution of this problem is slightly involved, the first-order condition immediately implies that:

$$\gamma f'(b_{t+\tau}) = f'(g_{t+\tau})$$

Therefore, teachers can be encouraged to exert good effort only at the cost of bad effort. As a result, the opportunity cost of inducing high effort is greater in the second-best problem than in the first-best.

Next consider a wage schedule of the form

$$w_t = \alpha m_t + \kappa,$$

which links teacher compensation to their contemporaneous perceived ability.

Given such a schedule, the privately optimal levels of good and bad effort are obtained as:

$$f'(g_{t+\tau}) = \gamma f'(b_{t+\tau}) = \frac{1 - \delta(1-\beta)}{\alpha\delta\beta}$$

for all $\tau \geq 0$.

Consequently, a greater α, i.e., higher-powered incentives, translate into greater good and bad effort, and for intuitive reasons, the magnitude of this effect depends both on the career concerns coefficient β and the discount factor δ. (Can you develop the intuition?)

Putting this together with the second-best above, we immediately see that

$$\alpha^{SB} = \frac{1 - \delta(1-\beta)}{\delta\beta f'(g^{SB})}, \tag{5.12}$$

would achieve the second-best for given level of β.

Interestingly, somehow if α was constant, but the planner could manipulate β, that is the degree to which teachers have "career concerns", the second-best could be achieved by setting:

$$\beta_{SB} = \frac{1-\delta}{\delta(\alpha f'(g^{SB})-1)}. \tag{5.13}$$

Now it is an immediate corollary of what we have seen so far that if all teachers work in "singleton teams," that is, if they work by themselves, the market wage for teacher i will be:

$$w_t^i = m_t^i + \mathbb{E}_t[f(g_t^i)]. \tag{5.14}$$

The market equilibrium is therefore similar to the second-best equilibrium, except that now α is fixed to be 1. This leads to a result that parallels the possibility of excess incentives in the multitask models.

In particular, the market equilibrium level of good effort will be g^M , given by:

$$f'\left(g^M\right) = \frac{1-\delta(1-\beta)}{\delta\beta}.$$

An interesting implication is that $g^M < g^{SB}$ if $\gamma < \underline{\gamma}$, and $g^M > g^{SB}$ if $\gamma > \underline{\gamma}$.

The result that $g^M < g^{SB}$ if $\gamma < \underline{\gamma}$ is similar to the result in Holmstrom discussed above that with discounting, career concerns are typically insufficient to induce the optimal level of effort. So in this case, even markets do not provide strong enough incentives.

The case where $\gamma > \underline{\gamma}$, on the other hand, leads to the opposite conclusion. Now the natural career concerns provided by the market equilibrium create too high-powered incentives relative to the second-best.

The extent to which the market provides excessively high-powered incentives depends on the career concerns coefficient, β, and via this, on σ_θ^2 and σ_ε^2. When σ_θ^2 is small relative to σ_ε^2, β is high, and teachers in the market care a lot about their pupils scores, giving them very high-powered incentives. In this case, since markets are encouraging too much bad effort, firms or governments may be useful by modifying the organization of production to dull incentives.

If indeed markets provide too high-powered incentives, one way of overcoming this may be to form teams of teachers to weaken the signaling ability of individual teachers.

Let us model the firm as a partnership of K teachers working together, engaged in joint teaching as captured in equation (5.5) above.

Crucially, parents only observe the aggregate or average test score of all the teachers (or pupils) in the firm.

The notation for the analysis in this case is somewhat involved, so we will not provided details, but simply highlight the main result. This is that when $\gamma > \underline{\gamma}$ so that markets provide excessive incentives, there exists a unique equilibrium where firms have size equal to $K^* = \beta/\beta_{SB} > 1$ and where teachers exert the second-best level of good effort, g^{SB}.

The paper also shows why this beneficial firm-equilibrium may be impossible to sustain because of inside information about the performance of employees within the firm (think about why such inside information will be problematic?), and how government-type organizations with dollar incentives may be useful (what are the things that the government can do and the private sector can not?).

10.1. Team Production. Finally, let us briefly discuss the Holmstrom 1982 paper where output is produced by a team, where every worker's contribution raises the total output of the firm. This seems like a good approximation to many production processes in practice.

The information structure is such that only total output is observed, that is the principal cannot tell the contribution of different workers to total production.

Given this assumption, the environment can be simplified by first removing uncertainty, because there is still a non-trivial problem for the principal, since she cannot invert the output-effort relationship to obtain the actions of all agents (since all of their efforts matter for output).

More formally, consider an organization consisting of n agents $\quad i \in \{1, \ldots, n\}$

They all choose effort $a_i \in [0, \infty)$

Let the vector of efforts be denoted by

$$a = (a_1, \ldots, a_n)$$

The key assumption is that of team production, so output is equal to

$$x = x(a_1, \ldots, a_n) \in \mathbb{R}$$

and does not depend on the stochastic variable θ. We make the natural assumption that higher effort leads to higher output, that is,

$$\frac{\partial x}{\partial a_i} > 0 \quad \forall\, i$$

All of the workers have risk neutral preferences:

$$U(w_i, a_i) = w_i - c_i(a_i)$$

with the usual assumption that, $c_i' > 0,\ c_i'' > 0 \quad \forall\, i$

What is a contract here?

Since only x is observable, a contract has to be a factor that specifies payments to each agent as a function of the realization of output.

Let us refer to this as a sharing rule, denoted by

$$s(x) = (s_1(x), ..., \underbrace{s_i(x)}_{\text{payment to agent } i \text{ when team output is } x}, \ldots, s_n(x))$$

It is natural to impose limited liability in this case, so

$$s_i(x) \geq 0 \quad \forall\, x, \forall\, i$$

Moreover, we may want to impose that the firm can never payout more than what it generates

$$\sum_{i=1}^{n} s_i(x) \leq x$$

(though this may be relaxed if the firm is represented by a risk-taking entrepreneur, who makes a loss in some periods and compensated by gains during other times).

The timing of events is as usual:

(1) Principal and agents sign $s(x)$

(2) n agents simultaneously choose efforts.

(3) Everybody observes x

(4) The payments specified by $s(x)$ are distributed and the principal keeps $x - \sum_{i=1}^{n} s_i(x)$

Before we analyze this game, let us imagine that there is no principal and the team manages itself, in the spirit of a labor-managed firm.

How should the labor-managed firm ideally set $s(x)$?

The key constraint is that of *budget balance*, i.e., the labor managed firm has to distribute all of the output between its employees (there is no principal to make additional payments or take a share of profits; money-burning type rules would be ex post non-credible). Thus, we have

$$\sum_{i=1}^{n} s_i(x) = x \qquad \forall\, x$$

Let us ask whether the labor-managed firm can achieve efficiency, that is effort levels such that

$$\frac{\partial x}{\partial a_i} = c_i'(a_i^*) \qquad \forall\, i \tag{5.15}$$

Thus the question is whether there exists a sharing rule $s(x)$ that achieves full efficiency.

To answer this question, we have to look at the first-order conditions of the agents, which take the natural form:

$$s_i'(x)\frac{\partial x}{\partial a_i} = c_i'(a_i)$$

Now for this condition to be consistent with (5.15), it must be that $s_i'(x) = 1 \quad \forall i$. But budget balance requires $\sum s_i'(x) = 1$ so we cannot have full efficiency.

One solution is to have a "budget breaker" so that the budget balance constraint is relaxed.

Consider the contract

$$s_i(x) = \begin{cases} b_i & \text{if} \quad x \geq x(a^*) \\ 0 & \text{if} \quad x < x(a^*) \end{cases}$$

where $a^* = (a_1^*, \ldots, a_n^*)$ is the vector of efficient effort levels and $\sum b_i = x(a^*)$. What is happening to the output when it is less than $x(a^*)$?

It can be verified easily that given this contract all agents will choose the efficient level of effort. If they do not, then they will all be punished severely.

In fact, with this contract, along the equilibrium path there is budget balance, but the principal can design a contract whereby off the equilibrium path, output is taken away from the workers as punishment.

Looked at in this light, the problem of the labor-managed firm (relative to the capitalist firm) is its inability to punish its employees by throwing away output.

10.2. Teams with Observed Individual Outputs. Let us end this discussion by going back to tournaments, relative performance evaluation and sufficient statistics. Let us consider the team production problem and investigate the role that the performance of other agents play in the optimal contract.

Let us assume that the principal is risk neutral while agents have utility

$$u_i(w_i) - c_i(a_i)$$

with the standard assumptions,

$$\begin{aligned} u_i' &> 0, \quad u_i'' < 0 \\ c_i' &> 0, \quad c_i'' > 0 \end{aligned}$$

In the baseline model, the output x_i of each agent is observed but not his effort level a_i. The output of the agent is again a function of his own effort and some state of nature θ_i

$$x_i(a_i, \theta_i) = \text{ output of } i$$

Notice that there is no "team production" here, since x_i only depends on the action of individual i, a_i.

As usual we assume that

$$\frac{\partial x_i}{\partial a_i} > 0,$$

and as a normalization,

$$\frac{\partial x_i}{\partial \theta_i} > 0$$

Finally, let us denote the vector of stochastic elements by

$$\theta = (\theta_1, \ldots, \theta_n) \sim F(\theta)$$

Possible sharing rules in the setup are vectors of the form

$$\{s_i(x_1, \ldots, x_i, \ldots, x_n)\}_{i=1}^n .$$

The general sharing rules are difficult to characterize. However, the sufficient statistic result from before enables us to answer an interesting question: when is a sharing rule that only depends on the individual agent's output $\{s_i(x_i)\}_{i=1}^n$ optimal?

To answer this, let us set up the principal's problem as

$$\max_{\substack{a=(a_1,\ldots,a_n) \\ \{s_1(x_1,\ldots,x_n)\}_{i=1}^n}} \int_\theta \left[\sum_{i=1}^n \{x_i(a_i, \theta_i) - s_i(x_1(a_1, \theta_1), \ldots, x_n(a_n, \theta_n))\} \right] dF(\theta)$$

subject to the participation constraint of each agent

$$\int_\theta u_i(s_i(x_1(a_1, \theta_1), \ldots, x_n(a_n, \theta_n)))dF(\theta) - c_i(a_i) \geq \overline{H}_i \text{ for all } i$$

and the incentive compatibility constraints (combined with Nash equilibrium in the tournament-like environments that we have already seen), that is,

$$a_i \in \max_{a'} \int_\theta u_i(s_i(x_1(a_1, \theta_1), \ldots, x_i(a_i', \theta_i), \ldots, x_n(a_n, \theta_n))dF(\theta) - c_i(a_i') \qquad \text{for all } i$$

Recall from our discussion above that for $n = 1$ we know x is a sufficient statistic for (x, y), if and only if $f(a \mid x, y) = f(a \mid x)$, which implies that the optimal contract $s(x, y) = S(x)$

Generalizing that analysis in a natural way, for $n > 1$, we are interested in the question: is x_i a sufficient statistic for $(x_1, \ldots, x_n)$ with respect to a_i?

Using the previous definition, let y now be a vector, defined as

$$y = x_{-i} = (x_1 \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

The sufficient statistic result says that

$$s_i(x_1, \ldots, x_n) = s_i(x_i) \text{ if and only if } f(a_i \mid x_i, x_{-i}) = f(a_i \mid x_i).$$

This leads to the natural result that the optimal sharing rules $\{s_i(x_i, \ldots, x_n)\}_{i=1}^n$ are functions of x_i alone iff the θ_i's are independent. i.e.,

$$F(\theta) = F_1(\theta_1)F_2(\theta_2)\ldots F_n(\theta_n).$$

What the proposition says is that forcing agents to compete with each other is *useless* if there exists no common uncertainty. This contrasts with the tournament results of Lazear and Rosen, and highlights that tournament-type contracts are generally not optimal.

Even when the different θs are not independent, the sharing rules might take simple forms.

For example, suppose that $\theta_i = \eta + \varepsilon_i$ so that individual uncertainty is the sum of an aggregate component and an individual component (independent across agents), and that both of these components are normally distributed

Then

$$\begin{aligned} x_i(a_i, \theta_i) &= a_i + \theta_i \\ &= a_i + \eta + \varepsilon_i \end{aligned}$$

In this case, it can be shown that optimal contracts are of the form

$$s_i(x_i, \overline{x})$$

where

$$\overline{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

is average output. This is because $(\overline{x}, x_i)$ is a sufficient statistic for $(x_1, \ldots, x_n)$ for the estimation of a_i.

11. Moral Hazard and Optimal Unemployment Insurance

Let us now consider an application of ideas related to moral hazard to the design of optimum unemployment insurance. The standard approach in the literature, first developed by Shavell and Weiss's classic paper in 1979, considers the problem of the design of optimal unemployment insurance as a dynamic moral hazard problem, where unemployed individuals have to exert effort to find jobs and the unemployment insurance system provides consumption insurance. Greater consumption insurance is desirable all else equal, but it tends to discourage search effort and thus increases unemployment duration.

Here I will present a slight generalization of Shavell and Weiss's approach based on a more recent paper by Hopenhayn and Nicolini (JPE, 1997). The interaction between a general equilibrium model of search and unemployment insurance is discussed in later chapters.

The model incorporates moral hazard regarding search effort (but there are no application decisions). Since the firm side is left implicit, it is essentially a partial equilibrium model. The preferences of the agent are

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^{t}\left[u(c_{t})-a_{t}\right]$$

where $c_t \in \mathbb{R}$ is consumption and $a_t \in A$ is search effort, which lead to a probability of finding a job $p_t = p(a_t)$. All jobs are homogeneous and pay w (the feature that rules out the application margin). We naturally assume that

$$p'' < 0,\ p' > 0.$$

We also assume that the individual has zero income when unemployed and does not have access to any savings or borrowing opportunities. This last assumption is crucial and simplifies the analysis by allowing the unemployment insurance authority to directly control the consumption level of the individual. Otherwise, there will be an additional constraint which determines the optimal consumption path of the individual.

Let s_j be state at time j

$$\begin{array}{lll} s_j = 0 & \rightarrow & \text{unemployed} \\ s_j = 1 & \rightarrow & \text{employed} \end{array}$$

The important object will be the history of the agent up to time t, which is denoted by $h_t = \{s_j\}_{j<t}$. Let $\mathcal{H}_t$ be this set of all such histories.

A general insurance contract can be represented as a mapping

$$\tau : \mathcal{H}_t \longrightarrow A \times \mathbb{R}$$

where the first element of the mapping is a_t, the "recommended search effort" and the second element z_t is the transfer to the worker, which will directly determine his consumption, since he has no access to an outside source of consumption and no savings opportunities.

Let $V_0(\tau)$ be the expected discounted utility at $t = 0$ associated with contract τ, and to prepare for setting up the dual of this problem, let $C_0(\tau)$ be the expected cost (net transfers) to the agent.

Now the optimal contract choice can be set up as

$$\max \ V_0(\tau)$$

s.t.

IC (incentive compatibility constraints) – if any

$$C_0(\tau) \leq \underline{C}$$

The last constraint for example may require the total cost to be equal to zero, i.e., all benefits to be financed by some type of payroll taxes or other taxation. E.g., budget balance as in the previous model.

Instead of this problem, we can look at the dual problem

$$\min C(V) = C_0(\tau)$$

s.t. IC

$$V_0(\tau) \geq V$$

Let us start with the full information case where the social planner (the unemployment insurance authority) can directly monitor the search effort of the unemployed individual, so the individual has no choice but to choose the recommended search effort. This implies that there are no IC constraints.

Then, it is straightforward that full insurance is optimal, i.e., $c_t = c \; \forall \; t$, and the level of search effort will solve:

$$a^* = \arg\max_a p(a) \sum_{t=0}^{\infty} \beta^t [1 - p(a)]^t \left[\frac{u(c)}{1-\beta} - a \right]$$

The more interesting case is the one with imperfect information, where a is the private information of the individual, so he will only follow the recommended search effort if this is incentive compatible for him. In other words, as in all types of implementation or optimal policy problems, there is an "argmax" constraint on the maximization problem.

Suppose $V_0(\tau) = V$. Let us introduce some useful notation

$$\begin{array}{ll} V^e = V_1(\tau) & \text{if } s_1 = 1 \\ V^u = V_1(\tau) & \text{if } s_1 = 0. \end{array}$$

This implies that we can write the value of the individual as

$$V = u(c) - a + \beta \left\{ p(a)V^e + (1 - p(a))V^u \right\}$$

Now the incentive compatibility constraints boil down to

$$\text{(IC)} \qquad a \in \arg\max_{a'} u(c) - a' + \beta \left\{ p(a')V^e + (1 - p(a'))V^u \right\}. \tag{5.16}$$

Naturally, (5.16) defines a very high dimensional object. It basically requires a to be better than or as good as any other feasible choice in A. These kinds of constraints are very difficult to work with, so the literature usually takes **the first-order approach**, which is to represent (5.16) with the corresponding first-order condition of the agent, i.e.,

$$\beta p'(a)(V^e - V^u) = 1 \tag{5.17}$$

This may seem innocuous, but in many situations it leads to the **wrong** solution. One has to be very careful in using the first-order approach. In this case, the situation is not so bad, because the individual only has a single choice, and given V^e and V^u, his maximization problem is strictly concave, so the first-order condition (5.17) is necessary and sufficient for the individual's maximization problem. Nevertheless, this constraint itself, i.e., (5.17), is non-linear and non-convex, so some of the difficulties of designing optimal contracts carry over to this case.

The problem is further simplified by noting that after the individual finds a job, there is no further incentive problem, so after that point there will be full consumption smoothing, i.e.,

$$V^e = \frac{u(c^e)}{1-\beta} \qquad \text{for some } c^e. \tag{5.18}$$

This is equivalent to a per-period transfer $c^e - w$ to the agent. In other words, there may be negative or positive transfers to the agent after he finds a job. The level of these transfers will be a function of its history, i.e., when (after how many periods of unemployment) he has found a job.

Now let

$$W(V^e) = \frac{c^e - w}{1-\beta}$$

be the discounted present value of the transfer from the principal to the agent. Inverting (5.18), we have

$$W(V^e) = \frac{-w + u^{-1}[(1-\beta)V^e]}{1-\beta}$$

Differentiating this equation, we obtain an intuitive formula

$$W'(V^e) = \frac{1}{u'(c^e)},$$

which states that the cost of providing greater utility is the reciprocal of the marginal utility of consumption for the individual. When $u'(c^e)$ is high, providing more utility to the individual is relatively cheap. From the concavity of the individual's utility function, u, W is also seen to be a convex function (it is clearly increasing).

Now let $C(V)$ be the cost of providing utility V to an unemployed individual. It can be written in a recursive form as

$$C(V) = \min_{a,c^u,V^e,V^u} c^u + \beta \{p(a)W(V^e) + [1-p(a)]C(V^u)\}$$

subject to

$$u(c^u) - a + \beta \{p(a)V^e + [1-p(a)]V^u\} = V \tag{5.19}$$

$$\beta p'(a)(V^e - V^u) = 1 \tag{5.20}$$

where c^u is utility given to unemployed individual, (5.19) is the promise keeping constraint, which makes sure that the agent indeed receives utility V. (5.20) is the IC constraint using the first-order approach. Note that this formulation makes it clear that the social planner or the unemployment insurance authority is directly controlling consumption. Otherwise, there would be another constraint corresponding to the Euler equation of the individual for example.

Also, notice that this is a standard recursive equation, so time has been dropped and everything has been written recursively. This creates quite a bit of economy in terms of notation. Moreover, the existence of a function $C(V)$ can be again guaranteed using the contraction mapping theorem (Theorem **??**).

An interesting question is whether $C(V)$ is convex. Recall that in the standard dynamic programming problems, concavity of the payoff function and the convexity of the constraints set were sufficient to establish concavity of the value function. Here we are dealing with a minimization problem, so the equivalent result would be convexity of the cost function. However, the constraint set is no longer convex, so the convexity of $C(V)$ is not guaranteed. This does not create a problem for the solution, but it implies that there may be a better policy than the one outlined above which would involve using lotteries.

Can you see why lotteries would improve the allocation in this case? Can you see how the problem should be formulated with lotteries?

Here, to simplify the analysis, let us ignore lotteries.

To make more progress, let us assign multiplier λ to (5.19) and η to (5.20). Then the first-order conditions (with respect to a, c^u, V^e and V^u) are

$$\beta p'(a)[W(V^e) - C(V^u)] - \lambda\left[\beta p'(a)(V^e - V^u) - 1\right] - \eta\beta p''(a)(V^e - V^u) = 0$$

$$1 - \lambda u'(c^u) = 0$$

$$\beta p(a)W'(V^e) - \lambda\beta p(a) - \eta\beta p'(a) = 0$$

$$\beta\left[1 - p(a)\right]C'(V^u) - \lambda\beta[1 - p(a)] + \eta\beta p'(a) = 0$$

The second first-order condition immediately implies

$$\lambda = 1/u'(c^u)$$

Now substituting this into the other conditions (and using constraint (5.20)), we have

$$p'(a)[W(V^e) - C(V^u)] = \eta p''(a)(V^e - V^u) \tag{5.21}$$

$$C'(V^u) = \frac{1}{u'(c^u)} - \eta\frac{p'(a)}{1 - p(a)} \tag{5.22}$$

$$W'(V^e) = \frac{1}{u'(c^e)} = \frac{1}{u'(c^u)} + \eta\frac{p'(a)}{p(a)} \tag{5.23}$$

In addition, we have the following envelope condition by differentiating the cost function with respect to V:

$$C'(V) = \frac{1}{u'(c^u)} = [1 - p(a)]C'(V^u) + p(a)W'(V^e) \tag{5.24}$$

We now have a key result of optimal unemployment insurance:

THEOREM 5.1. *The unemployment benefit and thus unemployed consumption, c^u, is decreasing over time. In addition, if $C(V)$ is convex, then $V^u < V$.*

PROOF. (sketch) From (5.22) and (5.23), we have that

$$W'(V^e) - C'(V^u) = \eta p'(a)\left[\frac{1}{1 - p(a)} + \frac{1}{p(a)}\right].$$

Since $\eta > 0$ (see the paper, or think intuitively), this immediately implies

$$W'(V^e) > C'(V^u)$$

Now use the Envelope condition (5.24), which immediately implies

$$W'(V^e) > C'(V) > C'(V^u) \tag{5.25}$$

Let $\hat{c}^u$ be next period's consumption. Then we have

$$C'(V^u) = \frac{1}{u'(\hat{c}^u)},$$

which combined with (5.25) and (5.24) and the concavity of the utility function u immediately implies

$$\hat{c}^u < c^u$$

as claimed. Moreover, (5.25) also implies that $V^u < V$ as long as C is convex, completing the proof of the theorem. □

What is the intuition? *Dynamic incentives:* the planner can give more efficient incentives by reducing consumption in the future.

A related question is what happens to the transfer/tax to employed workers. Is this a function of history?

THEOREM 5.2. *The wage tax/subsidy is a function of history, h_t, i.e., it is not constant.*

PROOF. (sketch) Let us revisit the envelope condition and rewrite it as

$$\begin{aligned} C'(V_t) &= [1 - p(a_t)]C'(V_{t+1}) + p(a_t)W'(V_t^e). \\ C'(V_t) &= \sum_{i=0}^{T-1} \left\{ \prod_{j=0}^{i-1} (1 - p(a_{t+j})) \right\} p(a_{t+i})W'(V_{t+i}^e) \\ &\quad + \left\{ \prod_{j=0}^{T-1} (1 - p(a_{t+j})) \right\} C'(V_{t+T}^u) \end{aligned}$$

Now to obtain a contradiction, suppose that $V_t^e = V^e$ for all t. From Theorem 5.1, V_t^u must eventually be decreasing (since consumption benefits are). Let the second term with $\prod_{j=0}^{T-1}$ be denoted by b_2. Since $C'(V_{t+T}^u)$ is bounded, so as $T \to \infty$, we have $b_2 \to 0$. Therefore,

$$C'(V_t) = \sum_{i=0}^{\infty} \left\{ \prod_{j=0}^{i-1} (1 - p(a_{t-1+j})) \right\} p(a_{t-1+i}) W'(V_{t+i}^e)$$

Since, by hypothesis, $W'(V_{t+i}^e)$ is constant, we have

$$\begin{aligned} C'(V_t) &= W'(V^e) \sum_{i=0}^{\infty} \left\{ \prod_{j=0}^{i-1} (1 - p(a_{t+j})) \right\} p(a_{t+i}) \\ &= W'(V^e), \end{aligned}$$

which contradicts (5.25), so V^e cannot be constant and therefore c^e cannot be constant. □

Under further assumptions, it can be established that generally c^e is a decreasing sequence, which implies that Optimal unemployment insurance schemes should make use of employment taxes conditional on history as well as allow for decreasing benefits.

Can you see the intuition for why wage taxes/subsidies are non-constant? Can you relate this result to decreasing benefits?

CHAPTER 6

Holdups, Incomplete Contracts and Investments

Before we discuss theories of investments in general and specific training, it is useful to review certain basic notions and models of holdups and investments in the absence of perfect markets and complete contracts. This chapter discusses the model by Grout where wage negotiations in the absence of a binding contract between workers and firms leads to underinvestment by firms, and the famous incomplete contract approach to the organization of the firm due to Williamson and Grossman and Hart.

1. Investments in the Absence of Binding Contracts

Consider the following simple setup. A firm and a worker are matched together, and because of labor market frictions, they cannot switch partners, so wages are determined by bargaining. As long as it employs the worker, the total output of the firm is

$$f(k)$$

where k is the amount of physical capital firm has, and f is an increasing, continuous and strictly concave production function.

The timing of events in this simple model is as follows:

- The firm decides how much to invest, at the cost rk.
- The worker and the firm bargain over the wage, w. We assume that bargaining can be represented by the *Nash solution* with asymmetric bargaining powers. In this bargaining problem, if there is disagreement, the worker receives an outside wage, $\bar{w}$, and the firm produces nothing, so its payoff is $-rk$.

The equilibrium has to be found by backward induction, starting in the second period. This involves first characterizing the Nash solution to bargaining. In this Nash bargaining problem, let the bargaining power of the worker be $\beta \in (0,1)$. Recall that the (asymmetric) Nash solution to bargaining between two players, 1 and 2, is given by maximizing

$$(\text{payoff}_1 - \text{outside option}_1)^{\beta} (\text{payoff}_2 - \text{outside option}_2)^{1-\beta}. \tag{6.1}$$

Digression: Before proceeding further, let us review where equation (6.1) comes from. Nash's bargaining theorem considers the bargaining problem of choosing a point x from a set $X \subset \mathbb{R}^N$ for some $N \geq 1$ by two parties with utility functions $u_1(x)$ and $u_2(x)$, such that if they cannot agree, they will obtain respective *disagreement payoffs* d_1 and d_2 (these are sometimes referred to as "outside options," though if they are literally modeled as outside options in a dynamic game of alternating offers bargaining, we would not necessarily in that with the Nash solution). The remarkable Nash bargaining theorem is as follows. Suppose we impose the following four axioms on the problem and solution: (1) $u_1(x)$ and $u_2(x)$ are Von Neumann-Morgenstern utility functions, in particular, unique up to positive linear transformations; (2) Pareto optimality, the agreement point will be along the frontier; (3) Independence of the Relevant Alternatives; suppose $X' \subset X$ and the choice when bargaining over the set X is $x' \in X'$, then x' is also the solution when bargaining over X'; (4) Symmetry; identities of the players do not matter, only their utility functions. Then, there exists a unique bargaining solution that satisfies these four axioms. This unique solution is given by

$$x^{NS} = \arg\max_{x \in X} (u_1(x) - d_1)(u_2(x) - d_2)$$

If we relax the symmetry axiom, so that the identities of the players can matter (e.g., worker versus firm have different "bargaining powers"), then we obtain:

$$x^{NS} = \arg\max_{x \in X} (u_1(x) - d_1)^{\beta} (u_2(x) - d_2)^{1-\beta} \tag{6.2}$$

where $\beta \in [0,1]$ is the bargaining power of player 1.

Next note that if both utilities are linear and defined over their share of some pie, and the set $X \subset \mathbb{R}^2$ is given by $x_1 + x_2 \leq 1$, then the solution to (6.2) is given by

$$(1-\beta)(x_1 - d_1) = \beta (x_2 - d_2),$$

with $x_1 = 1 - x_2$, which implies the linear sharing rule:

$$x_2 = (1-\beta)(1 - d_1 - d_2) + d_2.$$

Intuitively, player 2 receives a fraction $1-\beta$ of the net surplus $1 - d_1 - d_2$ plus his outside option, d_2.

In our context, the Nash bargaining solution amounts to choosing the wage, w, so as to maximize:

$$(f(k) - w)^{1-\beta} (w - \bar{w})^{\beta}.$$

An important observation is that the cost of investment, rk, does not feature in this expression, since these investment costs are sunk. In other words, the profits of the firm are $f(k) - w - rk$, while its outside option is $-rk$. So the difference between payoff and outside option for the firm is simply $f(k) - w$. The above argument immediately implies that the Nash solution, as a function of the capital investment k, can be expressed as

$$w(k) = \beta f(k) + (1-\beta)\bar{w}.$$

This expression emphasizes the dependence of the equilibrium wage on the capital stock of the firm. In contrast, in a competitive labor market, the wage that a worker of a given skill is paid is always independent of the physical capital level of his employer. Here this dependence arises because of wage bargaining, i.e., absence of a competitive market.

Therefore, at the point of investment, the profits of the firm are

$$\begin{aligned}\pi(k) &= f(k) - w(k) - rk \\ &= (1-\beta)(f(k) - \bar{w}) - rk\end{aligned}$$

The first-order condition of the profit maximization problem gives the equilibrium investment/physical capital level, k^e, as

$$(1-\beta) f'(k^e) = r$$

In comparison, the efficient level of investment that would have emerged in a competitive labor market, is given by

$$f'(k^*) = r$$

The concavity of f immediately implies that $k^e < k^*$, thus there will be underinvestment.

The reason for underinvestment is straightforward to see. Because of bargaining, the firm is not the full residual claimant of the additional returns it generates by its investment. A fraction β all the returns are received by the worker, since the wage that the firm has to pay is increasing in its capital stock.

That there are no binding contracts is important for this result. Imagine an alternative scenario where the firm and the worker first negotiate a wage contract $w(k)$ which specifies the wage that the worker will be paid for every level of physical capital. Assume this wage contract is binding, and to simplify discussion, let us limit attention to differentiable wage functions. Then, the equilibrium can be characterized as follows: first, given the wage function, find the firm's investment. This is clearly given by

$$f'(k) - w'(k) = r \tag{6.3}$$

Then, the wage function, $w(k)$, and the level of investment, k, will be chosen so as to maximize:

$$(f(k) - w(k) - rk)^{1-\beta} (w(k) - \bar{w})^{\beta}$$

where notice that now rk is now subtracted from the firm's payoff, since the negotiation is before investment costs are sunk. It is straightforward to see that the solution to this problem must have $w'(k) = 0$, so the efficient level of investment will be implemented (to see this consider changes in the functions $w(k)$ such that the derivative at the value k changes without $w(k)$ changing. By considering such changes, we can manipulate the level of k from (6.3). So the equilibrium has to satisfy the first-order condition with respect to k

$$(1-\beta)\frac{(f'(k)-w'(k)-r)}{f(k)-w(k)-rk}+\beta\frac{w'(k)}{w(k)-\bar{w}}=0$$

Using (6.3), immediately gives $w'(k)=0$).

This analysis establishes that underinvestment arises in this investment problem here because of the absence of binding contracts, which in turn lead to a *holdup problem*; once the firm invests a larger amount in physical capital, it is potentially "held up" by the worker who can demand higher wages to work for the firm, with the threat point that, if he does not accept to work for the firm, the investment of the firm will be wasted.

Is the assumption of "no binding contracts," which underlies this hold a problem reasonable? There are two reasonsfor why binding contracts are generally not possible and instead contracts have to be "incomplete":

(1) Such contracts require the level investment, k, to be easily observable by outside parties, so that the terms of a contract that makes payments conditional on k are easily enforceable (notice the important emphasis here; there is no asymmetric information between the parties, but outside courts cannot observe what the firm and the worker observe; can there be no contracts that transmits this information to outside parties in order to make contracts conditional on this information?).
(2) We need to rule out renegotiation.

A combination of these reasons imply that such binding contracts are not easy to write. This problem becomes even more serious when investments are not in physical capital, but human capital, which will be our focus below.

2. Incomplete Contracts and the Internal Organization of the Firm

The type of incompleteness of contracts discussed in the previous section plays an important role in thinking about the internal organization of the firm. This is the essence of the approach started by Coase, Williamson, and Grossman and Hart.

An important application of this approach is a theory of vertical integration. This theory provides potential answers to the question: when should two divisions be part of a single firm, in a vertically-integrated structure, and when should they function as separate firms at arm's length? Although issues related to vertical integration are not central to labor economics, they highlight the implications of incomplete contracts in other settings as well.

The answer that the incomplete contracts literature gives to the vertical integration question links the organization of the firm to the distribution of bargaining power. Whoever has the right to use physical assets has the residual rights of control, so in the event of disagreement in bargaining, he or she can use these assets. This improves the outside option of the party who owns more assets. In a vertically-integrated structure, the owner/manager of the (downstream) firm has residual rights of control, and the manager of the (upstream) division is an employee. He can be fired at will. In an arm's-length relationship, in case of disagreement, the owner of the separate (upstream) firm has the residual rights of control of his own production and assets. This improves his incentives, but harms the incentives of the owner of the other firm.

To make these ideas more specific, consider the following simple setup. A downstream firm will buy an input of quality h from an upstream supplier. He will then combine this with another input of quality s, to produce output of value

$$2\left(1-\delta+\delta e\right)s^{\alpha}h^{\gamma}, \tag{6.4}$$

where $e \in \{0,1\}$ is an ex post effort that the upstream manager exerts with some small cost ε. The cost of investing in quality for the downstream manager is equal to s, and for the upstream firm, it is h. Both qualities, as well as e, are unobserved by other parties, thus contracts conditional on h, s and e cannot be written.

We will consider two different organizational forms. The first organizational form is *vertical integration*, which involves the vesting of property rights over assets to the downstream manager. In particular, once the input is produced, the downstream manager owns the input and can appropriate the input from the upstream manager (naturally, there could also be the converse case in which there is vertical integration with the upstream manager having property rights over all the assets, but for our purposes here, it is sufficient to focus on vertical integration with property rights vested in the downstream manager). Now consider the case where the upstream manager has chosen h and the downstream firm has chosen s. If the two managers agree, then the upstream manager will exert the required effort, $e = 1$, and total surplus is given by (6.4) with $e = 1$:

$$2s^{\alpha}h^{\gamma}$$

If they disagree, the upstream firm obtains nothing, while the upstream from obtains

$$2\left(1-\delta\right)s^{\alpha}h^{\gamma}$$

Therefore, symmetric Nash bargaining gives the gross payoffs of the two parties as

$$\pi_d = \left[2\left(1-\delta\right)+\delta\right]s^{\alpha}h^{\gamma} \text{ and } \pi_u = \delta s^{\alpha}h^{\gamma}$$

Now going back to the investment stage, the two firms (managers) will choose h and s with the following first-order conditions:

$$\alpha\left[2\left(1-\delta\right)+\delta\right]s^{\alpha-1}h^{\gamma} = 1 \text{ and } \gamma\delta s^{\alpha}h^{\gamma-1} = 1,$$

which implies that

$$\frac{h}{s} = \frac{\gamma\delta}{\alpha\left[2\left(1-\delta\right)+\delta\right]}, \tag{6.5}$$

and thus

$$\begin{aligned} h &= \left(\left(\alpha\left[2\left(1-\delta\right)+\delta\right]\right)^{\alpha}\left(\gamma\delta\right)^{1-\alpha}\right)^{1/(1-\alpha-\gamma)} \text{ and} \\ s &= \left(\left(\alpha\left[2\left(1-\delta\right)+\delta\right]\right)^{1-\gamma}\left(\gamma\delta\right)^{\gamma}\right)^{1/(1-\alpha-\gamma)} \end{aligned}$$

Thus total surplus with vertical integration is:

(6.6)

$$\begin{aligned} V_{VI} &= 2\left(\left(\alpha\left[2\left(1-\delta\right)+\delta\right]\right)^{\alpha}\left(\gamma\delta\right)^{\gamma}\right)^{1/(1-\alpha-\gamma)} - \left(\left(\alpha\left[2\left(1-\delta\right)+\delta\right]\right)^{1-\gamma}\left(\gamma\delta\right)^{\gamma}\right)^{1/(1-\alpha-\gamma)} \\ &\quad - \left(\left(\alpha\left[2\left(1-\delta\right)+\delta\right]\right)^{\alpha}\left(\gamma\delta\right)^{1-\alpha}\right)^{1/(1-\alpha-\gamma)} \end{aligned}$$

Next consider the same problem with arm's-length relationship. Now, if there is disagreement, the upstream would not supply the input, so the output of the downstream firm would be 0. Similarly, the output of the upstream firm is also 0, since he's making no sales. Thus, gross payoffs are

$$\pi_d = s^\alpha h^\gamma \text{ and } \pi_u = s^\alpha h^\gamma$$

The ex ante maximization problem then gives:

$$\frac{h}{s} = \frac{\gamma}{\alpha} \tag{6.7}$$

and

$$\begin{aligned} s &= \left(\alpha^{1-\gamma}\gamma^\gamma\right)^{1/(1-\alpha-\gamma)} \text{ and} \\ h &= \left(\alpha^\alpha \gamma^{1-\alpha}\right)^{1/(1-\alpha-\gamma)} \end{aligned}$$

and total surplus is

$$\begin{aligned} V_{NI} &= 2\left(\alpha^\alpha\gamma^\gamma\right)^{1/(1-\alpha-\gamma)} \\ &\quad - \left(\alpha^{1-\gamma}\gamma^\gamma\right)^{1/(1-\alpha-\gamma)} - \left(\alpha^\alpha\gamma^{1-\alpha}\right)^{1/(1-\alpha-\gamma)} \end{aligned} \tag{6.8}$$

Comparison of (6.5) and (6.7) shows that the upstream firm is investing more relative to the downstream firm with arm's-length relationship. This is because it has better outside options with this arrangement.

Comparison of (6.6) and (6.8) in turn shows that as γ increases (while keeping $\gamma + \alpha$ constant), V_{NI} increases relative to V_{VI}. Thus for relatively high γ, implying that the quality of the input from the upsteam firm is relatively important, V_{NI} will exceed V_{VI}. It is also straightforward to see that in this world with no credit market problems when $V_{NI} > V_{VI}$, the equilibrium organization will be arm's-length and when $V_{NI} < V_{VI}$, it will be vertical integration. This is because in the absence of credit market problems, ex ante transfers will ensure that the equilibrium organizational form is chosen jointly to maximize total surplus.

Therefore, we now have a theory of vertical integration based on the relative importance of incentives of the downstream and the upstream firms.

CHAPTER 7

Efficiency Wage Models

Efficiency wage models are basically models with imperfect information where the participation constraint of employees are slack because of limited liability constraints, or sometimes because of other informational problems.

While basic agency models are widely used in contract theory, organizational economics and corporate finance, efficiency wage models are used mostly in macro and labor economics. But they are all part of the same family.

We start here with the Shapiro-Stiglitz model, which is the most famous macro/labor efficiency wage model, and provides a useful way of thinking about unemployment, which we will discuss in the context of search models as well later in the course.

1. The Shapiro-Stiglitz Model

The Shapiro-Stiglitz model is one of the workhorses of macro/labor. In this model, unemployment arises because wages need to be above the market clearing level in order to give incentives to workers. In fact, it is the combination of unemployment and high wages that make work more attractive for workers, hence the title on the paper "unemployment as a worker-discipline device".

Since the model is somewhat familiar, it is sufficient to sketch the main ingredients here: the model is in continuous time and all agents are infinitely lived.

Workers have to choose between two levels of effort, and are only productive if they exert effort.

$\longrightarrow$effort $\quad\longrightarrow 0 \quad \tilde{}\,\text{cost} = 0$, not productive

$\qquad\qquad\quad\;\longrightarrow 1 \quad \tilde{}\,\text{cost} = e$, productive

Without any informational problems firms would write contracts to pay workers only if they exert effort. The problem arises because firms cannot observe whether a worker has exerted effort or not, and cannot deduce it from output, since output is a function of all workers' efforts. This introduces the *moral hazard problem*

The model is in continuous time, so instead of probabilities we will be talking about flow rates.

If a worker "shirks", there is effort $= 0$, then there is probability (flow rate) q of getting detected and fired. [...For example, the worker's actions affect the probability distribution of some observable signal on the basis of which the firm compensates him. When the worker exerts effort, this signal takes the value 1. When he shirks, this signal is equal to 1 with probability $1 - q$ and 0 with probability q...]

All agents are risk neutral, and there are N workers

$b =$ exogenous separation rate

$a =$ job finding rate, which will be determined in equilibrium

$r =$interest rate/discount factor

These type of dynamic models are typically solved by using dynamic programming/Bellman equations. Although the theory of dynamic programming can be sometimes difficult, in the context of this model, its application is easy.

We will make the analysis even easier by focusing on steady states. In steady state, we can simply think of the present discounted value (PDV) of workers as a function of their "strategy" of shirking or working hard.

Denote the PDV of employed-shirker by V_E^S (recall we are in continuous time)

$$rV_E^S = w + (b + q)(V_U - V_E^S) \tag{7.1}$$

where we have imposed $\dot{V}_E^S = 0$, since here we will only characterize steady states. The intuition for this equation is straightforward. The worker always receives his wage (his compensation for this instant of work) w, but at the flow rate b, he separates from the firm exogenously, and at the flow rate q, he gets caught for shirking, and in both cases he becomes unemployed, receiving V_U and losing V_E^S.

[The full continuous-time dynamic programming equation would be

$$rV_E^S - \dot{V}_E^S = w + (b+q)(V_U - V_E^S)$$

but in steady state $\dot{V}_E^S = 0$...]

Denote the PDV of employed-nonshirker by V_E^N

$$rV_E^N = w - e + b(V_U - V_E^N), \tag{7.2}$$

which is different from (7.1) because the worker incurs the cost e, but loses his job at the slower rate b.

PDV of unemployed workers V_U is

$$rV_U = z + a(V_E - V_U),$$

where

$$V_E = \max\left\{V_E^S, V_E^N\right\}$$

and z is the utility of leisure + unemployment benefit.

Non-shirking condition is an incentive-compatibility constraint that requires the worker to prefer to exert effort. Combining these equations, we obtain it as

$$\begin{aligned} V_E^N &\geq V_E^S \\ w &\geq rV_U + [r+b+a]\,\frac{e}{q} \quad \text{[non-shirking condition].} \end{aligned}$$

This equation is intuitive. The greater is the unemployment benefit and the greater is the cost of effort, the greater should the wage be. More importantly, the more likely the worker is to be caught when he shirks, the lower is the wage. Also, the wages higher when r, b and a are higher. Why is this?

Not shirking is an investment (why is this?), so the greater is r, the less attractive it is. This also explains the effect of b; the greater is this parameter, the more likely one is to leave the job, so this is just like discounting.

Finally, the effect of a can be understood by thinking of unemployment as punishment (after all, if it weren't, why would the worker care about being fired?). The lower is a, the harder it is to move out of unemployment, the harsher is unemployment as a punishment, thus wages will not need to be as high.

Steady state requires that $\Longrightarrow$

$$\text{flow into unemployment} = \text{flow out of unemployment}$$

Again this is a type of equation we will see a lot when we study search models below.

In equilibrium, no one shirks because the non-shirking condition holds (similar to the agents doing the right thing in the agency models).

Therefore,

$$bL = aU$$

where L is employment, and U unemployment.

This equation immediately determines the flow rate out of unemployment as

$$a = \frac{bL}{U} = \frac{bL}{N-L}.$$

Now substituting for this we get the full non-shirking condition as

$$\text{Non-Shirking Condition}: w \geq z+e + \left[r + \frac{bN}{N-L}\right]\frac{e}{q}$$

Notice that a higher level of $\frac{N}{N-L}$, which corresponds to lower unemployment, necessitate a higher wage to satisfy the non-shirking condition. This is the sense in which unemployment is a worker-discipline device. Higher unemployment makes losing the job more costly, hence encourages workers not to shirk.

Next, let us consider the determination of labor demand in this economy. Let us suppose that there are M firms, each with access to a production function

$$AF(L),$$

where L denotes their labor. We make the standard assumptions on F, in particular, it is increasing and strictly concave, i.e. $F'' < 0$.

These firms maximize static profits (no firing/hiring costs).

This implies that the equilibrium will satisfy:

$$AF'(L) = w,$$

Aggregate Labor Demand will therefore be given by

$$AF'\left(\frac{L}{M}\right) = w.$$

Figure 7.1 shows the determination of the equilibrium diagrammatically. It plots the non-shirking condition, in the labor-wage space as an upward sloping curve asymptoting to infinity at $L = N$ (full employment) as well as the downward sloping "labor demand curve" given by $AF'(L)$ (for now ignore the average product line).

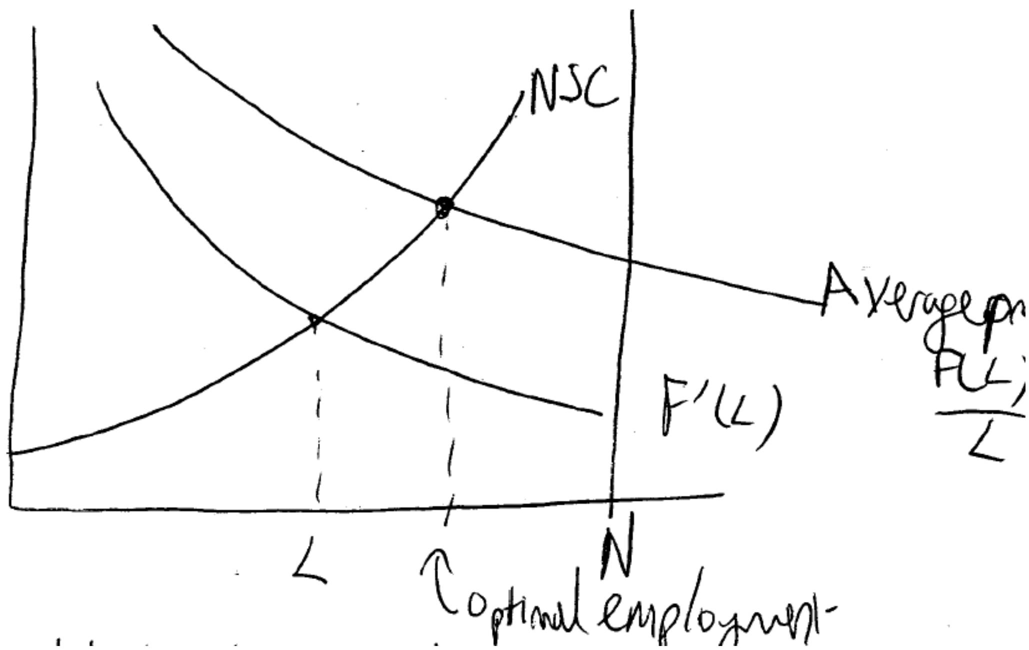


FIGURE 7.1

Set $M = 1$ as a normalization, then equilibrium will be given by the following (famous) equation:

$$z + e + \left[r + \frac{bN}{N-L}\right]\frac{e}{q} = AF'(L)$$

This equation basically equates labor demand to quasi-labor supply.

This is quasi-labor supply rather than real labor supply, because it is not determined by the work leisure trade-off of workers, but by the non-shirking condition—if wages did not satisfy this condition, workers will shirk and firms would lose money.

Given this equation, comparative statics are straightforward and intuitive:

$A \downarrow \Longrightarrow L \downarrow$: lower prod. $\Longrightarrow$ high unemployment

$z \uparrow \Longrightarrow L \downarrow$: high reservation wages $\Longrightarrow$ high unemployment

$q \downarrow \Longrightarrow L \downarrow$: bad monitoring $\Longrightarrow$ high unemployment

$r \uparrow \Longrightarrow L \downarrow$: high interest rates $\Longrightarrow$high unemployment

$b \uparrow \Longrightarrow L \downarrow$: high turnover $\Longrightarrow$ high unemployment

Since there is unemployment, rents and information problems here, it is also natural to ask the welfare question: is the level of unemployment too high? It depends on what notion of welfare we are using and whether firms are owned by nonworkers.

What are the externalities?

(1) By hiring one more worker, the firm is reducing unemployment, and forcing other firms to pay higher wages$\rightarrow$unemployment is too low.

(2) By hiring one more worker, the firm is increasing the worker's utility at the margin, since each worker is receiving a rent (wage > opportunity cost)$\rightarrow$unemployment is too high.

The diagram shows that the second effect always dominates (now consider the average product line). The unemployment is too high. A subsidy on wages financed by a tax on profits will increase output.

So can there be a Pareto-improving tax-subsidy scheme?

Contrary to what Shapiro and Stiglitz claim, the answer is *not necessarily.* If firms are owned by capitalists, the above policy will increase output, but will not constitute a Pareto improvement (think about whether Pareto improvement is the right criterion to look at in this case).

If firms are owned by workers, the above policy will constitute a Pareto improvement. But in this case workers have enough income. Why do they not already enter into "bonding" contracts or at least write better contracts as in our moral hazard models?

2. Other Solutions to Incentive Problems

In fact, the discussion at the end of the previous section immediately east of the question: Are firms behaving optimally? The answer to this question is also: not necessarily. Firms can write better contracts (from the viewpoint of maximizing profits) even when workers are severely credit constrained. In particular, *backloading compensation* for workers is always feasible and will be more effective in preventing shirking.

This is one of the main criticisms of the shirking model: the presence of the monitoring problem does not necessarily imply "rents" for workers, and it is the rents for the workers that lead to distortions and unemployment.

Moreover, if workers have wealth, they can enter into bonding contracts where they post a bond that they lose if they are caught shirking.

Problems: firm-side moral hazard—firms may claim workers have shirked and fire them either to reduce labor costs when the worker's wage has increased enough (above the opportunity cost), or to collect the bond payments.

In fact, in practice we observe a lot of upward sloping payment schedules for workers, and pensions and other benefits that they receive after retirement. Ed Lazear has argued that these are precisely responses to the incentive problems that workers face.

In any case, this discussion highlights that there are two empirical questions:

(1) Are monitoring problems important?
(2) Do more severe monitoring problems lead to greater rents for workers?

3. Evidence on Efficiency Wages

There are two types of evidence offered in the literature in support of efficiency wages.

The first type of evidence shows the presence of substantial inter-industry wage differences (e.g., Krueger and Summers). Such wage differentials are consistent with efficiency wage theories since the monitoring problem (q in terms of the model above) is naturally more serious in some industries than others.

Nevertheless, this evidence does not establish that efficiency wage considerations are important, since there are at least two other explanations for the inter-industry wage differentials:

(1) These differentials may reflect compensating wages (since some jobs may be less pleasant than others) or premia for unobserved characteristics of workers, which differ systematically across industries because workers select into industries based on their abilities.

It seems to be the case that a substantial part of the wage differentials are in fact driven by these considerations. Nevertheless, it also seems to be the case that part of the inter-industry wage differences do in fact correspond to "rents". Workers who move from a low wage to a high wage industry receive a wage increase in line with the wage differential between these two sectors (Krueger and Summers; Gibbons and Katz), suggesting that the differentials do not simply reflect unobserved ability (Nevertheless, this evidence is not watertight; what if workers move precisely when there is "good news" about their abilities?).

Compensating wage differentials also do not seem to be the whole story: workers are less likely to quit such jobs (Krueger and Summers), but let's see the discussion of Holzer, Katz, and Krueger below.

(2) Inter-industry wage differentials may correspond to differential worker rents in different industries, but not because of efficiency wages, but because of differences in unionization or other industry characteristics that give greater bargaining power to workers in some industries than others (e.g., capital intensity).

Therefore, the inter-industry wage differentials are consistent with efficiency wages, but do not prove that efficiency wage considerations are important.

An interesting paper by Holzer, Katz, and Krueger investigates whether higher wages attract more applicants (which would be an indication that these jobs might be more attractive). They find that when wages are exogenously higher because of minimum wages, there are in fact more applicants. However, interindustry wage differentials don't seem to induce more applications!

The second line of attack looks for direct evidence for efficiency wage considerations. A number of studies find support for efficiency wages. These include:

(1) Krueger compares wages and tenure premia in franchised and company-owned fast food restaurants. Krueger makes the natural assumption that there is less monitoring of workers in a franchised restaurant. He finds higher wages and steeper wage-tenure profiles in the franchised restaurants, which he interprets as evidence for efficiency wages.
(2) Cappelli and Chauvin provide more convincing evidence. They look at the number of disciplinary dismissals, which they interpret as a measure of shirking, in the different plants located in different areas, but all by the same automobile manufacturer (and covered by the same union). The firm pays the same nominal wage everywhere (because of union legislation). This nominal wage translates into greater wage premia in some areas because outside wages differ. They find that when wage premia are greater, there are fewer disciplinary dismissals. This appears to provide strong support to the basic implication of the shirking model.
(3) Campbell and Kamlani survey 184 firms and find that firms are often unwilling to cut wages because this will reduce worker effort and increase shirking.

These various pieces of evidence together suggest efficiency wage considerations are important. Nevertheless they do not indicate whether these efficiency wages are the main reason why wages are higher than market-cleaning levels and unemployment is high either in the U.S. or in Europe. Such an investigation requires more aggregate evidence.

4. Efficiency Wages, Monitoring and Corporate Structure

Next consider a simple model where we use the ideas of efficiency wages for think about the corporate structure.

For simplicity, take corporate structure to be the extent of monitoring (e.g., number of supervisors to production workers).

Consider a one-period economy consisting of a continuum of measure N of workers and a continuum of measure 1 of firm owners who are different from the workers.

Each firm i has the production function $AF(L_i)$.

Differently from the Shapiro-Stiglitz model, let the probability of catching a shirking worker be endogenous. In particular, let $q_i = q(m_i)$ where m_i is the degree of monitoring per worker by firm i.

The cost of monitoring for firm i which hires L_i workers is sm_iL_i.For example, m_i could be the number of managers per production worker and s as the salary of managers.

Since there is a limited liability constraint, workers cannot be paid a negative wage, and the worst thing that can happen to a worker is to receive zero income. Since all agents are risk-neutral, without loss of generality, restrict attention to the case where workers are paid zero when caught shirking.

Therefore, the incentive compatibility constraint of a worker employed in firm i can be written as:

$$w_i - e \geq (1 - q_i)w_i.$$

If the worker exerts effort, he gets utility $w_i - e$, which gives the left hand side of the expression. If he chooses to shirk, he gets caught with probability q_i and receives zero. If he is not caught, he gets w_i without suffering the cost of effort. This gives the right hand side of the expression.

Notice an important difference here from the Shapiro-Stiglitz model. Now if the worker is caught shirking, he does not receive the wage payment.

Firm i's maximization problem can be written as:

$$\max_{w_i, L_i, q_i} \Pi = AF(L_i) - w_i L_i - s m_i L_i \tag{7.3}$$

subject to:

$$w_i \geq \frac{e}{q(m_i)} \tag{7.4}$$

$$w_i - e \geq \underline{u} \tag{7.5}$$

The first constraint is the incentive compatibility condition rearranged. The second is the participation constraint where $\underline{u}$ is the *ex ante* reservation utility (outside option) of the worker; in other words, what he could receive from another firm in this market.

The maximization problem (7.3) has a recursive structure: m and w can be determined first without reference to L by minimizing the cost of a worker $w + sm$ subject to (7.4) and (7.5); then, once this cost is determined, the profit maximizing level of employment can be found. Each subproblem is strictly convex, so the solution is uniquely determined, and all firms will make the same choices: $m_i = m$, $w_i = w$ and $L_i = L$. In other words, the equilibrium will be symmetric.

Another useful observation is that the incentive compatibility constraint (7.4) will always bind [Why is this? Think as follows: if the incentive compatibility constraint, (7.4), did not bind, the firm could lower q, and increase profits without affecting anything else. This differs from the simplest moral hazard problem with fixed q in which the incentive compatibility constraint (7.4) could be slack...]

By contrast, the participation constraint (7.5) may or may not bind—hence there may or may not be be rents for workers; contrast this with the Shapiro-Stiglitz model. The comparative statics of the solution have a very different character depending on whether it does. The two situations are sketched in the figures.

(1) When (7.5) does not bind, the solution is characterized by the tangency of the (7.4) with the per-worker cost $w + sm$.

Call this solution (w^*, m^*), where:

$$\frac{eq'(m^*)}{(q(m^*))^2} = s \text{ and } w^* = \frac{e}{q(m^*)}. \tag{7.6}$$

In this case, because the participation constraint (7.5) does not bind, w and m are given by (7.6) and small changes in $\underline{u}$ leave these variables unchanged.

(2) In contrast, if (7.5) binds, w is determined directly from this constraint as equal to $\underline{u}+e$, and an increase in $\underline{u}$ causes the firm to raise this wage. Since (7.4) holds in this case, the firm will also reduce the amount of information gathering, m.

What determines whether (7.5) binds?

Let $\hat{w}$ and $\hat{m}$ be the per-worker cost minimizing wage and monitoring levels (which would not be equal to w^* and m^* when (7.5) binds). Then, labor demand of a representative firm solves:

$$AF'(\hat{L}) = \hat{w} + s\hat{m}. \tag{7.7}$$

Next, using labor demand, we can determine $\underline{u}$, workers' *ex ante* reservation utility from market equilibrium. It depends on how many jobs there are. If aggregate demand $\hat{L}$ is greater than or equal to N, then a worker who turns down a job is sure to get another. In contrast, if aggregate demand $\hat{L}$ is less than N, then a worker who turns down a job may end up without another. In particular, in this case, $\underline{u} = \frac{\hat{L}}{N}(\hat{w} - e) + (1 - \frac{\hat{L}}{N})z$, where z is an unemployment benefit that a worker who cannot find a job receives.

When $\hat{L} = N$, there are always firms who want to hire an unemployed worker at the beginning of the period, and thus $\underline{u} = \hat{w} - e$. If there is excess supply of workers, i.e. $\hat{L} < N$, then firms can set the wage as low as they want, and so they will choose the profit maximizing wage level w^* as given by (7.6). In contrast, with full employment, firms have to pay a wage equal to $\underline{u} + e$ which will generically exceed the (unconstrained) profit maximizing wage rate w^*. Therefore, we can think of labor demand as a function of $\underline{u}$, the reservation utility of workers: firms are "utility-takers" rather than price-takers. The figures show the two cases; the

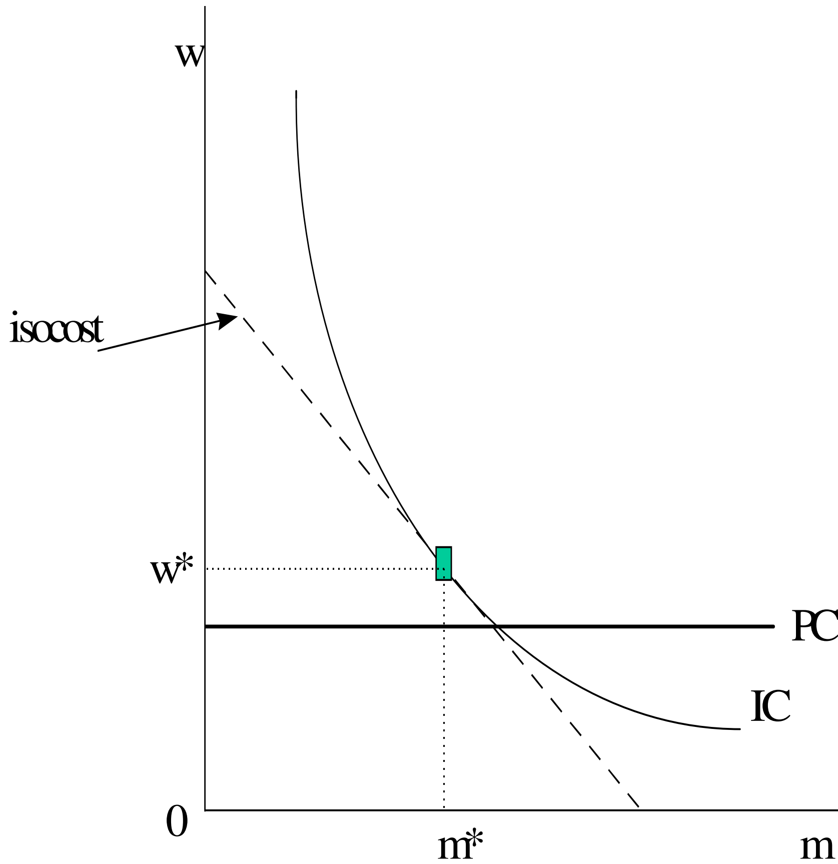


FIGURE 7.2. Participation Constraint is Slack.

outcome depends on the state of labor demand. More importantly, the comparative statics are very different in the two cases.

Now comparative statics are straightforward.

First, consider a small increase A and suppose that (7.5) is slack. The tangency between (7.4) and the per worker cost is unaffected. Therefore, neither w nor m change. Instead, the demand for labor shifts to the right and firms hire more workers. As long as (7.5) is slack, firms will continue to choose their (market) unconstrained optimum, (w^*, m^*), which is independent of the marginal product of labor. As a result, changes in labor demand do not affect the organizational form of the firm.

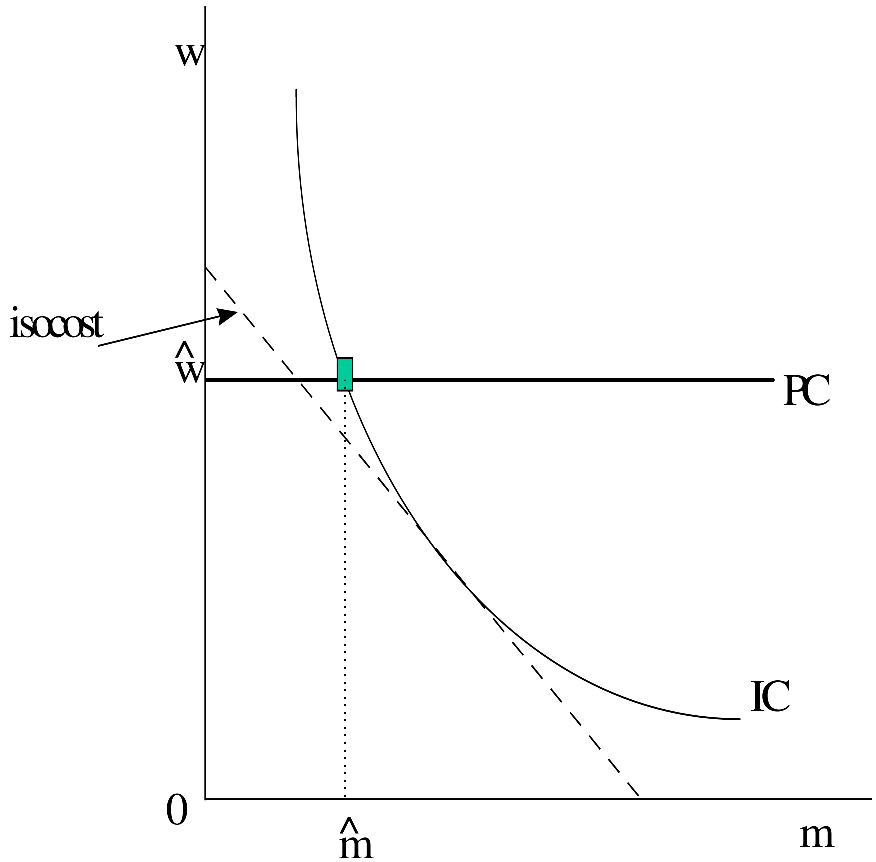


Figure 7.3. Participation Constraint is Binding.

If instead (7.5) holds as an equality, comparative static results will be different. In this case, (7.4), (7.7), and $L = N$ jointly determine $\hat{q}$ and $\hat{w}$. An increase in A induces firms to demand more labor, increasing $\hat{w}$. Since (7.4) holds, this reduces $\hat{q}$ as can be seen by shifting the PC curve up. Therefore, when (7.5) holds, an improvement in the state of labor demand reduces monitoring. The intuition is closely related to the fact that workers are subject to limited liability. When workers cannot be paid negative amounts, the level of their wages is directly related to the *power* of the incentives. The higher are their wages, the more they have to lose by being fired and thus the less willing they are to shirk.

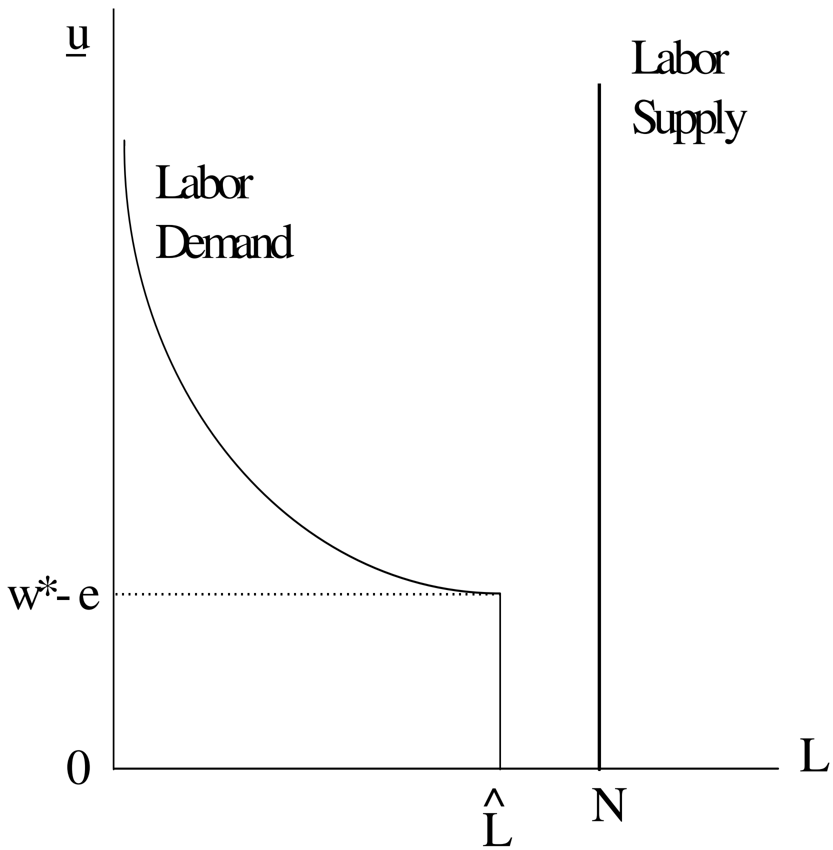


FIGURE 7.4. Participation Constraint is Slack.

Next, suppose that government introduces a wage floor $\underline{w}$ above the equilibrium wage (or alternatively, unions demand a higher wage than would have prevailed in the non-unionized economy). Since the incentive compatibility constraint (7.4) will never be slack, a higher wage will simply move firms along the IC curve in the figure and reduce m. However, this will also increase total cost of hiring a worker, reducing employment.

Can this model be useful in thinking about why the extent of monitoring appears to be behaving differently in continental Europe and the U.S.?

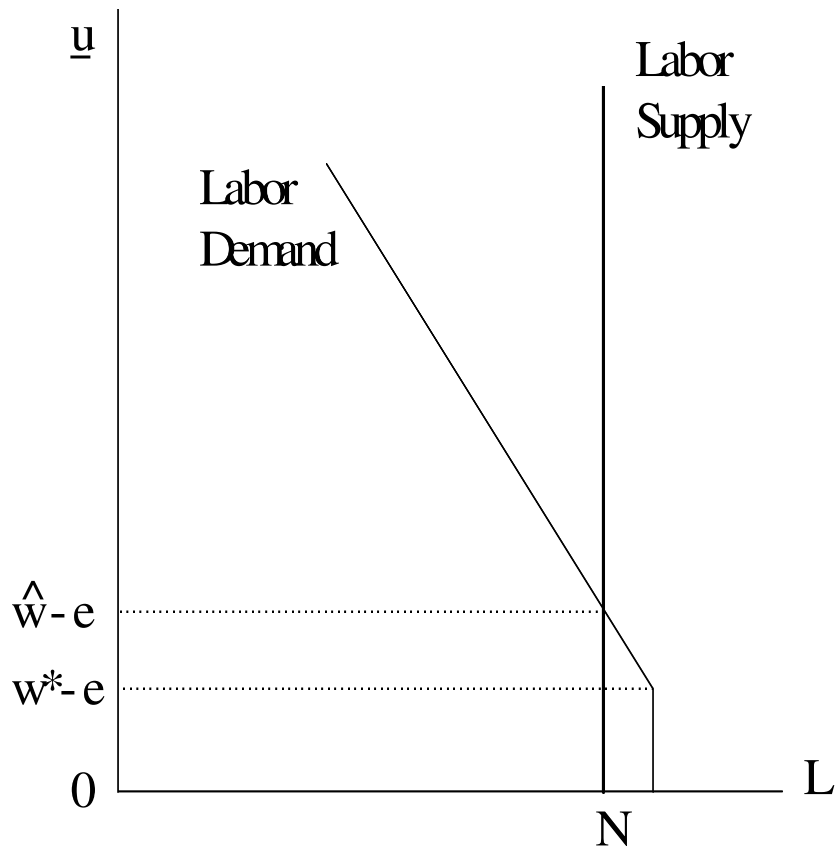


Figure 7.5. Participation Constraint is Binding.

Whether this model can explain these patterns or not is not clear. But cross-country differences in broad features of organizations are very stark, and investigation of these issues seems to be a very interesting area for future research.

Finally, let us look at welfare in this model.

Consider the aggregate surplus Y generated by the economy:

$$Y = AF(L) - smL - eL, \tag{7.8}$$

where $AF(L)$ is total output, and eL and smL are the (social) input costs.

In this economy, the equilibrium is constrained Pareto efficient: subject to the informational constraints, a social planner could not increase the utility of workers

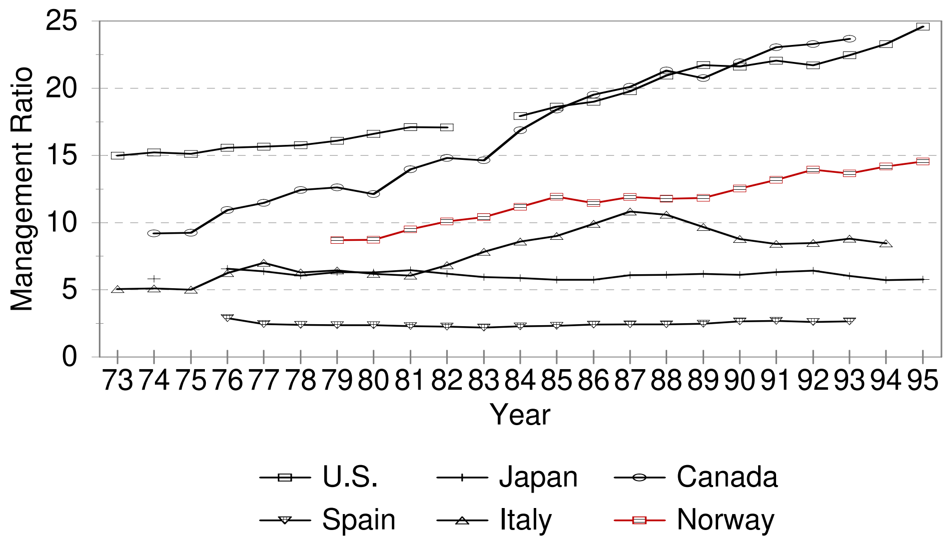


FIGURE 7.6. Trends in the ratio of managerial employees to non-managerial, non-agricultural workers in six countries. Source ILO Labor Statistics.

without hurting the owners. But total surplus Y is never maximized in laissez-faire equilibrium. This is because of the following reason: if we can reduce q without changing L, then Y increases. A tax on profits used to subsidize w relaxes the incentive constraint (7.4) and allows a reduction in monitoring. Indeed, the second-best allocation which maximizes Y subject to (7.4) would set wages as high as possible subject to zero profits for firms. Suppose that the second-best optimal level of employment is $\tilde{L}$, then we have:

$$\tilde{w} + sq^{-1}\left(\frac{e}{\tilde{w}}\right) = \frac{AF(\tilde{L})}{\tilde{L}} \tag{7.9}$$

In this allocation, all firms would be making zero-profits; since in the decentralized allocation, due to decreasing returns, they are always making positive profits, the two will never coincide.

A different intuition for why the decentralized equilibrium fails to maximize net output is as follows: part of the expenditure on monitoring, smL, can be interpreted as "rent-seeking" by firms. Firms are expending resources to reduce wages — they are trying to minimize the private cost of a worker $w + sm$ — which is to a first-order approximation, a pure transfer from workers to firms. A social planner who cares only about the size of the national product wants to minimize $e + sm$, and therefore would spend less on monitoring. Reducing monitoring starting from the decentralized equilibrium would therefore increase net output.

Part 3

Investment in Post-Schooling Skills

CHAPTER 8

The Theory of Training Investments

1. General Vs. Specific Training

In the Ben-Porath model, an individual continues to invest in his human capital after he starts employment. We normally think of such investments as "training", provided either by the firm itself on-the-job, or acquired by the worker (and the firm) through vocational training programs. This approach views training just as schooling, which is perhaps too blackbox for most purposes.

More specifically, two complications that arise in thinking about training are:

(1) Most of the skills that the worker acquires via training will not be as widely applicable as schooling. As an example, consider a worker who learns how to use a printing machine. This will only be useful in the printing industry, and perhaps in some other specialized firms; in this case, the worker will be able to use his skills only if he stays within the same industry. Next, consider the example of a worker who learns how to use a variety of machines, and the current employer is the only firm that uses this exact variety; in this case, if the worker changes employer, some of his skills will become redundant. Or more extremely, consider a worker who learns how to get along with his colleagues or with the customers of his employer. These skills are even more "specific", and will become practically useless if he changes employer.

(2) A large part of the costs of training consist of forgone production and other costs borne directly by the employer. So at the very least, training investments have to be thought as joint investments by the firm and the worker, and in many instances, they may correspond to the firm's decisions more than to that of the worker.

The first consideration motivates a particular distinction between two types of human capital in the context of training:

(1) Firm-specific training: this provides a worker with *firm-specific skills*, that is, skills that will increase his or her productivity *only* with the current employer.
(2) General training: this type of training will contribute to the worker's general human capital, increasing his productivity with a range of employers.

Naturally, in practice actual training programs could (and often do) provide a combination of firm-specific and general skills.

The second consideration above motivates models in which firms have an important say in whether or not the worker undertakes training investments. The extreme but not show case is the one where training costs are borne by the firm (for example, because the process of training reduces production), and in this case, the firm directly deciding whether and how much training the worker will obtain may be a good approximation to reality and a good starting point for our analysis.

2. The Becker Model of Training

Let us start with investments in general skills. Consider the following stylized model:

- At time $t = 0$, there is an initial production of y_0, and also the firm decides the level of training τ, incurring the cost $c(\tau)$. Let us assume that $c(0) = 0$, $c'(0) = 0$, $c'(\cdot) \geq 0$ and $c''(\cdot) > 0$. The second assumption here ensures that it is always socially beneficial to have some amount of positive training.
- At time $t = 1/2$, the firm makes a wage offer w to the worker, and other firms also compete for the worker's labor. The worker decides whether to quit and work for another firm. Let us assume that there are many identical firms who can use the general skills of the worker, and the worker does not incur any cost in the process of changing jobs. This assumption makes the

labor market essentially competitive. (Recall: there is no informational asymmetry here).

- At time $t = 1$, there is the second and final period of production, where output is equal to $y_1 + \alpha(\tau)$, with $\alpha(0) = 0$, $\alpha'(\cdot) > 0$ and $\alpha''(\cdot) < 0$. For simplicity, let us ignore discounting.

First, note that a social planner wishing to maximize net output would choose a positive level of training investment, $\tau^* > 0$, given by

$$c'(\tau^*) = \alpha'(\tau^*).$$

The fact that τ^* is strictly positive immediately follows from the fact that $c'(0) = 0$ and $\alpha'(0) > 0$.

Before Becker analyzed this problem, the general conclusion, for example conjectured by Pigou, was that there would be underinvestment in training. The reasoning went along the following lines. Suppose the firm invests some amount $\tau > 0$. For this to be profitable for the firm, at time $t = 1$, it needs to pay the worker at most a wage of

$$w_1 < y_1 + \alpha(\tau) - c(\tau)$$

to recoup its costs. But suppose that the firm was offering such a wage. Could this be an equilibrium? No, because there are other firms who have access to exactly the same technology, they would be willing to bid a wage of $w_1 + \varepsilon$ for this worker's labor services. Since there are no costs of changing employer, for ε small enough such that

$$w_1 + \varepsilon < y_1 + \alpha(\tau),$$

a firm offering $w_1 + \varepsilon$ would both attract the worker by offering this higher wage and also make positive profits. This reasoning implies that in any competitive labor market, we must have

$$w_1 = y_1 + \alpha(\tau).$$

But then, the firm cannot recoup any of its costs and would like to choose $\tau = 0$. Despite the fact that a social planner would choose a positive level of training investment, $\tau^* > 0$, the pre-Becker view was that this economy would fail to invest in training.

The mistake in this reasoning was that it did not take into account the worker's incentives to invest in his own training. In effect, the firm does not get any of the returns from training because the worker is receiving all of them. In other words, the worker is the *full residual claimant* of the increase in his own productivity, and in the competitive equilibrium of this economy without any credit market or contractual frictions, he would have the right incentives to invest in his training.

Let us analyze this equilibrium now. As is the case in all games of this sort, we are interested in the subgame perfect equilibria. So we have to solve the game by backward induction. First note that at $t = 1$, the worker will be paid $w_1 = y_1 + \alpha(\tau)$. Next recall that τ^* is the efficient level of training given by $c'(\tau^*) = \alpha'(\tau^*)$. Then in the unique subgame perfect equilibrium, in the first period the firm will offer the following package: training of τ^* and a wage of

$$w_0 = y_0 - c(\tau^*).$$

Then, in the second period the worker will receive the wage of

$$w_1 = y_1 + \alpha(\tau^*)$$

either from the current firm or from another firm.

To see why no other allocation could be an equilibrium, suppose that the firm offered (τ, w_0), such that $\tau \neq \tau^*$. For the firm to break even we need that $w_0 \leq y_0 - c(\tau)$, but by the definition of τ^*, we have

$$y_0 - c(\tau^*) + y_1 + \alpha(\tau^*) > y_0 - c(\tau) + y_1 + \alpha(\tau) \geq w_0 + y_1 + \alpha(\tau)$$

So the deviation of offering $(\tau^*, y_0 - c(\tau^*) - \varepsilon)$ for ε sufficiently small would attract the worker and make positive profits. Thus, the unique equilibrium is the one in which the firm offers training τ^*.

Therefore, in this economy the efficient level of training will be achieved with firms bearing none of the cost of training, and workers financing training by taking a wage cut in the first period of employment (i.e, a wage $w_0 < y_0$).

There are a range of examples for which this model appears to provide a good description. These include some of the historical apprenticeship programs where young individuals worked for very low wages and then "graduated" to become master craftsmen; pilots who work for the Navy or the Air Force for low wages, and then obtain much higher wages working for private sector airlines; securities brokers, often highly qualified individuals with MBA degrees, working at a pay level close to the minimum wage until they receive their professional certification; or even academics taking an assistant professor job at Harvard despite the higher salaries in other departments.

3. Market Failures Due to Contractual Problems

The above result was achieved because firms could commit to a wage-training contract. In other words, the firm could make a credible commitment to providing training in the amount of τ^*. Such commitments are in general difficult, since outsiders cannot observe the exact nature of the "training activities" taking place inside the firm. For example, the firm could hire workers at a low wage pretending to offer them training, and then employ them as cheap labor. This implies that contracts between firms and workers concerning training investments are naturally *incomplete*.

To capture these issues let us make the timing of events regarding the provision of training somewhat more explicit.

- At time $t = -1/2$, the firm makes a training-wage contract offer (τ', w_0). Workers accept offers from firms.
- At time $t = 0$, there is an initial production of y_0, the firm pays w_0, and also unilaterally decides the level of training τ, which could be different from the promised level of training τ'.
- At time $t = 1/2$, wage offers are made, and the worker decides whether to quit and work for another firm.

- At time $t = 1$, there is the second and final period of production, where output is equal to $y_1 + \alpha(\tau)$.

Now the subgame perfect equilibrium can be characterized as follows: at time $t = 1$, a worker of training τ will receive $w_1 = y_1 + \alpha(\tau)$. Realizing this, at time $t = 0$, the firm would offer training $\tau = 0$, irrespective of its contract promise. Anticipating this wage offer, the worker will only accept a contract offer of the form (τ', w_0), such that $w_0 \geq y_0$, and τ does not matter, since the worker knows that the firm is not committed to this promise. As a result, we are back to the outcome conjectured by Pigou, with no training investment by the firm.

A similar conclusion would also be reached if the firm could write a binding contract about training, but the worker were subject to credit constraints and $c(\tau^*) > y_0$, so the worker cannot take enough of a wage cut to finance his training. In the extreme case where $y_0 = 0$, we are again back to the Pigou outcome, where there is no training investment, despite the fact that it is socially optimal to invest in skills (which one of these problems, contractual incompleteness or credit market constraints, appears more important in the context of training?).

4. Training in Imperfect Labor Markets

4.1. Motivation. The general conclusion of both the Becker model with perfect (credit and labor) markets and the model with incomplete contracts (or severe credit constraints) is that there will be no firm-sponsored investment in general training. This conclusion follows from the common assumption of these two models, that the labor market is competitive, so the firm will never be able to recoup its training expenditures in general skills later during the employment relationship.

Is this a reasonable prediction? The answer appears to be no. There are many instances in which firms bear a significant fraction (sometimes all) of the costs of general training investments.

The first piece of evidence comes from the German apprenticeship system. Apprenticeship training in Germany is largely general. Firms training apprentices have to follow a prescribed curriculum, and apprentices take a rigorous outside exam in their trade at the end of the apprenticeship. The industry or crafts chambers certify whether firms fulfill the requirements to train apprentices adequately, while works councils in the firms monitor the training and resolve grievances. At least in certain technical and business occupations, the training curricula limit the firms' choices over the training content fairly severely. Estimates of the net cost of apprenticeship programs to employers in Germany indicate that firms bear a significant financial burden associated with these training investments. The net costs of apprenticeship training may be as high as DM 6,000 per worker (in the 1990s, equivalent of about \$6,000 today).

Another interesting example comes from the recent growth sector of the US, the temporary help industry. The temporary help firms provide workers to various employers on short-term contracts, and receive a fraction of the workers' wages as commission. Although blue-collar and professional temporary workers are becoming increasingly common, the majority of temporary workers are in clerical and secretarial jobs. These occupations require some basic computer, typing and other clerical skills, which temporary help firms often provide before the worker is assigned to an employer. Workers are under no contractual obligation to the temporary help firm after this training program. Most large temporary help firms offer such training to all willing individuals. As training prepares the workers for a range of different assignments, it is almost completely general. Although workers taking part in the training programs do not get paid, all the monetary costs of training are borne by the temporary help firms, giving us a clear example of firm-sponsored general training. This was first noted by Krueger and is discussed in more detail by David Autor.

Other evidence is not as clear-cut, but suggests that firm-sponsored investments in general skills are widespread. A number of studies have investigated whether workers who take part in general training programs pay for the costs by taking lower wages. The majority of these studies do not find lower wages for workers in training programs, and even when wages are lower, the amounts typically appear too small to compensate firms for the costs. Although this pattern can be explained within the paradigm of Becker's theory by arguing that workers selected for training were more skilled in unobserved dimensions, it is broadly supportive of widespread firm-sponsored-training.

There are also many examples of firms that send their employees to college, MBA or literacy programs, and problem solving courses, and pay for the expenses while the wages of workers who take up these benefits are not reduced. In addition, many large companies, such as consulting firms, offer training programs to college graduates involving general skills. These employers typically pay substantial salaries and bear the full monetary costs of training, even during periods of full-time classroom training.

How do we make sense of these firm-sponsored investments in general training? We will now illustrate how in frictional labor markets, firms may also be willing to make investments in the general skills of their employees.

4.2. A Basic Framework. Consider the following two-period model. In period 1, the worker and/or the employer choose how much to invest in the worker's general human capital, τ. There is no production in the first period. In period 2, the worker either stays with the firm and produces output $y = f(\tau)$, where $f(\tau)$ is a strictly increasing and concave function. The worker is also paid a wage rate, $w(\tau)$ as a function of his skill level (training) τ, or he quits and obtains an outside wage. The cost of acquiring τ units of skill is again $c(\tau)$, which is again assumed to be continuous, differentiable, strictly increasing and convex, and to satisfy $c'(0) = 0$. There is no discounting, and all agents are risk-neutral.

Assume that all training is *technologically general* in the sense that $f(\tau)$ is the same in all firms.

If a worker leaves his original firm, then he will earn $v(\tau)$ in the outside labor market. Suppose

$$v(\tau) < f(\tau).$$

That is, despite that fact that τ is general human capital, when the worker separates from the firm, he will get a lower wage than his marginal product in the current firm. The fact that $v(\tau) < f(\tau)$ implies that there is a surplus that the firm and the worker can share when they are together. Also note that $v(\tau) < f(\tau)$ is only possible in labor markets with frictions—otherwise, the worker would be paid his full marginal product, and $v(\tau) = f(\tau)$.

Let us suppose that this surplus will be divided by asymmetric Nash bargaining with worker bargaining power given by $\beta \in (0, 1)$. Recall from above that asymmetric Nash bargaining and risk neutral preferences imply that the wage rate as a function of training is

$$w(\tau) = v(\tau) + \beta\left[f(\tau) - v(\tau)\right]. \tag{8.1}$$

An important point to note is that the equilibrium wage rate $w(\tau)$ is independent of $c(\tau)$: the level of training is chosen first, and then the worker and the firm bargain over the wage rate. At this point the training costs are already sunk, so they do not feature in the bargaining calculations (bygones are bygones).

Assume that τ is determined by the investments of the firm and the worker, who independently choose their contributions, c_w and c_f, and τ is given by

$$c(\tau) = c_w + c_f.$$

Assume that \$1 investment by the worker costs \$$p$ where $p \geq 1$. When $p = 1$, the worker has access to perfect credit markets and when $p \to \infty$, the worker is severely constrained and cannot invest at all.

More explicitly, the timing of events are:

- The worker and the firm simultaneously decide their contributions to training expenses, c_w and c_f. The worker receives an amount of training τ such that $c(\tau) = c_w + c_f$.

- The firm and the worker bargain over the wage for the second period, $w(\tau)$, where the threat point of the worker is the outside wage, $v(\tau)$, and the threat point of the firm is not to produce.
- Production takes place.

Given this setup, the contributions to training expenses c_w and c_f will be determined noncooperatively. More specifically, the firm chooses c_f to maximize profits:

$$\pi(\tau) = f(\tau) - w(\tau) - c_f = (1-\beta)\left[f(\tau) - v(\tau)\right] - c_f.$$

subject to $c(\tau) = c_w + c_f$. The worker chooses c_w to maximize utility:

$$u(\tau) = w(\tau) - pc_w = \beta f(\tau) + (1-\beta)v(\tau) - pc_f$$

subject to the same constraint.

The first-order conditions are:

$$(1-\beta)\left[f'(\tau) - v'(\tau)\right] - c'(\tau) = 0 \quad \text{if } c_f > 0 \tag{8.2}$$

$$v'(\tau) + \beta\left[f'(\tau) - v'(\tau)\right] - pc'(\tau) = 0 \quad \text{if } c_w > 0 \tag{8.3}$$

Inspection of these equations implies that generically, one of them will hold as a strict inequality, therefore, one of the parties will bear the full cost of training.

The result of no firm-sponsored investment in general training by the firm obtains when $f(\tau) = v(\tau)$, which is the case of perfectly competitive labor markets. (8.2) then implies that $c_f = 0$, so when workers receive their full marginal product in the outside labor market, the firm will never pay for training. Moreover, as $p \to \infty$, so that the worker is severely credit constrained, there will be no investment in training. In all cases, the firm is not constrained, so one dollar of spending on training costs one dollar for the firm.

In contrast, suppose there are labor market imperfections, so that the outside wage is less than the productivity of the worker, that is $v(\tau) < f(\tau)$. Is this gap between marginal product and market wage enough to ensure firm-sponsored investments in training? The answer is no. To see this, first consider the case with no *wage compression*, that is the case in which a marginal increase in skills is valued appropriately in the outside market. Mathematically this corresponds to $v'(\tau) = f'(\tau)$ for all τ. Substituting for this in the first-order condition of the firm, (8.2), we immediately find that if $c_f > 0$, then $c'(\tau) = 0$. So in other words, there will be no firm contribution to training expenditures.

Next consider the case in which there is wage compression, i.e., $v'(\tau) < f'(\tau)$. Now it is clear that the firm may be willing to invest in the general training of the worker. The simplest way to see this is again to consider the case of severe credit constraints on the worker, that is, $p \to \infty$, so that the worker cannot invest in training. Then, $v'(0) < f'(0)$ is sufficient to induce the firm to invest in training.

This shows the importance of *wage compression* for firm-sponsored training. The intuition is simple: wage compression in the outside market translates into wage compression inside the firm, i.e., it implies $w'(\tau) < f'(\tau)$. As a result, the firm makes greater profits from a more skilled (trained) worker, and has an incentive to increase the skills of the worker.

To clarify this point further, the figure draws the productivity, $f(\tau)$, and wage, $w(\tau)$, of the worker. The gap between these two curves is the sector-period profit of the firm. When $f'(\tau) = w'(\tau)$, this profit is independent of the skill level of the worker, and the firm has no interest in increasing the worker's skill. A competitive labor market, $f(\tau) = v(\tau)$, implies this case. In contrast, if $f'(\tau) > w'(\tau)$, which follows is a direct implication of $f'(\tau) > v'(\tau)$ given Nash bargaining, the firm makes more profits from more skilled workers, and is willing to invest in the general skills of its employees.

Let τ_w be the level of training that satisfies (8.3) as equality, and τ_f be the solution to (8.2). Then, it is clear that if $\tau_w > \tau_f$, the worker will bear all the cost of training. And if $\tau_f > \tau_w$, then the firm will bear all the cost of training (despite the fact that the worker may have access to perfect capital markets, i.e. $p = 1$).

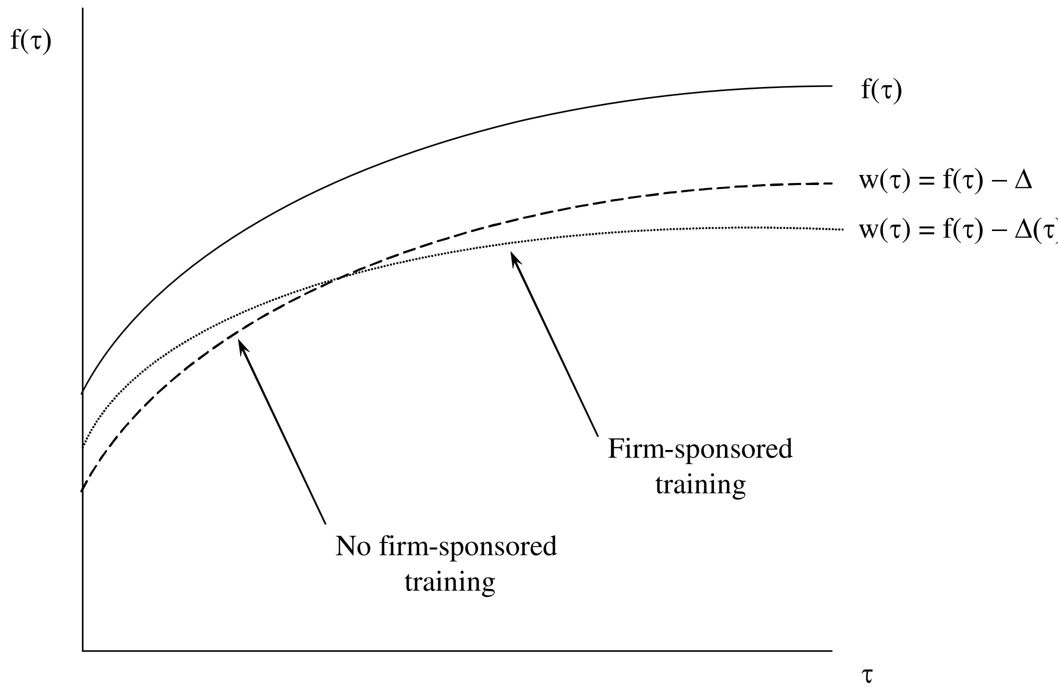


Figure 8.1

To derive the implications of changes in the skill premium on training, let $v(\tau) = af(\tau) - b$. A decrease in a is equivalent to a decrease in the price of skill in the outside market, and would also tilt the wage function inside the firm, $w(\tau)$, decreasing the relative wages of more skilled workers because of bargaining between the firm and in the worker, with the outside wage $v(\tau)$ as the threat point of the worker. Starting from $a = 1$ and $p < \infty$, a point at which the worker makes all investments, a decrease in a leads to less investment in training from (8.3). This is simply an application of the Becker reasoning; without any wage compression, the worker is the one receiving all the benefits and bearing all the costs, and a decline in the returns to training will reduce his investments.

As a declines further, we will eventually reach the point where $\tau_w = \tau_f$. Now the firm starts paying for training, and a further decrease in a increases investment in general training (from (8.2)). Therefore, there is a U-shaped relation between the skill premium and training—starting from a compressed wage structure, a further decrease in the skill premium may increase training. Holding $f(\tau)$ constant a tilting up of the wage schedule, $w(\tau)$, reduces the profits from more skilled workers, and the firm has less interest in investing in skills.

Changes in labor market institutions, such as minimum wages and unionization, will therefore affect the amount of training in this economy. To see the impact of a minimum wage, consider the next figure, and start with a situation where $v(\tau) = f(\tau) - \Delta$ and $p \to \infty$ so that the worker cannot invest in training, and there will be no training. Now impose a minimum wage as drawn in the figure. This distorts the wage structure and encourages the firm to invest in skills up to τ^*, as long as $c(\tau^*)$ is not too high. This is because the firm makes higher profits from workers with skills τ^* than workers with skills $\tau = 0$.

This is an interesting comparative static result, since the standard Becker model with competitive labor markets implies that *minimum wages should always reduce training*. The reason for this is straightforward. Workers take wage cuts to finance their general skills training, and minimum wages will prevent these wage cuts, thus reducing training. We will discuss this issue further below.

5. General Equilibrium with Imperfect Labor Markets

The above analysis showed how in imperfect labor market firms will find it profitable to invest in the general skills of their employees as long as the equilibrium wage structure is compressed. The equilibrium wage structure will be compressed, in turn, when the outside wage structure, $v(\tau)$, is compressed—that is, when $v'(\tau) < f'(\tau)$. The analysis was partial equilibrium in that this outside wage structure was taken as given. There are many reasons why in frictional labor markets we may expect this outside wage structure to be compressed. These include adverse selection, bargaining, and efficiency wages, as well as complementarity between general and specific skills. Here we will discuss how adverse selection leads to wage compression.

5.1. The Basic Model of Adverse Selection and Training. This is a simplified version of the model in Acemoglu and Pischke (1998). Suppose that fraction p of workers are high ability, and have productivity $\alpha(\tau)$ in the second period if they receive training τ in the first period. The remaining $1-p$ are low ability and produce nothing (in terms of the above model, we are setting $y=0$).

No one knows the worker's ability in the first period, but in the second period, the current employer learns this ability. Firms never observe the ability of the workers they have not employed, so outsiders will have to form beliefs about the worker's ability.

The exact timing of events is as follows:

- Firms make wage offers to workers. At this point, worker ability is unknown.
- Firms make training decisions, τ.
- Worker ability is revealed to the current employer and to the worker.
- Employers make second period wage offers to workers.
- Workers decide whether to quit.
- Outside firms compete for workers in the "secondhand" labor market. At this point, these firms observe neither worker ability nor whether the worker has quit or was laid off.
- Production takes place.

Since outside firms do not know worker ability when they make their bids, this is a (dynamic) game of incomplete information. So we will look for a Perfect Bayesian Equilibrium of this game, which is defined in the standard manner. We will characterize equilibria using backward induction conditional on beliefs at a given information set.

First, note that all workers will leave their current employer if outside wages are higher. In addition, a fraction λ of workers, irrespective of ability, realize that they form a bad match with the current employer, and leave whatever the wage is. The important assumption here is that firms in the outside market observe neither worker ability nor whether a worker has quit or has been laid off. However, worker training is publicly observed (what would happen to the model is training was not observed by outside employers?).

These assumptions ensure that in the second period each worker obtains his expected productivity *conditional* on his training. That is, his wage will be independent of his own productivity, but will depend on the average productivity of the workers who are in the secondhand labor market.

By Bayes's rule, the expected productivity of a worker of training τ, is

$$v(\tau) = \frac{\lambda p \alpha(\tau)}{\lambda p + (1-p)} \tag{8.4}$$

To see why this expression applies, note that all low ability workers will leave their initial employer, who will at most pay a wage of 0 (since this is the productivity of a low ability worker), and as we will see, outside wages are positive, low ability workers will quit (therefore, the offer of a wage of 0 is equivalent to a layoff; can there exist in equilibrium in which workers receive zero wage and stay at their job?). Those workers make up a fraction $1-p$ of the total workforce. In addition, of the high ability workers who make up a fraction p of the total workforce, a fraction λ of them will also leave. Therefore, the total size of the secondhand labor market is $\lambda p + (1-p)$, which is the denominator of (8.4). Of those, the low ability ones produce nothing, whereas the λp high ability workers produce $\alpha(\tau)$, which explains this expression.

Anticipating this outside wage, the initial employer has to pay each high ability worker $v(\tau)$ to keep him. This observation, combined with (8.4), immediately implies that there is wage compression in this world, in the sense that

$$v'(\tau) = \frac{\lambda p \alpha'(\tau)}{\lambda p + (1-p)} < \alpha'(\tau),$$

so the adverse selection problem introduces wage compression, and via this channel, will lead to firm-sponsored training.

To analyze this issue more carefully, consider the previous stage of the game. Now firm profits as a function of the training choice can be written as

$$\pi(\tau) = (1-\lambda)\, p\,[\alpha(\tau) - v(\tau)] - c(\tau).$$

The first-order condition for the firm is

$$\begin{aligned} \pi'(\tau) &= (1-\lambda)\, p\,[\alpha'(\tau) - v'(\tau)] - c'(\tau) = 0 \\ &= \frac{(1-\lambda)\, p\,(1-p)\,\alpha'(\tau)}{\lambda p + (1-p)} - c'(\tau) = 0 \end{aligned} \tag{8.5}$$

There are a number of noteworthy features:

(1) $c'(0) = 0$ is sufficient to ensure that there is firm-sponsored training (that is, the solution to (8.5) is interior).

(2) There is underinvestment in training relative to the first-best which would have involved $p\alpha'(\tau) = c'(\tau)$ (notice that the first-best already takes into account that only a fraction p of the workers will benefit from training). This is because of two reasons: first, a fraction λ of the high ability workers quit, and the firm does not get any profits from them. Second, even for the workers who stay, the firm is forced to pay them a higher wage, because they have an outside option that improves with their training, i.e., $v'(\tau) > 0$. This reduces profits from training, since the firm has to pay higher wages to keep the trained workers.

(3) The firm has *monopsony power* over the workers, enabling it to recover the costs of training. In particular, high ability workers who produce $\alpha(\tau)$ are paid $v(\tau) < \alpha(\tau)$.

(4) Monopsony power is not enough by itself. Wage compression is also essential for this result. To see this, suppose that we impose there is no wage compression, i.e., $v'(\tau) = \alpha'(\tau)$, then inspection of the first line of (8.5) immediately implies that there will be zero training, $\tau = 0$.

(5) But wage compression is also not automatic; it is a consequence of some of the assumptions in the model. Let us modify the model so that high ability workers produce $\eta + \alpha(\tau)$ in the second period, while low ability workers

produce $\alpha(\tau)$. This modification implies that training and ability are no longer complements. Both types of workers get exactly the same marginal increase in productivity (this contrasts with the previous specification where only high ability workers benefited from training, hence training and ability were highly complementary). Then, it is straightforward to check that we will have

$$v(\tau) = \frac{\lambda p \eta}{\lambda p + (1-p)} + \alpha(\tau),$$

and hence $v'(\tau) = \alpha'(\tau)$. Thus no wage compression, and firm-sponsored training. Intuitively, the complementarity between ability and training induces wage compression, because the training of high ability workers who are contemplating to leave their firm is judged by the market as the training of a relatively low ability worker (since low ability workers are overrepresented in the secondhand labor market). Therefore, the marginal increase in a (high ability) worker's productivity due to training is valued less in the outside market, which views this worker, on average, as low ability. Hence the firm does not have to pay as much for the marginal increase in the productivity of a high ability worker, and makes greater profits from more trained high-ability workers.

(6) What happens if

$$\pi(\tau) = (1-\lambda) p [\alpha(\tau) - v(\tau)] - c(\tau) > 0,$$

that is, if firms are making positive profits (at the equilibrium level of training)? If there is free entry at time $t = 0$, this implies that firms will compete for workers, since hiring a worker now guarantees positive profits in later periods. As a result, firms will have to pay a positive wage at time $t = 0$, precisely equal to

$$W = \pi(\tau)$$

as a result of this competition. This is because once a worker accepts a job with a firm, the firm acquires monopsony power over this worker's labor services at time $t = 1$ to make positive profits. Competition then implies that these profits have to be transferred to the worker at time $t = 0$. The interesting result is that not only do firms pay for training, but they may also pay workers extra in order to attract them.

5.2. Evidence. How can this model be tested? One way is to look for evidence of this type of adverse selection among highly trained workers. The fact that employers know more about their current employees may be a particularly good assumption for young workers, so a good area of application would be for apprentices in Germany.

According to the model, workers who quit or are laid off should get lower wages than those who stay in their jobs, which is a prediction that follows simply from adverse selection (and Gibbons and Katz tested in the U.S. labor market for all workers by comparing laid-off workers to those who lost their jobs as a result of plant closings). The more interesting implication here is that if the worker is separated from his firm for an exogenous reason that is clearly observable to the market, he should not be punished by the secondhand labor market. In fact, he's "freed" from the monopsony power of the firm, and he may get even higher wages than stayers (who are on average of higher ability, though subject to the monopsony power of their employer).

To see this, note that a worker who is exogenously separated from his firm will get to wage of $p\alpha(\tau)$ whereas stayers, who are still subject of the monopsony power of their employer, obtain the wage of $v(\tau)$ as given by (8.4), which could be less than $p\alpha(\tau)$. In the German context, workers who leave their apprenticeship firm to serve in the military provide a potential group of such exogenous separators. Interestingly, the evidence suggests that although these military quitters are on average lower ability than those who stay in the apprenticeship firm, the military quitters receive higher wages.

5.3. Mobility, training and wages. The interaction between training and adverse selection in the labor market also provides a different perspective in thinking about mobility patterns. To see this, change the above model so that $\lambda = 0$, but workers now quit if

$$w(\tau) - v(\tau) < \theta$$

where θ is a worker-specific draw from a uniform distribution over $[0, 1]$. θ, which can be interpreted as the disutility of work in the current job, is the worker's private information. This implies that the fraction of high ability workers who quit their initial employer will be

$$1 - w(\tau) + v(\tau),$$

so the outside wage is now

$$v(\tau) = \frac{p[1 - w(\tau) + v(\tau)]\,\alpha(\tau)}{p[1 - w(\tau) + v(\tau)] + (1 - p)} \tag{8.6}$$

Note that if $v(\tau)$ is high, many workers leave their employer because outside wages in the secondhand market are high. But also the right hand side of (8.6) is increasing in the fraction of quitters, $[1 - w(\tau) + v(\tau)]$, so $v(\tau)$ will increase further. This reflects the fact that with a higher quit rate, the secondhand market is not as adversely selected (it has a better composition).

This implies that there can be multiple equilibria in this economy. One equilibrium with a high quit rate, high wages for workers changing jobs, i.e. high $v(\tau)$, but low training. Another equilibrium with low mobility, low wages for job changers, and high training. This seems to give a stylistic description of the differences between the U.S. and German labor markets. In Germany, the turnover rate is much lower than in the U.S., and also there is much more training. Also, in Germany workers who change jobs are much more severely penalized (on average, in Germany such workers experience a substantial wage loss, while they experience a wage gain in the U.S.).

Which equilibrium is better? There is no unambiguous answer to this question. While the low-turnover equilibrium achieves higher training, it does worse in terms of matching workers to jobs, in that workers often get stuck in jobs that they do not like. In terms of the above model, we can see this by looking at the average disutility of work that workers receive (i.e., the average θ's).

5.4. Adverse selection and training in the temporary help industry. An alternative place to look for evidence is the temporary help industry in the U.S. Autor (2001) develops an extended version of this model, which also incorporates self-selection by workers, for the temporary help industry. Autor modifies the above model in four respects to apply it to the U.S. temporary help industry. These are:

(1) The model now lasts for three periods, and in the last period, all workers receive their full marginal products. This is meant to proxy the fact that at some point temporary-help workers may be hired into permanent jobs where their remuneration may better reflect their productivity.
(2) Workers have different beliefs about the probability that they are high ability. Some workers receive a signal which makes them believe that they are high ability with probability p, while others believe that they are high ability with probability $p' < p$. This assumption will allow self-selection among workers between training and no-training firms.
(3) Worker ability is only learned via training. Firms that do not offer training will not have superior information relative to the market. In addition, in contrast to the baseline version of the above model, it is also assumed that firms can offer different training levels and commit to them, so firms can use training levels as a method of attracting workers.
(4) The degree of competitiveness in the market is modeled by assuming that firms need to make a certain level of profits π, and a higher π corresponds to a less competitive market.

Autor looks for a "separating"/self-selection equilibrium in which p' workers select into no-training firms, whereas p workers go to training firms. In this context, self-selection equilibrium is one in which workers with different abilities (different beliefs) choose to accept jobs in different firms, because ability is rewarded differentially in different firms. This makes sense since training and ability are complements as before. Since firms that do not train their employees do not learn about employability, there is no adverse selection for workers who quit from no-training firms. Therefore, the second-period wage of workers who quit from no-training firms will be simply

$$v(0) = p'\alpha(0)$$

In contrast, the secondhand labor market wage of workers from training firms will be given by $v(\tau)$ from (8.4) above.

In the third period, all workers will receive their *expected* full marginal product. For workers who were employed by the non-training firms (and thus would did not receive training), this is $p'\alpha(0)$, whereas for workers with training, it is $p\alpha(\tau)$.

In the second-period, all workers receive their outside option in the secondhand market, so $v(0)$ for workers in no-training firms, and $v(\tau)$ for workers in training firms.

The condition for a self-selection equilibrium is

$$p(\alpha(\tau) - \alpha(0)) > v(0) - v(\tau) > p'(\alpha(\tau) - \alpha(0)),$$

that is, expected gain of third-period wages for high-belief workers should outweigh the loss (if any) in terms of second period wages (since there are no costs in the first-period by the assumption that there are no wages in the first-period). Otherwise, there could not be a separating equilibrium.

This immediately implies that if $v(0) - v(\tau) < 0$, that is, if workers with training receive higher wages in the second period, then there cannot be a self-selection equilibrium—all workers, irrespective of their beliefs, would like to take a job with training firms. Therefore, the adverse selection problem needs to be strong enough to ensure that $v(0) - v(\tau) > 0$. This is the first implication that Autor investigates empirically using data about the wages of temporary help workers in firms that offer free training compared to the wages of workers in firms that do not offer training. He finds that this is generally the case.

The second implication concerns the impact of greater competition on training. To see this more formally, simply return to the basic model, and look at the profits of a typical training firm. These are

$$\pi(\tau) = \frac{(1-\lambda)\,p\,(1-p)\,\alpha(\tau)}{\lambda p + (1-p)} - c(\tau).$$

Therefore, if in equilibrium we must have $\pi(\tau) = \pi$ for some exogenous level of profits π, and π increases exogenously, the training level offered by training firms must increase. To see this, note that in equilibrium we could never have $\pi'(\tau) > 0$, since then the firm can increase both its profits and attract more workers by simply increasing training. Therefore, the equilibrium must feature $\pi'(\tau) \leq 0$, and thus a decline in π, that is, increasing competitiveness, will lead to higher training.

Autor investigates this empirically using differences in temporary help firms concentration across MSAs, and finds that in areas where there is greater concentration, training is lower.

5.5. Labor market institutions and training. The theory developed here also implies that changes in labor market institutions, such as minimum wages and unionization, will therefore affect the amount of training in this economy. To see the impact of a binding minimum wage on training, let us return to the baseline framework and consider the next figure, and start with a situation where $v(\tau) = f(\tau) - \Delta$ and $p \to \infty$ so that the worker cannot invest in training, and there will be no training. Now impose a minimum wage as drawn in the figure. This distorts the wage structure and encourages the firm to invest in skills up to τ^*, as long as $c(\tau^*)$ is not too high. This is because the firm makes higher profits from workers with skills τ^* than workers with skills $\tau = 0$.

This is an interesting comparative static result, since the standard Becker model with competitive labor markets implies that *minimum wages should always reduce training.* The reason for this is straightforward. Workers take wage cuts to finance their general skills training, and minimum wages will prevent these wage cuts, thus reducing training.

Therefore, an empirical investigation of the relationship between minimum wage changes and worker training is a way of finding out whether the Becker channel

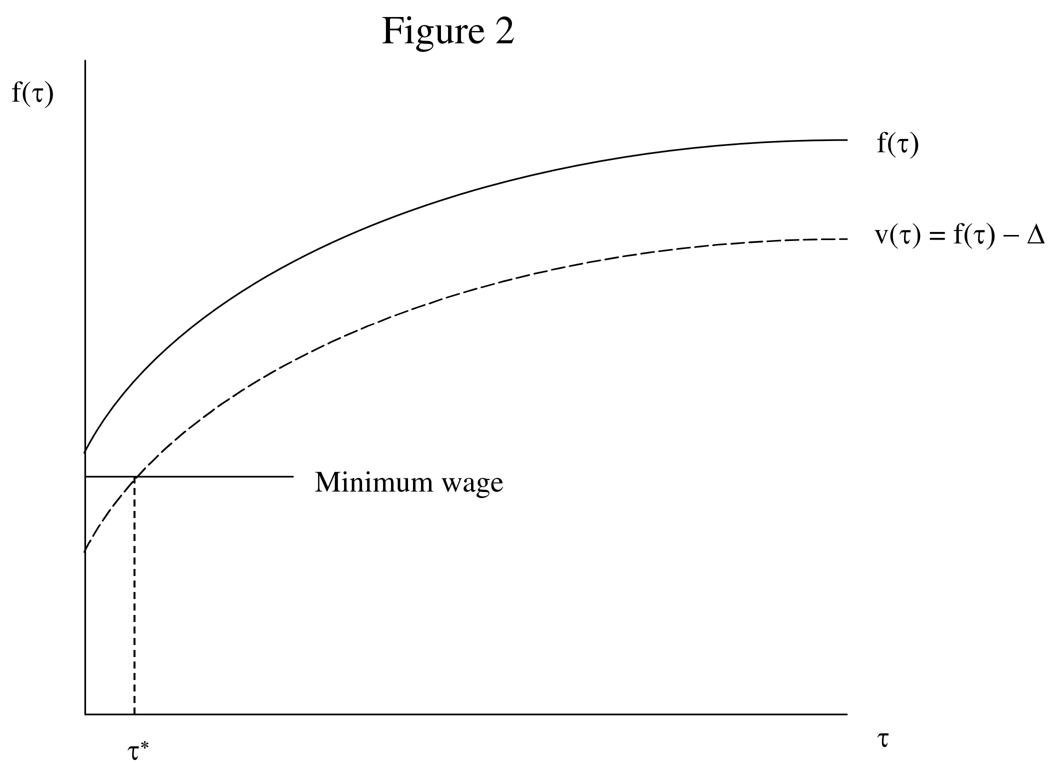


FIGURE 8.2

or the wage-compression channel is more important. Empirical evidence suggests that higher minimum wages are typically associated with more training for low-skill workers (though this relationship is not always statistically significant).

Table 9
The Effect of Minimum Wage Increases on Affected Workers

Independent Variable	*(1)*	*(2)*	*(3)*
Minimum wage increased and wage in prior year is below the current minimum wage	0.010 (0.014)	--	--
Minimum increased and wage in prior year is below the current minimum and above prior year minimum	--	0.016 (0.019)	--
Minimum wage increased and wage in prior year is below 150 % of the current minimum wage	--	--	0.005 (0.008)
Minimum wage increased and wage in prior year is below 130% of the current minimum wage	--	--	--
Change in high school graduation status	0.070 (0.053)	0.070 (0.054)	0.070 (0.054)
Change in new job status	0.032 (0.008)	0.032 (0.008)	0.032 (0.008)

Notes: Basic sample with all workers with high school education or less, who do not move between stat
the next. Dependent variable is the change in training incidence between two consecutive years. All re
include a constant and year dummies. Regressions are weighted by NSLY sampling weights. Standard
for the presence of state*time effects in the error term, and therefore robust to the MA structure of the e
17101

FIGURE 8.3

CHAPTER 9

Firm-Specific Skills and Learning

The analysis so far has focused on general skills, acquired in school or by investments in general training. Most labor economists also believe that there are also important firm-specific skills, acquired either thanks to firm-specific experience, or by investment in firm-specific skills, or via "matching". If such firm-specific skills are important we should observe worker productivity and wages to increase with *tenure*—that is, a worker who has stayed longer in a given job should earn more than a comparable worker (with the same schooling and experience) who has less tenure.

1. The Evidence On Firm-Specific Rents and Interpretation

1.1. Some Evidence. The empirical investigation of the importance of firm-specific skills and rents is a difficult and challenging area. There are two important conceptual issues that arise in thinking about the relationship between wages and tenure, as well as a host of econometric issues. The conceptual issues are as follows:

(1) We can imagine a world in which firm-specific skills are important, but there may be no relationship between tenure and wages. This is because, as we will see in more detail below, productivity increases due to firm-specific skills do not necessarily translate into wage increases. The usual reasoning for why high worker productivity translates into higher wages is that otherwise, competitors would bid for the worker and steal him. This argument does not apply when skills are firm-specific since such skills do not contribute to the worker's productivity in other firms. More generally, the

relationship between productivity and wages is more complex when firm-specific skills are a significant component of productivity. For example, we might have two different jobs, one with faster accumulation of firm-specific skills, but wages may grow faster in the other job because the outside option of the worker is improving faster.

(2) An empirical relationship between tenure and wages does not establish that there are imported from-specific effects. To start with, wages may increase with tenure because of *backloaded* compensation packages, which, as we saw above, are useful for dealing with moral hazard problems. Such a relationship might also result from the fact that there are some jobs with high "rents," and workers who get these jobs never quit, creating a positive relationship between tenure and wages. Alternatively, a positive relationship between tenure and wages may reflect the fact that high ability workers stay in their jobs longer (selection).

The existing evidence may therefore either overstate or understate the importance of tenure and firm-specific skills, and there are no straightforward ways of dealing with these problems. In addition, there are important econometric problems, for example, the fact that in most data sets most tenure spells are uncompleted (most workers are in the middle of their job tenure), complicating the analysis. A number of researchers have used the usual strategies, as well as some creative strategies, to deal with the selection and omitted variable biases, pointed out in the second problem. But is still requires us to ignore the first problem (i.e., be cautious in inferring the tenure-productivity relationship from the observed tenure-wage relationship).

In any case, the empirical relationship between tenure and wages is of interest in its own right, even if we cannot immediately deduce from this the relationship between tenure and firm-specific productivity.

With all of these complications, the evidence nevertheless suggests that there is a positive relationship between tenure and wages, consistent with the importance of firm-specific skills. Here we will discuss two different types of evidence.

The first type of evidence is from regression analyses of the relationship between wages and tenure exploiting within job wage growth. Here the idea is that by looking at how wages grow within a job (as long as the worker does not change jobs), and comparing this to the experience premium, we will get an estimate of the tenure premium. In other words, we can think of wages as given by the following model

$$\ln w_{it} = \beta_1 X_{it} + \beta_2 T_{it} + \varepsilon_{it} \tag{9.1}$$

where X_{it} this total labor market experience of individual i, and T_{it} is his tenure in the current job. Then, we have that his wage growth on this job is:

$$\Delta \ln w_{it} = \beta_1 + \beta_2 + \Delta \varepsilon_{it}$$

If we knew the experience premium, β_1, we could then immediately compute the tenure premium β_2. The problem is that we do not know the experience premium. Topel suggests that we can get an upper bound for the experience premium by looking at the relationship between entry-level wages and labor market experience (that is, wages in jobs with tenure equal to zero). This is an upper bound to the extent that workers do not randomly change jobs, but only accept new jobs if these offer a relatively high wage. Therefore, whenever $T_{it} = 0$, the disturbance term ε_{it} in (9.1) is likely to be positively selected. According to this reasoning, we can obtain a *lower bound* estimate of β_2, $\hat{\beta}_2$, using a two-step procedure—first estimate the rate of within-job wage growth, $\hat{\beta}_1$, and then subtract from this the estimate of the experience premium obtained from entry-level jobs (can you see reasons why this will lead to an upwardly biased estimate of the importance of tenure rather than a lower bound on tenure affects as Topel claims?).

Using this procedure Topel estimates relatively high rates of return to tenure. For example, his main estimates imply that ten years of tenure increase wages by about 25 percent, over and above the experience premium.

It is possible, however, that this procedure might generate tenure premium estimates that are upward biased. For example, this would be the case if the return to tenure or experience is higher among high-ability workers, and those are under-represented among the job-changers. Alternatively, returns to experience may be non-constant, and they may be higher in jobs to which workers are a better match. If this is the case, returns experience for new jobs will understate the average returns to experience for jobs in which workers choose to stay.

On the other hand, the advantage of this evidence is that it is unlikely to reflect simply the presence of some jobs that offer high-rents to workers, unless these jobs that provide high rents also have (for some reason) higher wage growth (one possibility might be that, union jobs pay higher wages, and have higher wage growth, and of course, workers do not leave union jobs, but this seems unlikely).

The second type of evidence comes from the wage changes of workers resulting from job displacement. A number of papers, most notably Jacobson, LaLonde and Sullivan, find that displaced workers experience substantial drop in earnings. This is shown in the next figure.

Part of this is due to non-employment following displacement, but even after three years a typical displaced worker is earning about \$1500 less (1987 dollars). Econometrically, this evidence is simpler to interpret than the tenure-premium estimates. Economically, the interpretation is somewhat more difficult than the tenure estimates, since it may simply reflect the loss of high-rent (e.g. union) jobs.

In any case, these two pieces of evidence together are consistent with the view that there are important firm-specific skills/expertises that are accumulated on the job.

1.2. What Are Firm-Specific Skills? If we are going to interpret the above evidence as reflecting the importance of firm-specific skills, then we have to be more specific about what constitutes firm-specific skills. Here are four different views:

(1) Firm-specific skills can be thought to result mostly from firm-specific training investments made by workers and firms. Here it is important to distinguish between firms' and workers' investments, since they will have different incentives.

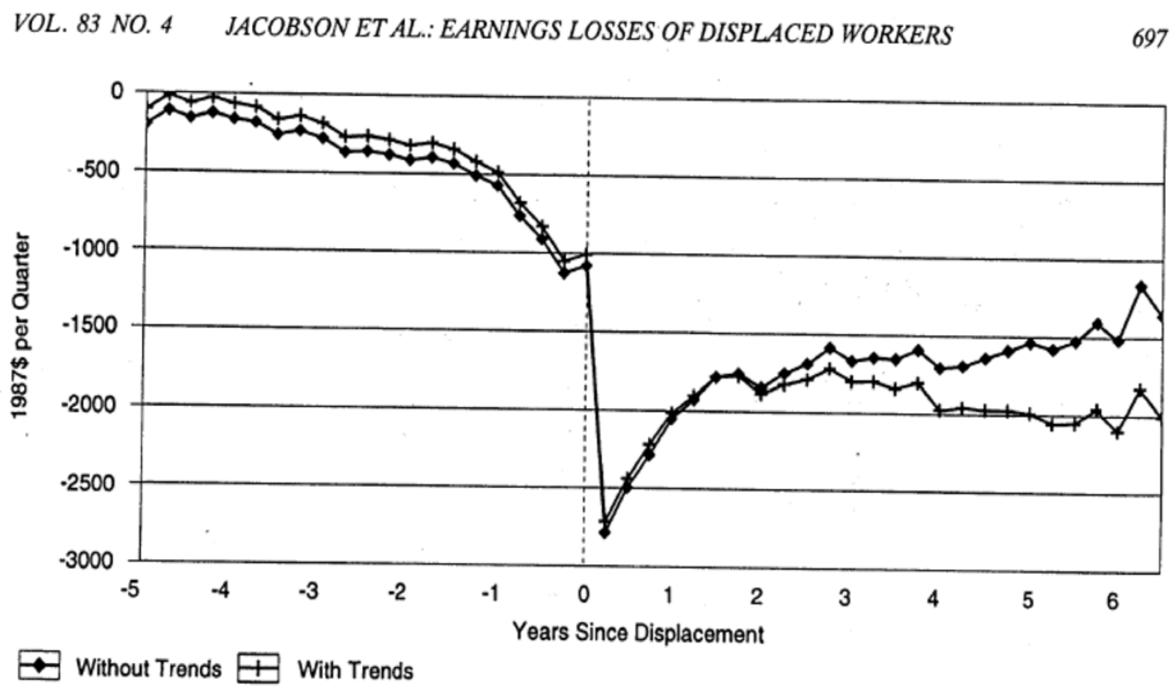


FIGURE 2. EARNINGS LOSSES FOR SEPARATORS IN MASS-LAYOFF SAMPLE

FIGURE 9.1

(2) Firm-specific skills simply reflect what the worker learns on-the-job without making any investments. In other words, they are simply unintentional byproducts of working on the job. The reason why it is useful to distinguish this particular view from the firm-specific investments view is that according to this view, we do not need to worry about the incentives to acquire firm-specific skills. However, most likely, even for simple skills that workers can acquire on-the-job, they need to exert some effort, so this view may have relatively little applicability.

(3) Firm-specific skills may reflect "matching" as in Jovanovic's approach. Here, there is no firm-specific skill, but some workers are better matches to some firms. Ex ante, neither the firm nor the worker knows this, and the information is revealed only slowly. Only workers who are revealed to be good

matches to a particular job will stay on that job, and as a result, they will be more productive in this job than a randomly chosen worker. We can think of this process of learning about the quality of the match as the "accumulation of firm-specific skills".

(4) There may be no technologically firm-specific skills. Instead, you may think of all skills as technologically general, in the sense that if the worker is more productive in a given firm, another firm that adopts exactly the same technologies and organizational structure, and hires the same set of co-workers will also be able to benefit from this high productivity. These technologically general skills are transformed into *de facto* firm-specific skills because of market imperfections. For example, if worker mobility is costly, or if it is difficult or unprofitable for firms to copy some other firms' technology choices, these skills will be de facto specific to the firm that has first made the technology/organizational choices. But if this is the case, we are back to the model of general training investments under imperfect markets we studied above. The reason why it is important to distinguish this view of de facto firm-specific skills from the first view above is that now changes in technology/market organization will affect which skills are specific and how much of a given bundle of technologically-determined skills are "specific".

2. Investment in Firm-Specific Skills

2.1. The basic problem. The problem with general training investments was that part of the costs had to be borne by the firm, but, at least in competitive labor markets, the worker was the residual claimant. The worker, in turn, was the residual claimant because the skills were general, and other firms could compete for this worker's labor services. In contrast, with specific skills, the current employer is the only (or at least the main) "consumer," so there is no competition from other firms to push up the worker's wages. As a result, firm-specific skills will make the firm the ex post monopsonist. This creates the converse problem. Now the worker also bears some (perhaps most) of the costs of investment, but may not have the right incentives to invest, since the firm will get most of the benefits.

To capture these problems, consider the following very simple model:

- At time $t = 0$, the worker decides how much to invest in firm-specific skills, denoted by s, at the cost $\gamma(s)$. $\gamma(s)$ is strictly increasing and convex, with $\gamma'(0) = 0$.
- At time $t = 1$, the firm makes a wage offer to the worker.
- The worker decides whether to accept this wage offer and work for this firm, or take another job.
- Production takes place and wages are paid.

Let the productivity of the worker be $y_1 + f(s)$ where y_1 is also what he would produce with another firm. Since s is specific skills, it does not affect the worker's productivity in other firms.

First, note that the first-best level of firm-specific skills is given by

$$\gamma'(s^*) = f'(s^*).$$

Here s^* is strictly positive since $\gamma'(0) = 0$.

Let us next solve this game by backward induction again, starting in the last period. The worker will accept any wage offer $w_1 \geq y_1$, since this is what he can get in an outside firm. Knowing this, the firm simply offers $w_1 = y_1$. In the previous period, realizing that his wage is independent of his specific skills, the worker makes no investment in specific skills, even though the first best level of firm-specific skills s^* is strictly positive.

What is the problem here? By investing in his firm-specific skills, the worker is increasing the firm's profits. Therefore, the firm would like to encourage the worker to invest. However, given the timing of the game, wages are determined by a take-it-leave-it offer by the firm after the investment. Therefore, it will always be in the interest of the firm to offer a low wage to the worker after the investment, in other words, the firm will *hold* the worker *up*. The worker anticipates this holdup problem and does not invest in his firm-specific skills.

Why is there not a contractual solution to this underinvestment problem? For example, the firm could write a contract ex ante promising a certain payment to the worker. Leaving aside the problems of enforcing such contracts (the firm could always try to fire the worker, or threaten to fire him), there is and more fundamental problem. If the employment contract does *not* make the wage of the worker conditional on his firm-specific skills, it will not encourage investment. So the only contracts that could help with the underinvestment problem are those that make the worker's wages contingent on his firm-specific skills. However, such skills are very difficult to observe or verify by outside parties. This motivates the assumption in this literature, as well as in the incomplete contracts literature, that such contingent contracts cannot be written (they cannot be enforced, and hence are useless). Therefore, contractual solutions to the underinvestment problem are difficult to devise

As a result, there is a severe underinvestment problem here, driven by exactly the converse of the underinvestment problem in general training. The worker will not undertake the required investments, because he's afraid of being held up by the firm.

2.2. Worker power and investment. How can we improve the worker's investment incentives?

At a very general level, the answer is simple. The worker's earnings have to be conditioned on his specific skills. There are a number of ways of achieving this. Perhaps the simplest is to give the worker some "power" in the employment relationship. This power may come simply because the worker can bargain with his employer effectively (either individually or via unions—though the latter would probably be not useful in this context, since union bargaining does not typically will link a worker's wage to his productivity). The worker may be able to bargain with the firm, in turn, for a variety of reasons. Here are some:

(1) Because of regulations, such as employment protection legislation, or precisely because of his specific skills, the firm needs the worker, hence we are

in the bilateral monopoly situation, and the rents will be shared (rather than the firm making a take-it-leave-it offer).

(2) The firm may purposefully give access to some important assets of the firm to the worker, so that the worker may feel secure that he will not be held up. This is basically the insight that follows from the incomplete contracting approach to property rights, which we discussed previously. Recall that in the Grossman-Hart-Moore approach to the internal organization of the firm, the allocation of property rights determine who can use assets and the use of the firm's assets is a way of manipulating ex post bargaining and via this channel ex ante investment incentives.

(3) The firm may change its organizational form in order to make a credible commitment not to hold up the worker.

(4) The firm may develop a reputation for not holding up workers who have invested in firm-specific human capital.

Here let us consider a simple example of investment incentives with bargaining power, and show why firms may preferred to give more bargaining power to their employees in order to ensure high levels of firm-specific investments. In the next section, we discuss alternative "organizational" solutions to this problem.

Modify the above game simply by assuming that in the final period, rather than the firm making a take-it-leave-it offer, the worker and the firm bargain over the firm-specific surplus, so the worker's wage is

$$w_1(s) = y_1 + \beta f(s)$$

Now at time $t = 0$, the worker maximizes

$$y_1 + \beta f(s) - \gamma(s),$$

which gives his investment as

$$\beta f'(\hat{s}) = \gamma'(\hat{s}) \tag{9.2}$$

Here $\hat{s}$ is strictly positive, so giving the worker bargaining power has improved investment incentives. However, $\hat{s}$ is strictly less than the first-best investment level s^*.

To investigate the relationship between firm-specific skills, firm profits and the allocation of power within firms, now consider an extended game, where at time $t = -1$, the firm chooses whether to give the worker access to a key asset. If it does, ex post the worker has bargaining power β, and if it does not, the worker has no bargaining power and wages are determined by a take-it-leave-it offer of the firm. Essentially, the firm is choosing between the game in this section and the previous one. Let us look at the profits of the firm from choosing the two actions. When it gives no access, the worker chooses zero investment, and since $w_1 = y_1$, the firm profits are $\pi_0 = 0$. In contrast, with the change in organizational form giving access to the worker, the worker undertakes investment $\hat{s}$, and profits are

$$\pi_\beta = (1 - \beta) f(\hat{s}).$$

Therefore, the firm would prefer to give the worker some bargaining power in order to encourage investment in specific skills.

Notice the contrast in the role of worker bargaining power between the standard framework and the one here. In the standard framework, worker bargaining power always reduces profits and causes inefficiency. Here, it may do the opposite. This suggests that in some situations reducing worker bargaining power may actually be counterproductive for efficiency.

Note another interesting implication of the framework here. If the firm could choose the bargaining power of the worker without any constraints, it would set $\bar{\beta}$ such that

$$\frac{\partial \pi_\beta}{\partial \beta} = 0 = -f\left(\hat{s}\left(\bar{\beta}\right)\right) + \left(1 - \bar{\beta}\right) f'\left(\hat{s}\left(\bar{\beta}\right)\right) \frac{d\hat{s}\left(\bar{\beta}\right)}{d\beta},$$

where $\hat{s}(\beta)$ and $d\hat{s}/d\beta$ are given by the first-order condition of the worker, (9.2).

One observation is immediate. The firm would certainly choose $\bar{\beta} < 1$, since with $\bar{\beta} = 1$, we could never have $\partial \pi_\beta / \partial \beta = 0$ (or more straightforwardly, profits would be zero). In contrast, a social planner who did not care about the distribution of income between profits and wages would necessarily choose $\beta = 1$. The reason why the firm would not choose the structure of organization that achieves the best investment outcomes is that it cares about its own profits, not total income or surplus.

If there were an ex ante market in which the worker and the firm could "transact', the worker could make side payments to the firm to encourage it to choose $\beta = 1$, then the efficient outcome would be achieved. This is basically the solution that follows from the analysis of the incomplete contracts literature discussed above, but this literature focuses on vertical integration, and attempts to answer the question of who among many entrepreneurs/managers should own the firm or its assets. In the context of worker-firm relationships, such a solution is not possible, given credit constraints facing workers. Perhaps more importantly, such an arrangement would effectively amount to the worker buying the firm, which is not possible for two important reasons:

- the entrepreneur/owner of the firm most likely has some essential knowledge for the production process and transferring all profits to workers or to a single worker is impractical and would destroy the value-generating capacity of the firm;
- in practice there are many workers, so it is impossible to improves their investment incentives by making each worker the residual claimant of the firm's profits.

2.3. Promotions. An alternative arrangement to encourage workers to invest in firm-specific skills is to design a promotion scheme. Consider the following setup. Suppose that there are two investment levels, $s = 0$, and $s = 1$ which costs c.

Suppose also that at time $t = 1$, there are two tasks in the firm, difficult and easy, D and E. Assume outputs in these two tasks as a function of the skill level are

$$y_D(0) < y_E(0) < y_E(1) < y_D(1)$$

Therefore, skills are more useful in the difficult task, and without skills the difficult task is not very productive.

Moreover, suppose that

$$y_D(1) - y_E(1) > c$$

meaning that the productivity gain of assigning a skilled worker to the difficult task is greater than the cost of the worker obtaining skills.

In this situation, the firm can induce firm-specific investments in skills if it can commit to a wage structure attached to promotions. In particular, suppose that the firm commits to a wage of w_D for the difficult task and w_E for the easy task. Notice that the wages do not depend on whether the worker has undertaken the investment, so we are assuming some degree of commitment on the side of the firm, but not modifying the crucial incompleteness of contracts assumption.

Now imagine the firm chooses the wage structures such that

$$y_D(1) - y_E(1) > w_D - w_E > c, \tag{9.3}$$

and then ex post decides whether the worker will be promoted.

Again by backward induction, we have to look at the decisions in the final period of the game. When it comes to the promotion decision, and the worker is unskilled, the firm will naturally choose to allocate him to the easy task (his productivity is higher in the easy task and his wage is lower). If the worker is skilled, and the firm allocates him to the easy task, his profits are $y_E(1) - w_E$. If it allocates him to the difficult task, his profits are $y_D(1) - w_D$. The wage structure in (9.3) ensures that profits from allocating him to the difficult task are higher. Therefore, with this wage structure the firm has made a credible commitment to pay the worker a higher wage if he becomes skilled, because it will find it profitable to promote the worker.

Next, going to the investment stage, the worker realizes that when he does not invest he will receive w_E , and when he invests, he will get the higher wage w_D. Since, again by (9.3), $w_D - w_E > c$, the worker will find it profitable to undertake the investment.

2.4. Investments and layoffs—The Hashimoto model. Consider the following model which is useful in a variety of circumstances. The worker can invest in $s = 1$ at time $t = 0$ again at the cost c. The investment increases the worker's productivity by an amount $m + \eta$ where η is a mean-zero random variable observed only by the firm at $t = 1$. The total productivity of the worker is $x + m + \eta$ (if he does not invest, his productivity is simply x). The firm unilaterally decides whether to fire the worker, so the worker will be fired if

$$\eta < \eta^* \equiv w - x - m,$$

where w is his wage. This wage is assumed to be fixed, and cannot be renegotiated as a function of η, since the worker does not observe η. (There can be other more complicated ways of revealing information about η, using stochastic contracts, whereby workers and firms make direct reports about the values of η and θ, and different values of these variables map into a wage level and a probability that the relationship will continue; using the Revelation Principle we can restrict attention to truthful reports subject incentive compatibility constraints and solve for the most efficient contracts of this form; nevertheless, to keep the discussion simple, we ignore these stochastic contracts here).

If the worker is fired or quits, he receives an outside wage v. If he stays, he receives the wage paid by the firm, w, and also disutility, θ, only observed by him. The worker unilaterally decides whether to quit or not, so he will quit if

$$\theta > \theta^* \equiv w - v$$

Denoting the distribution function of θ by Q and that of η by F, and assuming that the draws from these distributions are independent, the expected profit of the firm is

$$Q(\theta^*)[1 - F(\eta^*)][x + m - w + E(\eta \mid \eta \geq \eta^*)]$$

The expected utility of the worker is

$$v + Q(\theta^*)[1 - F(\eta^*)][w - v - E(\theta \mid \theta \leq \theta^*)]$$

In contrast, if the worker does not invest in skills, he will obtain

$$\begin{array}{rcl} v \text{ if } w & > & x \\ v + Q(\theta^*)[w - v - E(\theta \mid \theta \leq \theta^*)] \text{ if } w & \leq & x \end{array}$$

So we can see that a high wage promise by the firm may have either a beneficial or an adverse effect on investment incentives. If $w = x + \varepsilon > v$, the worker realizes that he can only keep his job by investing. But on the other hand, a high wage makes it more likely that $\eta < \eta^*$, so it may increase the probability that given the realization of the productivity shock, profits will be negative, and the worker will be fired. This will reduce the worker's investment incentives. In addition, a lower wage would make it more likely that the worker will quit, and through this channel increase inefficiency and discourage investment.

According to Hashimoto, the wage structure has to be determined to balance these effects, and moreover, the ex post wage structure chosen to minimize inefficient separations may dictate a particular division of the costs of firm-specific investments.

An interesting twist on this comes from Carmichael, who suggests that commitment to a promotion ladder might improve incentives to invest without encouraging further layoffs by the firm. Suppose the firm commits to promote N_h workers at time $t = 1$ (how such a commitment is made is an interesting and difficult question). Promotion comes with an additional wage of B. So the expected wage of the worker, if he keeps his job, is now

$$w + \frac{N_h}{N} B,$$

where N is employment at time $t = 1$, and this expression assumes that a random selection of the workers will be promoted. A greater N_h or B, holding the layoff rate of the firm constant, increases the incentive of the worker to stay around, and encourages investment.

Next think about the layoff rate of the firm. The total wage bill of the firm at time $t = 1$ is then

$$W = Nw + N_h B.$$

The significance of this expression is that if the firm fires a worker, this will only save the firm w, since it is still committed to promote N_h workers. Therefore, this commitment to (an absolute number of) promotions, reduces the firm's incentive to fire, while simultaneously increasing the reward to staying in the firm for the worker.

This is an interesting idea, but we can push the reasoning further, perhaps suggesting that it is not as compelling as it first appears. If the firm can commit to promote N_h workers, why can it not commit to employing N' workers, and by manipulating this number effectively make a commitment not to fire workers? So if this type of commitment to employment level is allowed, promotions are not necessary, and if such a commitment is not allowed, it is not plausible that the firm can commit to promoting N_h workers.

3. A Simple Model of Labor Market Learning and Mobility

An important idea related to firm-specific skills is that these skills are (at least in part) a manifestation of the quality of the match between a worker and his job. Naturally, if workers could costlessly learn about the quality of the matches between themselves and all potential jobs, they would immediately choose the job for which they are most suited to. In practice, however, jobs are "experience goods," meaning that workers can only find out whether they are a good match to a job (and to a firm) by working in that firm and job. Moreover, this type of learning does not take place immediately.

What makes these ideas particularly useful for labor economics is that a simple model incorporating this type of match-specific learning provides a range of useful results and also opens up even a larger set of questions for analysis. Interestingly, however, after the early models on these topics, there has been relatively little research.

The first model to formalize these ideas is due to Jovanovic. Jovanovic considered a model in which match-specific productivity is the draw from a normal distribution, and the output of the worker conditional on his match-specific productivity is also normally distributed. Though, as we have seen, normal distributions are often very convenient, in this particular context the normal distribution has a disadvantage, which is that as the worker learns about his match-specific productivity, we need to keep track of both his belief about the level of the quality and also the precision of his beliefs. This makes the model somewhat difficult to work with.

Instead, let us consider a simpler version of the same model.

Each worker is infinitely lived in discrete time and maximizes the expected discounted value of income, with a discount factor $\beta < 1$. There is no ex ante heterogeneity among the workers. But worker-job matches are random.

In particular, the worker may be a good match for a job (or a firm) or a bad match. Let the (population) probability that the worker is a good match be $\mu_0 \in (0,1)$. A worker in any given job can generate one of two levels of output, high, y_h, and low $y_l < y_h$. In particular, suppose that we have

$$\text{good match} \rightarrow \begin{array}{cc} y_h & \text{with probability } p \\ y_l & \text{with probability } 1-p \end{array}$$

and

$$\text{bad match} \rightarrow \begin{array}{cc} y_h & \text{with probability } q \\ y_l & \text{with probability } 1-q \end{array}$$

where, naturally,

$$p > q.$$

Let us assume that all learning is symmetric (as in the career concerns model). This is natural in the present context, since there is only learning about the match quality of the worker and the firm will also observe the productivity realizations of the worker since the beginning of their employment relationship. This implies that the firm and the worker will share the same *posterior probability* that the worker is a good match to the job. For worker i job j and time t, we can denote this posterior probability (belief) as μ_{ijt}. When there is no risk of confusion, we will denote this simply by μ.

Jovanovic assumes that workers always receive their full marginal product in each job. This is a problematic assumption, since match-specific quality is also firm specific, thus there is no reason for the worker to receive this entire firm-specific surplus. As in the models with firm-specific investments, the more natural assumption would be to have some type of wage bargaining. Let us assume the simplest bargaining structure in which a firm will pay the worker a fraction $\phi \in (0,1]$ of his expected productivity at that point. In particular, the wage of a worker whose posterior of a good match is μ will be

$$w(\mu) = \phi\left[\mu y_h + (1-\mu) y_l\right].$$

Note that this is different from the Nash bargaining solution, which would have to take into account the outside option and also the future benefits to the worker from being in this job (which result from learning). But having such a simple expression facilitates the analysis and the exposition here. [Alternatively, we could have assume that bargaining takes place after the realization of output, in which case the wage would be equal to ϕy_h with probability μ and to ϕy_l with probability $1-\mu$; since both the worker and affirm our risk neutral, there is no difference between these two cases].

To make progress, let us consider a worker with belief μ. If this worker produces output y_h, then Bayes's rule implies that his posterior (belief) next period should be

$$\mu'_h(\mu) \equiv \frac{\mu p}{\mu p + (1-\mu) q} > \mu,$$

where the fact that this is greater than μ immediately follows from the assumption that $p > q$. Similarly, following an output realization of y_l, the belief of the worker will be

$$\mu'_l(\mu) \equiv \frac{\mu(1-p)}{\mu(1-p) + (1-\mu)(1-p)} < \mu.$$

Finally, let us also assume that every time a worker changes jobs, he has to incur a training or mobility cost equal to $\gamma \geq 0$.

Under these assumptions, we can write the net present discounted value of a worker with belief μ recursively using simple dynamic programming arguments. In particular, this is

$$\begin{aligned} V(\mu) &= w(\mu) + \beta[(\mu p + (1-\mu) q) V(\mu_h'(\mu)) \\ &\quad + (\mu(1-p) + (1-\mu)(1-q)) \max\{V(\mu_l'(\mu)); V(\mu_0) - \gamma\}. \end{aligned}$$

Intuitively, the worker receives the wage $w(\mu)$ as a function of the (symmetrically held) belief about the quality of his match at the moment.

The continuation value, which is discounted with the discount factor $\beta < 1$, has the following explanation: with probability μ, the match is indeed good and then the worker will produce an output equal to y_h with probability p. With probability $1-\mu$, the match is not good and the worker will produce high output with probability q. In either case, the posterior about match quality will be $\mu_h'(\mu)$, and using the recursive reasoning, his value will be $V(\mu_h'(\mu))$. Since he was happy to be in this job with belief μ, $\mu_h'(\mu) > \mu$ as stated above, and clearly (can you prove this?) $V(\mu)$ is increasing in μ, he will not want to quit after a good realization and thus his value is written as $V(\mu_h'(\mu))$.

With probability $(\mu(1-p) + (1-\mu)(1-q))$, on the other hand, he will produce low output, y_l, and in this case the posterior will be $\mu_l'(\mu)$. Since $\mu_l'(\mu) < \mu$, at this point the worker may prefer to quit and take another job. Since a new job is a new draw from the match-quality distribution, the probability that he will be good at this job is μ_0. Subtracting the cost of mobility, γ, the value of taking a new job is therefore $V(\mu_0) - \gamma$. The worker chooses the maximum of this and this continuation value in the same firm, $V(\mu_l'(\mu))$.

An immediate result from dynamic programming is that if the instantaneous reward function, here $w(\mu)$, is strictly increasing in the state variable, which here is the belief μ, then the value function $V(\mu)$ will also be strictly increasing. This implies that there will exist some cutoff level of belief μ^* such that workers will stay in their job as long as

$$\mu \geq \mu^*,$$

and they will quit if $\mu < \mu^*$.

Let $\bar{\mu} = \inf \{\mu{:}\mu'_l(\mu) < \mu^*\}$. Then a worker with beliefs $\mu > \bar{\mu}$ will not quit irrespective of the realization of output. Workers with $\mu < \mu^*$ should have quit already. Therefore, the only remaining range of beliefs is $\mu \in [\mu^*, \bar{\mu}]$. A worker with beliefs in this range will quit the job if he generates low output.

Now a couple of observations are immediate.

(1) Provided that $\mu_0 \in (0,1)$, μ will never converge to 0 or 1 in finite time. Therefore, a worker who generates high output will have higher wages in the following period, and a worker who generates low output will have lower wages in the following period. Thus, in this model worker *wages will move with past performance.*

(2) It can be easily proved that if $\gamma = 0$, then $\mu^* = \mu_0$. This implies that when γ is equal to 0 or is very small, a worker who starts a job and generates low output will quit immediately. Therefore, as long as γ is not very high, there will be *a high likelihood of separation in new jobs.*

(3) Next consider a worker who has been in a job for a long time. Such workers will on average have high values of μ, since they have never experienced (on this job) a belief less than μ^*. This implies that the average value of their beliefs must be high. Therefore, *workers with long tenure are unlikely to quit or separate from their job.* [Here average refers to the average among the set of workers who have been in a job for a given length of time; for example, the average value of μ for all workers who have been a job for T periods].

(4) With the same argument,workers who have been in a job for a long time will have high average μ and thus high wages. This implies that in equilibrium *there will be a tenure premium.*

(5) Moreover, because Bayesian updating immediately implies that the gaps between $\mu'_h(\mu)$ and μ and between $\mu'_l(\mu)$ and μ are lowest when μ is close to 1 (and symmetrically when it is close to 0, but workers are never in jobs where their beliefs are close to 0), workers with long tenure will not

experience large wage changes. In contrast, *workers at the beginning of their tenure will have higher wage variability.*

(6) What will happen to wages when workers quit? If $\gamma = 0$, wages will necessarily fall when workers quit (since before they quit $\mu > \mu_0$, whereas in the new job $\mu = \mu_0$). If, on the other hand, γ is non-infinitesimal, workers will experience a wage gain when they change jobs, since in this case $\mu^* < \mu_0$ because they are staying in their current job until this job is sufficiently unlikely to be a good match. This last prediction is also consistent with the data, where on average workers who change jobs experience in increase in wages. [But is this a reasonable explanation for wage increases when workers change jobs?].

What is missing from this model is differential learning opportunities in different jobs. If we assume that output and underlying job quality are normally distributed, we already obtain some of amount of differential learning, since the value of learning is higher in new jobs because the precision of the posterior is smaller. Another possibility will be to have heterogeneous jobs, where some jobs have greater returns to match quality, or perhaps some jobs enable faster learning (e.g., more informative signals). Even more interesting would be to allow some amount of learning about general skills. For example, an academic will not be learning and revealing only about his match-specific quality but also about his industry-specific quality (e.g., his research potential). When this is the case, some jobs may play the role of "stepping stones" because they reveal information about the skills and productivity of the worker in a range on other jobs.

Finally, if instead of the reduced-form wage equation, we incorporate competition among firms into this model, some of the predictions change again. For example, we can consider a world in which a finite number of firms with access to the same technology compete a la Bertrand for the worker. Clearly the worker will start working for the firm where the prior of a good match is greatest. Bertrand competition implies that this firm will pay the worker his value at the next best job. Once the worker receives bad news and decides to quit, then he will switch to the job that was previously his next best option. But this implies that his wage, which will now be determined by the third best option (which may in fact be his initial employer) is necessarily smaller, thus job changes will always be associated with wage *declines*. This discussion shows that wage determination assumptions in these models are not innocuous, and more realistic wage determination schemes may lead to results that are not entirely consistent with the data (wage declines rather than wage increases upon job changes). This once again highlights the need for introducing some amount of general skills and job heterogeneity, so that workers quit not only because they have received bad news in their current job but also because they have learned about their ability and can therefore go and work for "higher-quality" jobs.

Part 4

Search and Unemployment

Let us start with the classical McCall model of search. This model is not only elegant, but has also become a workhorse for many questions in macro, labor and industrial organization. An important feature of the model is that it is much more tractable than the original Stigler formulation of search, as one of sampling multiple offers, but we will return to this theme below.

CHAPTER 10

The Partial Equilibrium Model

1. Basic Model

Imagine a partial equilibrium setup with a risk neutral individual in discrete time. At time $t = 0$, this individual has preferences given by

$$\sum_{t=0}^{\infty} \beta^t c_t$$

where c_t is his consumption. He starts life as unemployed. When unemployed, he has access to consumption equal to b (from home production, value of leisure or unemployment benefit). At each time period, he samples a job. All jobs are identical except for their wages, and wages are given by an exogenous stationary distribution of $F(w)$ with finite (bounded) support $\mathbb{W}$, i.e., F is defined only for $w \in \mathbb{W}$. Without loss of any generality, we can take the lower support of $\mathbb{W}$ to be 0, since negative wages can be ruled out. In other words, at every date, the individual samples a wage $w_t \in W$, and has to decide whether to take this or continue searching. Draws from $\mathbb{W}$ over time are independent and identically distributed.

This type of sequential search model can also be referred to as a model of *undirected search*, in the sense that the individual has no ability to seek or direct his search towards different parts of the wage distribution (or towards different types of jobs). This will contrast with models of *directed search* which we will see later.

Let us assume for now that there is no recall, so that the only thing the individual can do is to take the job offered within that date (with recall, the individual would be able to accumulate offers, so at time t, he can choose any of the offers he has received up at that point). If he accepts a job, he will be employed at that job forever, so the net present value of accepting a job of wage w_t is

$$\frac{w_t}{1-\beta}.$$

This is a simple decision problem. Let us specify the class of decision rules of the agent. In particular, let

$$a_t : \mathbb{W} \rightarrow [0,1]$$

denote the action of the agent at time t, which specifies his acceptance probability for each wage in $\mathbb{W}$ at time t. Let $a_t' \in \{0,1\}$ be the realization of the action by the individual (thus allowing for mixed strategies). Let also A_t denote the set of realized actions by the individual, and define $A^t = \prod_{s=0}^{t} A_s$. Then a strategy for the individual in this game is

$$p_t : A^{t-1} \times \mathbb{W} \rightarrow [0,1]$$

Let $\mathcal{P}$ be the set of such functions (with the property that $p_t(\cdot)$ is defined only if $p_s(\cdot) = 0$ for all $s \leq t$) and $\mathcal{P}^\infty$ the set of infinite sequences of such functions. The most general way of expressing the problem of the individual would be as follows. Let $\mathbb{E}$ be the expectations operator. Then the individual's problem is

$$\max_{\{p_t\}_{t=0}^{\infty} \in \mathcal{P}^\infty} \mathbb{E} \sum_{t=0}^{\infty} \beta^t c_t$$

subject to $c_t = b$ if $t < s$ and $c_t = w_s$ if $t \geq s$ where $s = \inf\{n \in \mathbb{N} : a_n' = 1\}$. Naturally, written in this way, the problem looks complicated. Nevertheless, the dynamic programming formulation of this problem will be quite tractable.

To develop this approach, let us analyze this problem by writing it recursively using dynamic programming techniques. First, let us define the value of the agent when he has sampled a job of $w \in \mathbb{W}$. This is clearly given by

$$v(w) = \max\left\{\frac{w}{1-\beta}, \beta v + b\right\}, \tag{10.1}$$

where

$$v = \int_{\mathbb{W}} v(\omega)\, dF(\omega) \tag{10.2}$$

is the continuation value of not accepting a job. Here we have made no assumptions about the structure of the set $\mathbb{W}$, which could be an interval, or might have a mass point, and the density of the distribution F may not exist. Therefore, the integral in (10.2) should be interpreted as a Lebesgue integral.

Equation (10.1) follows from the observation that the individual will either accept the job, receiving a constant consumption stream of w (valued at $w/(1-\beta)$) or will turn down this job, in which case he will enjoy the consumption level b, and receive the continuation value v. Maximization implies that the individual takes whichever of these two options gives higher net present value.

Equation (10.2), on the other hand, follows from the fact that from tomorrow on, the individual faces the same distribution of job offers, so v is simply the expected value of $v(w)$ over the stationary distribution of wages.

We are interested in finding both the value function $v(w)$ and the optimal policy of the individual.

Combining these two equations, we can write

$$v(w) = \max\left\{\frac{w}{1-\beta}, b + \beta \int_{\mathbb{W}} v(\omega)\, dF(\omega)\right\}. \tag{10.3}$$

We can now deduce the existence of optimal policies using standard theorems from dynamic programming. But in fact, (10.3) is simple enough that, one can derive these results without appealing to these theorems. In particular, this equation makes it clear that $v(w)$ must be piecewise linear with first a flat portion and then an increasing portion.

The next task is to determine the optimal policy. But the fact that $v(w)$ is non-decreasing and is piecewise linear with first a flat portion, immediately tells us that the optimal policy will take a *reservation wage* form, which is a key result of the sequential search model. More explicitly, there will exist some reservation wage R such that all wages above R will be accepted and those $w < R$ will be turned down. Moreover, this reservation wage has to be such that

$$\frac{R}{1-\beta} = b + \beta \int_{\mathbb{W}} v(\omega)\, dF(\omega), \tag{10.4}$$

so that the individual is just indifferent between taking $w = R$ and waiting for one more period. Next we also have that since $w < R$ are turned down, for all $w < R$

$$\begin{aligned} v(w) &= b + \beta \int_{\mathbb{W}} v(\omega)\, dF(\omega) \\ &= \frac{R}{1-\beta}, \end{aligned}$$

and for all $w \geq R$,

$$v(w) = \frac{w}{1-\beta}$$

Therefore,

$$\int_{\mathbb{W}} v(\omega)\, dF(\omega) = \frac{RF(R)}{1-\beta} + \int_{w \geq R} \frac{w}{1-\beta} dF(w).$$

Combining this with (10.4), we have

$$\frac{R}{1-\beta} = b + \beta \left[\frac{RF(R)}{1-\beta} + \int_{w \geq R} \frac{w}{1-\beta} dF(w) \right]$$

Manipulating this equation, we can write

$$R = \frac{1}{1-\beta F(R)} \left[b(1-\beta) + \beta \int_R^{+\infty} w dF(w) \right],$$

which is one way of expressing the reservation wage. More useful is to rewrite this equation as

$$\int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{R}{1-\beta} dF(w) = b + \beta \left[\int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{w}{1-\beta} dF(w) \right]$$

Now subtracting $\beta R \int_{w \geq R} dF(w) / (1-\beta) + \beta R \int_{w<R} dF(w) / (1-\beta)$ from both sides, we obtain

$$\begin{aligned} & \int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{R}{1-\beta} dF(w) \\ & -\beta \int_{w \geq R} \frac{R}{1-\beta} dF(w) - \beta \int_{w<R} \frac{R}{1-\beta} dF(w) \\ = \; & b + \beta \left[\int_{w \geq R} \frac{w-R}{1-\beta} dF(w) \right] \end{aligned}$$

Collecting terms, we obtain

$$R - b = \frac{\beta}{1-\beta} \left[\int_{w \geq R} (w-R)\, dF(w) \right], \tag{10.5}$$

which is a particularly useful and economically intuitive way of characterizing the reservation wage. The left-hand side is best understood as the cost of foregoing the wage of R, while the right hand side is the expected benefit of one more search. Clearly, at the reservation wage, these two are equal.

One implication of the reservation wage policy is that the assumption of no recall, made above, was of no consequence. In a stationary environment, the worker will have a constant reservation wage, and therefore has no desire to go back and take a job that he had previously rejected.

Let us define the right hand side of equation (10.5) as

$$g(R) \equiv \frac{\beta}{1-\beta}\left[\int_{w \geq R}(w-R)\,dF(w)\right],$$

which represents the expected benefit of one more search as a function of the reservation wage. Clearly,

$$\begin{aligned} g'(R) &= -\frac{\beta}{1-\beta}(R-R)\,f(R) - \frac{\beta}{1-\beta}\left[\int_{w \geq R} dF(w)\right] \\ &= -\frac{\beta}{1-\beta}\left[1-F(R)\right] < 0 \end{aligned}$$

This implies that equation (10.5) has a unique solution. Moreover, by the implicit function theorem,

$$\frac{dR}{db} = \frac{1}{1-g'(R)} > 0,$$

so that as expected, higher benefits when unemployed increase the reservation wage, making workers more picky.

Moreover, for future reference, also note that when the density of $F(R)$, denoted by $f(R)$, exists, the second derivative of g also exists and is

$$g''(R) = \frac{\beta}{1-\beta}f(R) \geq 0,$$

so that the right hand side of equation (10.5) is also convex.

The next question is to investigate how changes in the distribution of wages F affect the reservation wage. Before doing this, however, we will use this partial equilibrium McCall model to derive a very simple theory of unemployment.

2. Unemployment with Sequential Search

Let us now use the McCall model to construct a simple model of unemployment. In particular, let us suppose that there is now a continuum 1 of identical individuals sampling jobs from the same stationary distribution F. Moreover, once a job is created, it lasts until the worker dies, which happens with probability s. There is a mass of s workers born every period, so that population is constant, and these workers start out as unemployed. The death probability means that the effective discount factor of workers is equal to $\beta(1-s)$. Consequently, the value of having accepted a wage of w is:

$$v^a(w) = \frac{w}{1-\beta(1-s)}.$$

Moreover, with the same reasoning as before, the value of having a job offer at wage w at hand is

$$v(w) = \max\{v^a(w), b + \beta(1-s)v\}$$

with

$$v = \int_{\mathbb{W}} v(w)\, dF.$$

Therefore, the same steps lead to the reservation wage equation:

$$R - b = \frac{\beta(1-s)}{1-\beta(1-s)}\left[\int_{w \geq R} (w-R)\, dF(w)\right].$$

Now what is interesting is to look at the law of motion of unemployment. Let us start time t with U_t unemployed workers. There will be s new workers born into the unemployment pool. Out of the U_t unemployed workers, those who survive and do not find a job will remain unemployed. Therefore

$$U_{t+1} = s + (1-s)F(R)U_t,$$

where $F(R)$ is the probability of not finding a job (i.e., a wage offer below the reservation wage), so $(1-s)F(R)$ is the joint probability of not finding a job and surviving, i.e., of remaining unemployed. This is a simple first-order linear difference equation (only depending on the reservation wage R, which is itself independent of the level of unemployment, U_t) and determines the law of motion of unemployment. Moreover, since $(1-s)\,F(R)<1$, it is asymptotically stable, and will converge to a unique steady-state level of unemployment.

To get more insight, subtract U_t from both sides, and rearrange to obtain

$$U_{t+1}-U_t=s\,(1-U_t)-(1-s)\,(1-F(R))\,U_t.$$

This is the simplest example of the *flow approach* to the labor market, where unemployment dynamics are determined by flows in and out of unemployment. In fact is equation has the canonical form for change in unemployment in the flow approach. The left hand-side is the change unemployment (which can be either indiscreet or continuous time), while the right hand-side consists of the job destruction rate (in this case s) multiplied by $(1-U_t)$ minus the rate at which workers leave unemployment (in this case $(1-s)\,(1-F(R))$) multiplied with U_t.

The unique steady-state unemployment rate where $U_{t+1}=U_t$ is given by

$$U=\frac{s}{s+(1-s)\,(1-F(R))}.$$

This is again the canonical formula of the flow approach. The steady-state unemployment rate is equal to the job destruction rate (here the rate at which workers die, s) divided by the job destruction rate plus the job creation rate (here in fact the rate at which workers leave unemployment, which is different from the job creation rate). Clearly, an increases in s will raise steady-state unemployment. Moreover, an increase in R, that is, a higher reservation wage, will also depress job creation and increase unemployment.

3. Aside on Riskiness and Mean Preserving Spreads

To investigate the effect of changes in the distribution of wages on the reservation wage, let us introduce the concept of *mean preserving spreads*. Loosely speaking, a mean preserving spread is a change in distribution that increases risk. Let a family of distributions over some set $X\subset\mathbb{R}$ with generic element x be denoted by $F\,(x,r)$, where r is a shift variable, which changes the distribution function. An example will be $F(x, r)$ to stand for mean zero normal variables, with r parameterizing the variance of the distribution. In fact, the normal distribution is special in the sense that, the mean and the variance completely describe the distribution, so the notion of risk can be captured by the variance. This is generally not true. The notion of "riskier" is a more stringent notion than having a greater variance. In fact, we will see that "riskier than" is a partial order (while, clearly, comparing variances is a complete order).

Here is a natural definition of one distribution being riskier than another, first introduced by Blackwell, and then by Rothschild and Stiglitz.

DEFINITION 10.1. *$F(x, r)$ is less risky than $F(x, r')$, written as $F(x, r) \succeq_R F(x, r')$, if for all concave and increasing $u : \mathbb{R} \to \mathbb{R}$, we have*

$$\int_X u(x)\, dF(x, r) \geq \int_X u(x)\, dF(x, r').$$

At some level, it may be a more intuitive definition of "riskiness" to require that $F(x, r)$ and $F(x, r')$ to have the same mean, i.e., $\int_X x dF(x, r) = \int_X x dF(x, r')$, while still $F(x, r) \succeq_R F(x, r')$. However, whether we do this or not is not important for our focus.

A related definition is that of second-order stochastic dominance.

DEFINITION 10.2. *$F(x, r)$ second order stochastically dominates $F(x, r')$, written as $F(x, r) \succeq_{SD} F(x, r')$, if*

$$\int_{-\infty}^{c} F(x, r)\, dx \leq \int_{-\infty}^{c} F(x, r')\, dx, \text{ for all } c \in X.$$

In other words, this definition requires the distribution function of $F(x, r)$ to start lower and always keep a lower integral than that of $F(x, r')$. One easy case where this will be satisfied is when both distribution functions have the same mean and they intersect only once: "single crossing") with $F(x, r)$ cutting $F(x, r')$ from below.

The definitions above use weak inequalities. Alternatively, they can be strengthened to strict inequalities. In particular, the first definition would require a strict inequality for functions that are strictly concave over some range, while the second definition will require strict inequality for some c.

THEOREM 10.1. ***(Blackwell, Rothschild and Stiglitz)*** *$F(x,r) \succeq_R F(x,r')$ if and only if $F(x,r) \succeq_{SD} F(x,r')$.*

Therefore, there is an intimate link between second-order stochastic dominance and the notion of riskiness. This also shows that variance is not a good measure of riskiness, since second order stochastic dominance is a partial order.

Now **mean preserving spreads** are essentially equivalent to second-order stochastic dominance with the additional restriction that both distributions have the same mean. As the term suggests, a mean preserving spread is equivalent to taking a given distribution and shifting some of the weight from around the mean to the tails. Alternative representations also include one distribution being obtained from the other by adding "white noise" to the other.

Second-order stochastic dominance plays a very important role in the theory of learning, and also more generally in the theory of decision-making under uncertainty. Here it will be useful for comparative statics.

4. Back to the Basic Partial Equilibrium Search Model

Let us return to the McCall search model. To investigate the effect of changes in the riskiness (or dispersion) of the wage distribution on reservation wages, and thus on search and unemployment behavior, let us express the reservation wage somewhat differently. Start with equation (10.5) above, which is reproduced here for convenience,

$$R - b = \frac{\beta}{1-\beta}\left[\int_{w \geq R} (w - R)\, dF(w)\right].$$

Rewrite this as

$$\begin{aligned}
R-b &= \frac{\beta}{1-\beta}\left[\int_{w\geq R}(w-R)\,dF(w)\right]+\frac{\beta}{1-\beta}\left[\int_{w\leq R}(w-R)\,dF(w)\right] \\
&\quad -\frac{\beta}{1-\beta}\left[\int_{w\leq R}(w-R)\,dF(w)\right], \\
&= \frac{\beta}{1-\beta}(Ew-R)-\frac{\beta}{1-\beta}\left[\int_{w\leq R}(w-R)\,dF(w)\right],
\end{aligned}$$

where Ew is the mean of the wage distribution, i.e.,

$$Ew=\int_{\mathbb{W}} wdF(w).$$

Now rearranging this last equation, we have

$$R-b=\beta(Ew-b)-\beta\int_{w\leq R}(w-R)\,dF(w).$$

Applying integration by parts to the integral on the right hand side, in particular, noting that

$$\begin{aligned}
\int_{w\leq R} wdF(w) &= \int_0^R wdF(w) \\
&= wF(w)|_0^R-\int_0^R F(w)\,dw \\
&= RF(R)-\int_0^R F(w)\,dw,
\end{aligned}$$

this equation can be rewritten as

$$R-b=\beta(Ew-b)+\beta\int_0^R F(w)\,dw. \tag{10.6}$$

Now consider a shift from F to $\tilde{F}$ corresponding to a mean preserving spread. This implies that Ew is unchanged, but by definition of a mean preserving spread (second-order stochastic dominance), the last integral increases. Therefore, the mean preserving spread induces a shift in the reservation wage from R to $\tilde{R}>R$. This reflects the greater option value of waiting when faced with a more dispersed wage distribution; lower wages are already turned down, while higher wages are now more likely.

A different way of viewing this result is that the analysis above established that the value function $v(w)$ is convex. While Theorem 10.1 shows that concave utility functions like less risky distributions, convex functions like more risky distributions.

5. Paradoxes of Search

The search framework is attractive especially when we want to think of a world without a Walrasian auctioneer, or alternatively a world with "frictions". How do prices get determined? How do potential buyers and sellers get together? Can we think of Walrasian equilibrium as an approximation to such a world under some conditions?

Search theory holds the promise of potentially answering these questions, and providing us with a framework for analysis.

5.1. The Rothschild Critique. The McCall model is an attractive starting point. It captures the intuition that individuals may be searching for the right types of job (e.g., jobs offering higher wages), trading off the prospects of future benefits (high wages) for the costs of foregoing current wages.

But everything hinges on the distribution of wages, $F(w)$. Where does this come from? Presumably somebody is offering every wage in the support of this distribution.

The basis of the Rothschild critique is that it is difficult to rationalize the distribution function $F(w)$ as resulting from profit-maximizing choices of firms.

Imagine that the economy consists of a mass 1 of identical workers similar to our searching agent. On the other side, there are N firms that can productively employ workers. Imagine that firm j has access to a technology such that it can employ l_j workers to produce

$$y_j = x_j l_j$$

units of output (with its price normalized to one as the numeraire, so that w is the real wage). Suppose that each firm can only attract workers by posting a single vacancy. Moreover, to simplify life, suppose that firms post a vacancy at the beginning of the game at $t = 0$, and then do not change the wage from then on. This will both simplify the strategies, and imply that the wage distribution will be stationary, since all the same wages will remain active throughout time. [Can you see why this simplifies the discussion? Imagine, for contrast, the case in which each firm only hires one worker; then think of the wage distribution at time t, $F_t(w)$, starting with some arbitrary $F_0(w)$. Will it remain constant?]

Suppose that the distribution of x in the population of firms is given by $G(x)$ with support $X \subset \mathbb{R}_+$. Also assume that there is some cost $\gamma > 0$ of posting a vacancy at the beginning, and finally, that $N >> 1$ (i.e., $N = \int_{-\infty}^{\infty} dG(x) >> 1$) and each worker samples one firm from the distribution of posting firms.

As before, we will assume that once a worker accepts a job, this is permanent, and he will be employed at this job forever. Moreover let us set $b = 0$, so that there is no unemployment benefits. Finally, to keep the environment entirely stationary, assume that once a worker accepts a job, a new worker is born, and starts search.

Will these firms offer a non-degenerate wage distribution $F(w)$?

The answer is no.

First, note that an endogenous wage distribution equilibrium would correspond to a function

$$p : X \to \{0, 1\},$$

denoting whether the firm is posting a vacancy or not, and if it is, i.e., $p = 1$,

$$h : X \to \mathbb{R}_+,$$

specifying the wage it is offering.

It is intuitive that $h(x)$ should be non-decreasing (higher wages are more attractive to high productivity firms). Let us suppose that this is so, and denote its set-valued inverse mapping by h^{-1}. Then, the along-the-equilibrium path wage distribution is

$$F(w) = \frac{\int_{-\infty}^{h^{-1}(w)} p(x)\, dG(x)}{\int_{-\infty}^{\infty} p(x)\, dG(x)}.$$

Why?

In addition, the strategies of workers can be represented by a function

$$a : \mathbb{R}_+ \to [0,1]$$

denoting the probability that the worker will accept any wage in the "potential support" of the wage distribution, with 1 standing for acceptance. This is general enough to nest non-symmetric or mixed strategies.

The natural equilibrium concept is subgame perfect Nash equilibrium, whereby the strategies of firms (p, h) and those of workers, a, are best responses to each other in all subgames.

The same arguments as above imply that all workers will use a reservation wage, so

$$\begin{aligned} a(w) &= 1 \text{ if } w \geq R \\ &= 0 \text{ otherwise} \end{aligned}$$

Since all workers are identical and the equation above determining the reservation wage, (10.5), has a unique solution, all workers will all be using the same reservation rule, accepting all wages $w \geq R$ and turning down those $w < R$. Workers' strategies are therefore again characterized by a reservation wage R.

Now take a firm with productivity x offering a wage $w' > R$. Its net present value of profits from this period's matches is

$$\pi(p = 1, w' > R, x) = -\gamma + \frac{1}{n}\frac{(x - w')}{1 - \beta}$$

where

$$n = \int_{-\infty}^{\infty} p(x)\, dG(x)$$

is the measure of active firms, $1/n$ is the probability of a match within each period (since the population of active firms and searching workers are constant), and $x - w'$ is the profit from the worker discounted at the discount factor β.

Notice two (implicit) assumptions here: (1) wage posting: each job comes with a commitment to a certain wage; (2) undirected search: the worker makes a random draw from the distribution F, and the only way he can seek higher wages is by turning down lower wages that he samples.

This firm can deviate and cut its wage to some value in the interval $[R, w')$. All workers will still accept this job since its wage is above the reservation wage, and the firm will increase its profits to

$$\pi\left(p=1, w \in [R, w'), x\right) = -\gamma + \frac{1}{n}\frac{x-w}{1-\beta} > \pi\left(p=1, w', x\right)$$

So there should not be any wages strictly above R.

Next consider a firm offering a wage $\tilde{w} < R$. This wage will be rejected by all workers, and the firm would lose the cost of posting a vacancy, i.e.,

$$\pi\left(p=1, w<R, x\right) = -\gamma,$$

and this firm can deviate to $p=0$ and make zero profits. Therefore, in equilibrium when workers use the reservation wage rule of accepting only wages greater than R, all firms will offer the same wage R, and there is no distribution and no search.

This establishes

THEOREM 10.2. *When all workers are homogeneous and engage in undirected search, all equilibrium distributions will have a mass point at their reservation wage R.*

In fact, the paradox is even deeper.

5.2. The Diamond Paradox. The following result is one form of the Diamond paradox:

THEOREM 10.3. ***(Diamond Paradox)*** *For all $\beta < 1$, the unique equilibrium in the above economy is $R = 0$.*

Given the Theorem 10.2, this result is easy to understand. Theorem 10.2 implies that all firms will offer the same wage, R.

Suppose $R > 0$, and $\beta < 1$. What is the optimal acceptance function, a, for a worker?

If the answer is

$$\begin{aligned} a(w) &= 1 \text{ if } w \geq R \\ &= 0 \text{ otherwise} \end{aligned}$$

then we can support all firms offering $w = R$ as an equilibrium (notice that the acceptance function needs to be defined for wages "off-the-equilibrium path"). Why is this important?

However, we can prove:

LEMMA 10.1. *There exists $\varepsilon > 0$ such that when "almost all" firms are offering $w = R$, it is optimal for each worker to use the following acceptance strategy:*

$$\begin{aligned} a(w) &= 1 \textit{ if } w \geq R - \varepsilon \\ &= 0 \textit{ otherwise} \end{aligned}$$

Note: think about what "almost all" means here and why it is necessary.

PROOF. If the worker accepts the wage of $R - \varepsilon$ today his payoff is

$$u^{accept} = \frac{R - \varepsilon}{1 - \beta}$$

If he rejects and waits until next period, then since "almost all" firms are offering R, he will receive the wage of R, so

$$u^{reject} = \frac{\beta R}{1 - \beta}$$

where the additional β comes in because of the waiting period. For all $\beta < 1$, there exists $\varepsilon > 0$ such that

$$u^{accept} > u^{reject},$$

proving the claim. $\square$

What is the intuition for this lemma?

But this implies that, starting from an allocation where all firms offer R, any firm can deviate and offer a wage of $R - \varepsilon$ and increase its profits. This proves that no wage $R > 0$ can be the equilibrium, proving the proposition.

Notice that subgame perfection is important here. We know that these are non-subgame perfect Nash equilibria, and this highlights the importance of using the right equilibrium concept in the context of dynamic economies.

So now we are in a conundrum. Not only does there fail to be a wage distribution, but irrespective of the distribution of productivities or the degree of discounting, all firms offer the lowest possible wage, i.e., they are full monopsonists.

How do we resolve this paradox?

(1) By assumption: assume that $F(w)$ is not the distribution of wages, but the distribution of "fruits" exogenously offered by "trees". This is clearly unsatisfactory, both from the modeling point of view, and from the point of view of asking policy questions from the model (e.g., how does unemployment insurance affect the equilibrium? The answer will depend also on how the equilibrium wage distribution changes).

(2) Introduce other dimensions of heterogeneity: to be done later.

(3) Modify the wage determination assumptions: to be done in a little bit.

CHAPTER 11

Basic Equilibrium Search Framework

1. Motivation

Importance of labor market flows, job creation, job destruction.

Need for a framework that can be used for equilibrium analysis, but allows for unemployment→Equilibrium search models.

More reduced form than a partial equilibrium model in order to avoid the "paradoxes" mentioned above.

2. The Basic Search Model

Now we discuss the basic search-matching model, or sometimes called the flow approach to the labor market.

Here the basic idea is that that are frictions in the labor market, making it costly (time-consuming) for workers to find firms and vice versa. This will lead to what is commonly referred to as "frictional unemployment". However, as soon as there are these types of frictions, there are also quasi-rents in the relationship between firms and workers, and there will be room for rent-sharing. In the basic search model, the main reason for high unemployment may not be the time costs of finding partners, but bargaining between firms and workers which leads to non-market-clearing equilibrium prices.

Here is a simple version of the basic search model.

The first important object is the matching function, which gives the number of matches between firms and workers as a function of the number of unemployed workers and number of vacancies.

Matching Function: Matches $= x(U, V)$

This function captures the frictions inherent in the process of assigning workers to jobs in a very *reduced form* way. This reduced-form structure is its advantage and disadvantage. It is difficult to have microfoundations for this function, but it is very tractable, fairly easy to map to data (at least to data on job flows and worker flows), and captures the intuitive notion that job finding rates for workers should depend on how many unemployed workers are chasing how many vacancies.

Of course the form of the matching function will also depend on what the time horizon is.

Following our treatment of the Shapiro-Stiglitz model, we will work with continuous time, so we should think of $x(U, V)$ as the flow rate of matches.

We typically assume that this matching function exhibits constant returns to scale (CRS), that is,

$$\begin{aligned} \text{Matches} &= xL = x(uL, vL) \\ &\implies x = x(u, v) \end{aligned}$$

Here we have adopted the usual notation:

U =unemployment;

u =unemployment rate

V =vacancies;

v = vacancy rate (per worker in labor force)

L = labor force

Existing aggregate evidence suggests that the assumption of x exhibiting CRS is reasonable (Blanchard and Diamond, 1989)

Using the constant returns assumption, we can express everything as a function of the tightness of the labor market.

Therefore;

$$q(\theta) \equiv \frac{x}{v} = x\left(\frac{u}{v}, 1\right),$$

where $\theta \equiv v/u$ is the tightness of the labor market

Since we are in continuous time, these things immediately map to flow rates. Namely

$$q(\theta) : \text{Poisson arrival rate of match for a vacancy}$$
$$q(\theta)\theta : \text{Poisson arrival rate of match for an unemployed worker}$$

What does Poisson mean?

Take a short period of time Δt, then the Poisson process is defined such that during this time interval, the probability that there will be one arrival, for example one arrival of a job for a worker, is

$$\Delta t q(\theta)\theta$$

The probability that there will be more than one arrivals is vanishingly small (formally, of order $o(\Delta t)$).

Therefore,

$1 - \Delta t\theta q(\theta)$: probability that a worker looking for a job will not find one during Δt

This probability depends on θ, thus leading to a potential externality—the search behavior of others affects my own job finding rate.

The search model is also sometimes called the flow approach to unemployment because it's all about job flows. That is about job creation and job destruction.

This is another dividing line between labor and macro. Many macroeconomists look at data on job creation and job destruction following Davis and Haltiwanger. Most labor economists do not look at these data. Presumably there is some information in them.

Job creation is equal to

$$\text{Job creation} = u\theta q(\theta)L$$

What about job destruction?

Let us start with the simplest model of job destruction, which is basically to treat it as "exogenous".

Think of it as follows, firms are hit by adverse shocks, and then they decide whether to destroy or to continue.

$\longrightarrow$ Adverse Shock$\longrightarrow$destroy

$\longrightarrow$ continue

Exogenous job destruction: Adverse shock $= -\infty$ with "probability " s

As in the Shapiro-Stiglitz model, we will focus on steady states.

Steady State:

flow into unemployment = flow out of unemployment

Therefore, with exogenous job destruction:

$$s(1-u) = \theta q(\theta)u$$

This gives the steady-state unemployment rate as

$$u = \frac{s}{s + \theta q(\theta)}$$

This relationship is sometimes referred to as the Beveridge Curve, or the U-V curve. It draws a downward sloping locus of unemployment-vacancy combinations in the U-V space that are consistent with flow into unemployment being equal with flow out of unemployment. Some authors interpret shifts of this relationship is reflecting structural changes in the labor market, but we will see that there are many factors that might actually shift at a generalized version of such relationship.

It is a crucial equation even if you don't like the search model. It relates the unemployment rate to the rate at which people leave their jobs and and unemployment and the rate at which people leave the unemployment pool.

In a more realistic model, of course, we have to take into account the rate at which people go and come back from out-of-labor force status.

Let's next turn to the production side.

Let the output of each firm be given by neoclassical production function combining labor and capital:

$$Y = AF(K, N)$$

where the production function F is assumed to exhibit constant returns, K is the capital stock of the economy, and N is employment (different from labor force because of unemployment).

Defining $k \equiv K/N$ as the capital labor ratio, we have that output per worker is:

$$\frac{Y}{N} = Af(k) \equiv AF(\frac{K}{N}, 1)$$

because of constant returns.

Two interpretations $\longrightarrow$ each firm is a "job" hires one worker
each firm can hire as many worker as it likes

For our purposes either interpretation is fine

Hiring: Vacancy costs γ_0: fixed cost of hiring

r: cost of capital

δ: depreciation

The key assumption here is that capital is perfectly reversible.

As in the Shapiro Stiglitz model, we will solve everything by using dynamic programming, or in other words by writing the asset value equations. As in there, let us define those in terms of the present discounted values.

Namely, let

J^V : PDV of a vacancy

J^F :PDV of a "job"

J^U :PDV of a searching worker

J^E :PDV of an employed worker

More generally, we have that worker utility is: $EU_0 = \int_0^\infty e^{-rt} U(c_t)$, but for what we care here, risk-neutrality is sufficient.

Utility $U(c) = c$, in other words, linear utility, so agents are risk-neutral.

Perfect capital market gives the asset value for a vacancy (in steady state) as

$$rJ^V = -\gamma_0 + q(\theta)(J^F - J^V)$$

Intuitively, there is a cost of vacancy equal to γ_0 at every instant, and the vacancy turns into a filled job at the flow rate $q(\theta)$.

Notice that in writing this expression, we have assumed that firms are risk neutral. Why is this important?
$\longrightarrow$ workers risk neutral, or
$\longrightarrow$ complete markets

The question is how to model job creation (which is the equivalent of how to model labor demand in a competitive labor market).

Presumably, firms decide to create jobs when there are profit opportunities.

The simplest and perhaps the most extreme form of endogenous job creation is to assume that there will be a firm that creates a vacancy as soon as the value of a vacancy is positive (after all, unless there are scarce factors necessary for creating vacancies anybody should be able to create one).

This is sometimes referred to as the free-entry assumption, because it amounts to imposing that whenever there are potential profits they will be eroded by entry.

Free Entry $\Longrightarrow$

$$J^V \equiv 0$$

The most important implication of this assumption is that job creation can happen really "fast", except because of the frictions created by matching searching workers to searching vacancies.

Alternative would be: $\gamma_0 = \Gamma_0(V)$ or $\Gamma_1(\theta)$, so as there are more and more jobs created, the cost of opening an additional job increases.

Free entry implies that

$$J^F = \frac{\gamma_0}{q(\theta)}$$

Next, we can write another asset value equation for the value of a field job:

$$r(J^F + k) = Af(k) - \delta k - w - s(J^F - J^V)$$

Intuitively, the firm has two assets: the fact that it is matched with a worker, and its capital, k. So its asset value is $J^F + k$ (more generally, without the perfect reversability, we would have the more general $J^F(k)$). Its return is equal to production, $Af(k)$, and its costs are depreciation of capital and wages, δk and w. Finally, at the rate s, the relationship comes to an end and the firm loses J^F.

Perfect Reversability implies that w does not depend on the firm's choice of capital

$\Longrightarrow$ equilibrium capital utilization $f'(k) = r + \delta$— Modified Golden Rule

[...Digression: Suppose k is not perfectly reversible then suppose that the worker captures a fraction β all the output in bargaining. Then the wage depends on the capital stock of the firm, as in the holdup models discussed before.

$$\begin{aligned} w(k) &= \beta Af(k) \\ Af'(k) &= \frac{r+\delta}{1-\beta} \text{ ; capital accumulation is distorted} \end{aligned}$$

...]

Now, ignoring this digression

$$Af(k) - (r-\delta)k - w - \frac{(r+s)}{q(\theta)}\gamma_0 = 0$$

Now returning to the worker side, the risk neutrality of workers gives

$$rJ^U = z + \theta q(\theta)(J^E - J^U)$$

where z is unemployment benefits. The intuition for this equation is similar. We also have

$$rJ^E = w + s(J^U - J^E)$$

Solving these equations we obtain

$$\begin{aligned} rJ^U &= \frac{(r+s)z + \theta q(\theta) w}{r+s+\theta q(\theta)} \\ rJ^E &= \frac{sz + [r + \theta q(\theta)]\, w}{r+s+\theta q(\theta)} \end{aligned}$$

How are wages determined? Nash Bargaining.

Why do we need bargaining? Answer: bilateral monopoly or much more specifically: match specific surplus.

Think of a competitive labor market, at the margin the firm is indifferent between employing the marginal worker or not, and the worker is indifferent between supplying the marginal hour or not (or working for this firm or another firm). We can make both parties in different at the same time—no match-specific surplus.

In a frictional labor market, if we choose the wage such that $J^E = 0$, we will typically have $J^F > 0$ and vice versa. There is some surplus to be shared.

Nash solution to bargaining is again the natural benchmark. Let us assume that the worker has bargaining power β.

Applying this formula, for pair i, we have

$$\begin{aligned} rJ_i^F &= Af(k) - (r+\delta)k - w_i - sJ_i^F \\ rJ_i^E &= w_i - s(J_i^E - J_0^U). \end{aligned}$$

The Nash solution will solve

$$\max (J_i^E - J^U)^\beta (J_i^F - J^V)^{1-\beta}$$

$$\beta \quad = \quad \text{bargaining power of the worker}$$

Since we have linear utility, thus "transferable utility", this implies

$$\Longrightarrow \quad J_i^E - J^U = \beta(J_i^F + J_i^E - J^V - J^U)$$

$$\Longrightarrow \quad w = (1-\beta)z + \beta\left[Af(k) - (r+\delta)k + \theta\gamma_0\right]$$

Here $[Af(k) - (r+\delta)k + \theta\gamma_0]$ is the quasi-rent created by a match that the firm and workers share. Why is the term $\theta\gamma_0$ there?

Now we are in this position to characterize the steady-state equilibrium.

Steady State Equilibrium is given by four equations

(1) The Beveridge curve:

$$u = \frac{s}{s + \theta q(\theta)}$$

(2) Job creation leads zero profits:

$$Af(k) - (r+\delta)k - w - \frac{(r+s)}{q(\theta)}\gamma_0 = 0$$

(3) Wage determination:

$$w = (1-\beta)z + \beta\left[Af(k) - (r+\delta)k + \theta\gamma_0\right]$$

(4) Modified golden rule:

$$Af'(k) = r + \delta$$

These four equations define a block recursive system

$(4) + r \longrightarrow k$

$k + r + (2) + (3) \longrightarrow \theta, w$

$\theta + (1) \longrightarrow u$

Alternatively, combining three of these equations we obtain the zero-profit locus, the VS curve, and combine it with the Beveridge curve. More specifically,

$(2), (3), (4) \Longrightarrow$ the VS curve

$$(1-\beta)\left[Af(k)-(r+\delta)k-z\right]-\frac{r+\delta+\beta\theta q(\theta)}{q(\theta)}\gamma_0=0$$

Therefore, the equilibrium looks very similar to the intersection of "quasi-labor demand" and "quasi-labor supply".

Quasi-labor supply is given by the Beveridge curve, while labor demand is given by the zero profit conditions.

Given this equilibrium, comparative statics (for steady states) are straightforward.

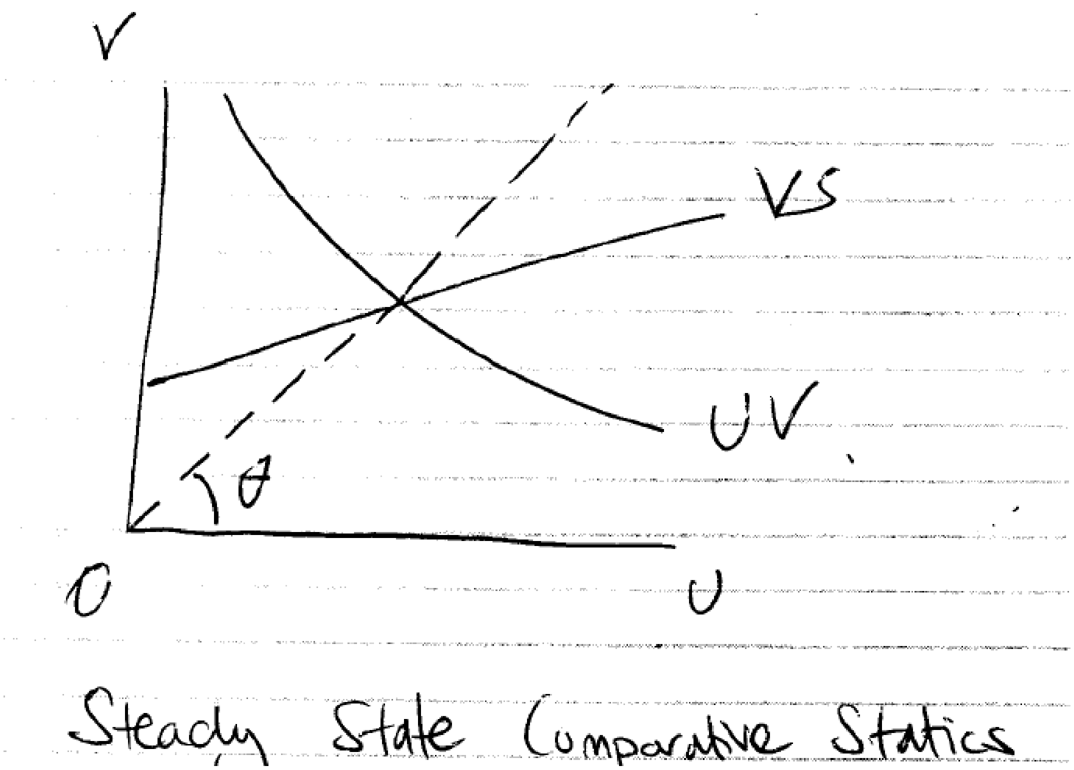


FIGURE 11.1

For example:

$s\uparrow \quad U\uparrow \quad V\uparrow \quad \theta\downarrow \quad w\downarrow$

$r\uparrow \quad U\uparrow \quad V\downarrow \quad \theta\downarrow \quad w\downarrow$

$\gamma_0\uparrow \quad U\uparrow \quad V\downarrow \quad \theta\downarrow \quad w\downarrow$

$$\begin{array}{lllll} \beta\uparrow & U\uparrow & V\downarrow & \theta\downarrow & w\uparrow \\ z\uparrow & U\uparrow & V\downarrow & \theta\downarrow & w\uparrow \\ A\uparrow & U\downarrow & V\uparrow & \theta\uparrow & w\uparrow \end{array}$$

Thus, a greater exogenous separation rate, higher discount rates, higher costs of creating vacancies, higher bargaining power of workers, higher unemployment benefits lead to higher unemployment. Greater productivity of jobs, leads to lower unemployment.

Interestingly, some of those, notably the greater separation rate also increases the number of vacancies.

Can we think of any of these factors is explaining the rise in unemployment in Europe during the 1980s, or the lesser rise in unemployment in 1980s in in the United States?

3. Efficiency of Search Equilibrium

Is the search equilibrium efficient? Clearly, it is inefficient relative to a first-best alternative, e.g., a social planner that can avoid the matching frictions.

However, this is not an interesting benchmark. Much more interesting is whether a social planner affected by exactly the same externalities as the market economy can do better than the decentralized equilibrium.

An alternative way of asking this question is to think about externalities. In this economy there are two externalities

$$\begin{array}{lll} \theta\uparrow & \Longrightarrow & \text{workers find jobs more easily} \\ & \hookrightarrow & \text{thick-market externality} \\ & \Longrightarrow & \text{firms find workers more slowly} \\ & \hookrightarrow & \text{congestion externality} \end{array}$$

Therefore, the question of efficiency boils down to whether these two externalities cancel each other or whether one of them dominates.

To analyze this question more systematically, consider a social planner subject to the same constraints, intending to maximize “total surplus”, in other words, pursuing a utilitarian objective.

First ignore discounting, i.e., $r \to 0$, then the planner's problem can be written as

$$\begin{aligned}\max_{u,\theta} SS &= (1-u)y + uz - u\theta\gamma_0. \\ &\text{s.t.} \\ u &= \frac{s}{s+\theta q(\theta)}.\end{aligned}$$

where we assumed that z corresponds to the utility of leisure rather than unemployment benefits (how would this be different if z were unemployment benefits?)

The form of the objective function is intuitive. For every employed worker, a fraction $1-u$ of the workers, the society receives an output of y; for every unemployed worker, a fraction u of the population, it receives z, and in addition for every vacancy it pays the cost of γ_0 (and there are $u\theta$ vacancies).

The constraint on this problem is that imposed by the matching frictions, i.e. the Beveridge curve, capturing the fact that lower unemployment can only be achieved by creating more vacancies, i.e., higher θ.

Holding $r = 0$, turns this from a dynamic into a static optimization problem, and it can be analyzed by forming the Lagrangian, which is

$$\mathcal{L} = (1-u)y + uz - u\theta\gamma_0 + \lambda\left[u - \frac{s}{s+\theta q(\theta)}\right]$$

The first-order conditions with respect to u and θ are straightforward:

$$(y-z) + \theta\gamma_0 = \lambda$$

$$u\gamma_0 = \lambda s \frac{\theta q'(\theta) + q(\theta)}{(s+\theta q(\theta))^2}$$

Since the constraint will clearly binding (why is this? Otherwise reduce θ, and social surplus increases), we can substitute for u from the Beveridge curve, and obtain:

$$\lambda = \frac{\gamma_0 (s + \theta q(\theta))}{\theta q'(\theta) + q(\theta)}$$

Now substitute this into the first condition to obtain

$$[\theta q'(\theta) + q(\theta)](y-z) + [\theta q'(\theta) + q(\theta)]\theta\gamma_0 - \gamma_0(s + \theta q(\theta)) = 0$$

Now simplifying and dividing through by $q\left(\theta\right)$, we obtain

$$[1-\eta(\theta)]\left[y-z\right]-\frac{s+\eta(\theta)\theta q(\theta)}{q(\theta)}\gamma_0=0.$$

where

$$\eta\left(\theta\right)=-\frac{\theta q'\left(\theta\right)}{q\left(\theta\right)}=\frac{\frac{\partial M(U,V)}{\partial U}U}{M\left(U,V\right)}$$

is the elasticity of the matching function respect to unemployment.

Recall that in equilibrium, we have (with $r=0$)

$$(1-\beta)(y-z)-\frac{s+\beta\theta q(\theta)}{q(\theta)}\gamma_0=0.$$

Comparing these two conditions we find that efficiency obtains if and only if

$$\beta=\eta(\theta).$$

In other words, efficiency requires the bargaining power of the worker to be equal to the elasticity of the matching function with respect to unemployment.

We can also note that this result is made possible by the fact that the matching function is constant returns to scale, and efficiency would never obtain if it exhibited increasing or increasing returns to scale. (Why is this? How would go about proving this?)

The condition $\beta=\eta(\theta)$ is the famous *Hosios condition.* It requires the bargaining power of a factor to be equal to the elasticity of the matching function with respect to the corresponding factor.

What is the intuition?

It is not easy to give an intuition for this result, but here is an attempt: as a planner you would like to increase the number of vacancies to the point where the marginal benefit in terms of additional matches is equal to the cost. In equilibrium, vacancies enter until the marginal benefits in terms of their bargained returns is equal to the cost. So if β is too high, they are getting too small a fraction of the return, and they will not enter enough. If β is too low, then they are getting too much of the surplus, so there will be excess entry. The right value of β turns out to be the one that is equal to the elasticity of the matching function with respect to unemployment (thus $1-\beta$ is equal to the elasticity of the matching function with respect to vacancies, by constant returns to scale).

Exactly the same result holds when we have discounting, i.e., $r>0$

In this case, the objective function is

$$SS^* = \int_0^\infty e^{-rt}\left[Ny - zN - \gamma_0\theta(L-N)\right]dt$$

and will be maximized subject to

$$\dot{N} = q(\theta)\theta(L-N) - sN$$

The first-order condition is

$$y - z - \frac{r+s+\eta(\theta)q(\theta)\theta}{q(\theta)\left[1-\eta(\theta)\right]}\gamma_0 = 0$$

Compared to the equilibrium where

$$(1-\beta)[y-z] + \frac{r+s+\beta q(\theta)\theta}{q(\theta)}\gamma_0 = 0$$

Again, $\eta(\theta)=\beta$ would decentralized the constrained efficient allocation.

At this point, you may be puzzled. Isn't there unemployment in equilibrium? So the equilibrium being efficient means that the social planner likes unemployment too. This raises the question: What is the use of unemployment?

The answer to this question is quite revealing. Unemployment in fact has a social role in this model. Its role is to facilitate trade at low transaction costs; the greater is unemployment, the less costly this is to fill vacancies (which are in turn costly to open). This highlights why the bargaining parameter should be related to the elasticity of the matching function. The greater is this elasticity, it means that the more important it is to have more unemployed workers around to facilitate matching, and that means a high shadow value of unemployed workers, which corresponds to a high β in equilibrium.

4. Endogenous Job Destruction

So far we treated the rate at which jobs get destroyed as a constant, s, giving us a simple equation

$$\dot{u} = s(1-u) - \theta q(\theta) u$$

But presumably thinking of job destruction as exogenous is not satisfactory. Firms decide when to expand and contract, so it's a natural next step to endogenize s.

To do this, suppose that each firm consists of a single job (so we are now taking a position on for size). Also assume that the productivity of each firm consists of two components, a common productivity and a firm-specific productivity.

In particular

$$\text{productivity for firm } i = \underbrace{p}_{\text{common productivity}} + \underbrace{\sigma \times \varepsilon_i}_{\text{firm-specific}}$$

where

$$\varepsilon_i \sim F(\cdot)$$

over support $\underline{\varepsilon}$ and $\bar{\varepsilon}$, and σ is a parameter capturing the importance of firm-specific shocks.

Moreover, suppose that each new job starts at $\varepsilon = \bar{\varepsilon}$, but does not necessarily stay there. In particular, there is a new draw from $F(\cdot)$ arriving at the flow the rate λ.

To simplify the discussion, let us ignore wage determination and set

$$w = b$$

This then gives the following value function (written in steady state) for a an active job with productivity shock ε (though this job may decide not to be active):

$$rJ^F(\varepsilon) = p + \sigma\varepsilon - b + \lambda\left[\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \max\{J^F(x), J^V\} dF(x) - J^F(\varepsilon)\right]$$

where J^V is the value of a vacant job, which is what the firm becomes if it decides to destroy. The max operator takes care of the fact that the firm has a choice after the realization of the new shock, x, whether to destroy or to continue.

Since with free entry $J^V = 0$, we have

$$rJ^F(\varepsilon) = p + \sigma\varepsilon - b + \lambda \left[E(J^F) - J^F(\varepsilon)\right] \tag{11.1}$$

where now we write $J^F(\varepsilon)$ to denote the fact that the value of employing a worker for a firm depends on firm-specific productivity.

$$E(J^F) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \max\left\{J^F(x), 0\right\} dF(x) \tag{11.2}$$

is the expected value of a job after a draw from the distribution $F(\varepsilon)$.

Given the Markov structure, the value conditional on a draw does not depend on history.

What is the intuition for this equation?

Differentiation of (11.1) immediately gives

$$\frac{dJ^F(\varepsilon)}{d\varepsilon} = \frac{\sigma}{r+\lambda} > 0 \tag{11.3}$$

Greater productivity gives greater values the firm.

When will job destruction take place?

Since (11.3) establishes that J^F is monotonic in ε, job destruction will be characterized by a cut-off rule, i.e.,

$$\exists\, \varepsilon_d : \varepsilon < \varepsilon_d \longrightarrow \text{ destroy}$$

Clearly, this cutoff threshold will be defined by

$$rJ^F(\varepsilon_d) = 0$$

But we also have $rJ^F(\varepsilon_d) = p + \sigma\varepsilon_d - b + \lambda\left[E(J^F) - J^F(\varepsilon_d)\right]$, which yields an equation for the value of a job after a new draw:

$$E(J^F) = -\frac{p + \sigma\varepsilon_d - b}{\lambda} > 0$$

This is an interesting result; it implies that since the expected value of continuation is positive (remember equation (11.2)), the flow profits of the marginal job, $p+\sigma\varepsilon_d-b$, must be negative. Why is this? The answer is option value. Continuing as a productive unit means that the firm has the option of getting a better draw in the future, which is potentially profitable. For this reason it waits until current profits are sufficiently negative to destroy the job; in other words there is a natural form of labor hoarding in this economy.

Furthermore, we have a tractable equation for $J^F(\varepsilon)$:

$$J^F(\varepsilon) = \frac{\sigma}{r+\lambda}(\varepsilon - \varepsilon_d)$$

Let us now make more progress towards characterizing $E(J^F)$

By definition, we have

$$E(J^F) = \int_{\varepsilon_d}^{\bar{\varepsilon}} J^F(x)dF(x)$$

(where we have used the fact that when $\varepsilon < \varepsilon_d$, the job will be destroyed).

Now doing integration by parts, we have

$$\begin{aligned} E(J^F) &= \int_{\varepsilon_d}^{\bar{\varepsilon}} J^F(x)dF(x) = J^F(x)F(x)\big|_{\varepsilon_d}^{\bar{\varepsilon}} - \int_{\varepsilon_d}^{\bar{\varepsilon}} F(x)\frac{dJ^F(x)}{dx}dx \\ &= J^F(\bar{\varepsilon}) - \frac{\sigma}{\lambda+r}\int_{\varepsilon_d}^{\bar{\varepsilon}} F(x)dx \\ &= \frac{\sigma}{\lambda+r}\int_{\varepsilon_d}^{\bar{\varepsilon}} \left[1 - F(x)\right]dx \end{aligned}$$

where the last line use the fact that $J^F(\varepsilon) = \frac{\sigma}{\lambda+r}(\varepsilon - \varepsilon_d)$, so incorporates $J^F(\bar{\varepsilon})$ into the integral

Next, we have that

$$\begin{aligned} \underbrace{p + \sigma\varepsilon_d - b}_{\text{profit flow from marginal job}} &= -\frac{\lambda\sigma}{r+\lambda}\int_{\varepsilon_d}^{\bar{\varepsilon}} \left[1 - F(x)\right]dx \\ &< 0 \text{ due to option value} \end{aligned}$$

which again highlights the hoarding result. More importantly, we have

$$\frac{d\varepsilon_d}{d\sigma} = \frac{p-b}{\sigma}\left[\sigma\left(\frac{r+\lambda F(\varepsilon_d)}{r+\lambda}\right)\right]^{-1} > 0.$$

which implies that when there is more dispersion of firm-specific shocks, there will be more job destruction

The job creation part of this economy is similar to before. In particular, since firms enter at the productivity $\bar{\varepsilon}$, we have

$$\begin{aligned} q(\theta) J^F(\bar{\varepsilon}) &= \gamma_0 \\ &\Longrightarrow \frac{\gamma_0(r+\lambda)}{\sigma(\bar{\varepsilon}-\varepsilon_d)} = q(\theta) \end{aligned}$$

Recall that as in the basic search model, job creation is "sluggish", in the sense that it is dictated by the matching function; it cannot jump it can only increase by investing more resources in matching.

On the other hand, job destruction is a jump variable so it has the potential to adjust much more rapidly (this feature was emphasized a lot when search models with endogenous job-destruction first came around, because at the time the general belief was that job destruction rates were more variable than job creation rates; now it's not clear whether this is true; it seems to be true in manufacturing, but not in the whole economy).

The Beveridge curve is also different now. Flow into unemployment is also endogenous, so in steady-state we need to have

$$\lambda F(\varepsilon_d)(1-u) = q(\theta)\theta u$$

In other words:

$$u = \frac{\lambda F(\varepsilon_d)}{\lambda F(\varepsilon_d) + q(\theta)\theta},$$

which is very similar to our Beveridge curve above, except that $\lambda F(\varepsilon_d)$ replaces s.

The most important implication of this is that shocks (for example to productivity) now also shift the Beveridge curve shifts. For example, an increase in p will cause an inward shift of the Beveridge curve; so at a given level of creation, unemployment will be lower.

How do you think endogenous job destruction affects efficiency?

5. A Two-Sector Search Model

Now consider a two-sector version of the search model, where there are skilled and unskilled workers. In particular, suppose that the labor force consists of L_1 and L_2 workers, i.e.

L_1 : unskilled worker
L_2 : skilled worker

Firms decide whether to open a skilled vacancy or an unskilled vacancy.

$$\left.\begin{array}{l} M_1 = x(U_1, V_1) \\ M_2 = x(U_2, V_2) \end{array}\right\} \text{ the same matching function in both sectors.}$$

Opening vacancies is costly in both markets with

$$\begin{array}{ll} \gamma_1 : & \text{cost of vacancy for unskilled worker} \\ \gamma_2 : & \text{cost of vacancy for skilled worker.} \end{array}$$

As before, shocks arrive at some rate, here assumed to be exogenous and potentially different between the two types of jobs

$$s_1, s_2 : \text{ separation rates}$$

Finally, we allow for population growth of both skilled unskilled workers to be able to discuss changes in the composition of the labor force. In particular, let the rate of population growth of L_1 and L_2 be n_1 and n_2 respectively.

$$n_1, n_2 : \text{ population growth rates}$$

This structure immediately implies that there will be two separate Beveridge curves for unskilled and skilled workers, given by

$$u_1 = \frac{s_1 + n_1}{s_1 + n_1 + \theta_1 q(\theta_1)} \qquad u_2 = \frac{s_2 + n_2}{s_2 + n_2 + \theta_2 q(\theta_2)}.$$

(can you explain these equations? Derive them?)

So different unemployment rates are due to three observable features, separation rates, population growth and job creation rates.

The production side is largely the same as before

$$\text{output} \quad Af(K, N)$$

where N is the effective units of labor, consisting of skilled and unskilled workers.

We assumed that each unskilled worker has one unit of effective labor, while each skilled worker has $\eta > 1$ units of effective labor.

Finally, the interest rate is still r and the capital depreciation rate is δ.

Asset Value Equations are as before.

For filled jobs,

$$\begin{aligned} rJ_1^F &= Af(k) - (r+\delta)k - w_1 - s_1 J_1^F \\ rJ_2^F &= Af(k)\eta - (r+\delta)k\eta - w_2 - s_2 J_2^F \end{aligned}$$

While for vacancies, we have

$$\begin{aligned} rJ_1^V &= -\gamma_1 + q(\theta_1)(J_1^F - J_1^V) \\ rJ_2^V &= -\gamma_2 + q(\theta_2)(J_2^F - J_2^V) \end{aligned}$$

Zero profit for opening jobs in both sectors implies

$$J_1^V = J_2^V = 0$$

Using this, we have the value of filled jobs in the two sectors

$$J_1^F = \frac{\gamma_1}{q(\theta_1)} \quad \text{and} \quad J_2^F = \frac{\gamma_2}{q(\theta_2)}$$

The worker side is also identical, especially since workers don't have a choice affecting their status. In particular,

$$\begin{aligned} rJ_1^U &= z + \theta_1 q(\theta_1)(J_1^E - J_1^U) \\ rJ_2^U &= z + \theta_2 q(\theta_2)(J_2^E - J_2^U) \end{aligned}$$

where we have assumed the unemployment benefit is equal for both groups (this is not important, what's important is that unemployment benefits are not proportional to equilibrium wages).

Finally, the value of being employed for the two types of workers are

$$rJ_i^E = w_i - s(J_i^E - J_i^U)$$

The structure of the equilibrium is similar to before, in particular the modified golden rule and the two wage equations are:

$$\begin{aligned} Af'(k) &= r+\delta \qquad \text{M.G.R.} \\ w_1 &= (1-\beta)z+\beta\left[Af(k)-(r+\delta)k+\theta_1\gamma_1\right] \\ w_2 &= (1-\beta)z+\delta\left[Af(k)\eta-(r+\delta)k\eta+\theta_2\gamma_2\right] \end{aligned}$$

The most important result here is that wage differences between skilled unskilled workers are compressed.

To illustrate this, let us take a simple case and suppose first that

$$\gamma_1=\gamma_2,\ n_1=n_2,\ s_1=s_2,\ z=0.$$

Thus there are no differences in costs of creating vacancies, separation rates, unemployment benefits, and population growth rates between skilled and unskilled workers.

Then we have

$$u_2 > u_1$$

Why? Let's see

$$\begin{aligned} J_1^F &= \frac{\gamma}{q(\theta_1)} \qquad \text{and} \qquad J_2^F=\frac{\gamma}{q(\theta_2)} \\ J_2^F &> J_1^F \Longrightarrow \theta_1<\theta_2 \Longrightarrow u_1>u_2. \end{aligned}$$

High skill jobs yield higher rents, so everything else equal firms will be keener to create these types of jobs, and the only thing that will equate their marginal profits is a slower rate of finding skilled workers, i.e., a lower rate of unemployment for skilled than unskilled workers

There are also other reasons for higher unemployment for unskilled workers.

Also, $s_1 > s_2$ but lately $n_1 < n_2$ so the recent fall in n_1 and increase in n_2 should have helped unskilled unemployment.

But $z\uparrow$ has more impact on unskilled wages.

$\eta\uparrow\Longrightarrow$ "skill-biased" technological change.

$$\Longrightarrow \quad u_1 = cst,\ w_1 = cst$$
$$u_2 \downarrow,\ w_2 \uparrow$$

A set of interesting effects happen when r are endogenous. What are they?

Suppose we have $\eta \uparrow$, this implies that demand for capital goes up, and this will increase the interest rate, i.e., $r \uparrow$

The increase in the interest rate will cause

$$u_1 \uparrow, w_1 \downarrow .$$

What about labor force participation? Can this model explain non-participation?

Suppose that workers have outside opportunities distributed in the population, and they decide to take these outside opportunities if the market is not attractive enough. Suppose that there are N_1 and N_2 unskilled and skilled workers in the population. Each unskilled worker has an outside option drawn from a distribution $G_1(v)$, while the same distribution is $G_2(v)$ for skilled workers. In summary:

$$\begin{array}{ll} G_1(v) & N_1 : \text{ unskilled} \\ G_2(v) & N_2 : \text{ skilled} \end{array}$$

Given v; the worker has a choice between J_i^U and v.

Clearly, only those unskilled workers with

$$J_1^U \geq v$$

will participate and only skilled workers with

$$J_2^U \geq v$$

(why are we using the values of unemployed workers and not employed workers?)

Since L_1 and L_2 are irrelevant to steady-state labor market equilibrium above (because of constant returns to scale), the equilibrium equations are unchanged. Then,

$$\begin{aligned} L_1 &= N_1 \int_0^{J_1^U} dG_1(v) \\ L_2 &= N_2 \int_0^{J_2^U} dG_2(v). \end{aligned}$$

$\eta \uparrow,\ r \uparrow \Longrightarrow u_1 \uparrow,\ w_1 \downarrow\ J_1^U \downarrow$

$\Longrightarrow$ unskilled participation falls. (consistent with Juhn-Murphy and Topel's findings on US labor markets in the 1980s).

But this mechanism requires an interest rate response. Is the interest rate higher in the '80s?

Alternative formulation: the skilled do the unskilled jobs and there are not so many jobs (demand??). This takes us the next topic.

CHAPTER 12

Composition of Jobs

Search models, and more generally models with frictional labor markets, also provided a useful perspective for thinking about the endogenous composition of jobs. The "composition of jobs" here refers to the quality distribution of jobs, for example, some jobs may involve higher quality or newer vintage machines or more physical capital, and the same worker will be more productive in these jobs than others with lower quality machines or less physical capital. An investigation of the composition of jobs is interesting in part because this is one of the main margins in which labor markets may have different degrees of success in achieving and efficient allocation. For example, depending on labor market institutions or other features of the environment, the equilibrium may or may not involve the "appropriate" allocation of workers to firms, or the creation of the right types of jobs.

1. Endogenous Composition of Jobs with Homogeneous Workers

Let us start with the simplest setup, in which workers are homogeneous, but they can be employed in two different types of jobs. Labor and capital are used to produce two non-storable intermediate goods that are then sold in a competitive market and immediately transformed into the final consumption good. Preferences of all agents are defined over the final consumption good alone. Let us normalize the price of the final good to 1.

There is a continuum of identical workers with measure normalized to 1. All workers are infinitely lived and risk-neutral. They derive utility from the consumption of the unique final good and maximize the present discounted value of their utility. Time is continuous and the discount rate of workers is equal to r. On the other side of the market, there is a larger continuum of firms that are also risk-neutral with discount rate r.

The technology of production for the final good is:

$$Y = \left(\alpha Y_b^\rho + (1-\alpha) Y_g^\rho\right)^{1/\rho} \tag{12.1}$$

where Y_g is the aggregate production of the first input, and Y_b is the aggregate production of the second input, and $\rho < 1$. The elasticity of substitution between Y_g and Y_b is $1/(1-\rho)$ and α parameterizes the relative importance of Y_b. The subscripts g and b refer "good" and "bad" jobs as it will become clear shortly. This formulation captures the idea that there is some need for diversity in overall consumption/production, and is also equivalent to assuming that (12.1) is the utility function defined over the two goods.

Since the two intermediate goods are sold in competitive markets, their prices are:

$$\begin{aligned} p_b &= \alpha Y_b^{\rho-1} Y^{1-\rho} \\ p_g &= (1-\alpha) Y_g^{\rho-1} Y^{1-\rho} \end{aligned} \tag{12.2}$$

The technology of production for the inputs is Leontieff. When matched with a firm with the necessary equipment (capital k_b or k_g), a worker produces 1 unit of the respective good. The equipment required to produce the first input costs k_g while the cost of equipment for the second input is k_b. Let us assume that

$$k_g > k_b.$$

Before we move to the search economy, it is useful to consider the perfectly competitive benchmark. Since $k_g > k_b$, in equilibrium, we will have

$$p_g > p_b.$$

But firms hire workers at the common wage, w, irrespective of their sector. Thus, there will be neither wage differences nor bad nor good jobs. Also, since the first welfare theorem applies to this economy, the composition of output will be optimal.

Given the setup so far we can obtain the main idea before presenting the detailed analysis. As soon as we enter the world of search, there will be some rent-sharing. This implies that a worker who produces a higher valued output will receive a higher wage. As noted above, because $k_g > k_b$, the input which costs more to produce will command a higher price, thus in equilibrium $p_g > p_b$. Rent-sharing, then, leads to equilibrium wage differentials across identical workers. That is, $w_g > w_b$. Hence, the terms *good* and *bad* jobs. Next, it is intuitive that since, compared to the economy with competitive labor markets, good jobs have higher relative labor costs, their relative production will be less than optimal. In other words, the proportion of good (high-wage) jobs will be too low compared to what a social planner would choose. The rest of this section will formally analyze the search economy and establish these claims. It will then demonstrate that higher minimum wages and more generous unemployment benefits will improve the composition of jobs and possibly welfare.

1.1. The Technology of Search. As in the canonical search model, firms and workers come together via a matching technology $M(u, v)$ where u is the unemployment rate, and v is the vacancy rate (the number of vacancies). Once again, we assume that search is *undirected*, thus both types of vacancies have the same probability of meeting workers, and it is the total number of vacancies that enters the matching function. $M(u, v)$ is twice differentiable and increasing in its arguments and exhibits constant returns to scale. This enables me to write the flow rate of match for a vacancy as

$$\frac{M(u, v)}{v} = q(\theta),$$

where $q(.)$ is a differentiable decreasing function and

$$\theta = \frac{v}{u}$$

is the tightness of the labor market. It also immediately follows from the constant returns to scale assumption that the flow rate of match for an unemployed worker is

$$\frac{M(u, v)}{u} = \theta q(\theta).$$

In general, $q(\theta)$, $\theta q(\theta) < \infty$, thus it takes time for workers and firms to find suitable production partners. We also make the standard Inada-type assumptions on $M(u, v)$ which ensure that $\theta q(\theta)$ is increasing in θ, and that $\lim_{\theta\to\infty} q(\theta) = 0$, $\lim_{\theta\to 0} q(\theta) = \infty$, $\lim_{\theta\to\infty} q(\theta)\theta = 0$ and $\lim_{\theta\to 0} q(\theta)\theta = \infty$.

All jobs end at the exogenous flow rate s, and in this case, the firm becomes an unfilled vacancy and the worker becomes unemployed. Finally, there is free entry into both good and bad job vacancies, therefore both types of vacancies should expect zero net profits.

Let us denote the flow return from unemployment by z which will be thought as the level of unemployment benefit financed by lump-sum taxation. As usual, we assume that wages are determined by asymmetric Nash Bargaining where the worker has bargaining power β. Nash Bargaining per se is not essential, though rent-sharing is crucial for the results.

Firms can choose either one of two types of vacancies: *(i)* a vacancy for a intermediate good 1 - a *good job*; *(ii)* a vacancy for an intermediate good 2 - *a bad job*. Therefore, before opening a vacancy a firm has to decide which input it will produce, and at this point, it will have to buy the equipment that costs either k_b or k_g. The important aspect is that these *creation* costs are incurred before the firm meets its employees; this is a reasonable assumption, since, in practice, k corresponds to the costs of machinery, which are sector and occupation specific.

1.2. The Basic Bellman Equations. As usual, we will solve the model via a series of Bellman equations. We denote the discounted value of a vacancy by J^V, of a filled job by J^F, of being unemployed by J^U and of being employed by J^E. We will use subscripts b and g to denote good and bad jobs. We also denote the proportion of bad job vacancies among all vacancies by ϕ. Then, in steady state:

$$rJ^U = z + \theta q(\theta)\left[\phi J_b^E + (1-\phi)J_g^E - J^U\right] \tag{12.3}$$

Being unemployed is similar to holding an asset; this asset pays a dividend of z, the unemployment benefit, and has a probability $\theta q(\theta)\phi$ of being transformed into a bad job in which case the worker obtains J_b^E, the asset value of being employed in a bad job, and loses J^U; it also has a probability $\theta q(\theta)(1-\phi)$ of being transformed into a good job, yielding a capital gain $J_g^E - J^U$ (out of steady state, $\dot{J}^U$ has to be added to the right-hand side to capture future changes in the value of unemployment). Observe that this equation is written under the implicit assumption that workers will not turn down jobs, which we will discuss further below. The steady state discounted present value of employment can be written as:

$$rJ_i^E = w_i + s(J^U - J_i^E) \tag{12.4}$$

for $i = b, g$. (12.4) has a similar intuition to (12.3).

Similarly, when matched, both vacancies produce 1 unit of their goods, so:

$$rJ_i^F = p_i - w_i + s\left(J_i^V - J_i^F\right) \tag{12.5}$$

$$rJ_i^V = q(\theta)\left(J_i^F - J_i^V\right) \tag{12.6}$$

for $i = b, g$, where we have ignored the possibility of voluntary job destruction which will never take place in steady state.

Since workers and firms are risk-neutral and have the same discount rate, Nash Bargaining implies that w_b and w_g will be chosen so that:

$$\begin{aligned}(1-\beta)(J_b^E - J^U) &= \beta(J_b^F - J_b^V)\\ (1-\beta)(J_g^E - J^U) &= \beta(J_g^F - J_g^V)\end{aligned} \tag{12.7}$$

Note that an important feature is already incorporated in these expressions: workers cannot pay to be employed in high wage jobs: due to search frictions, at the moment a worker finds a job, there is bilateral monopoly, and this leads to rent-sharing over the surplus of the match.

As there is free-entry on the firm side, it should not be possible for an additional vacancy to open and make expected net profits. Hence:

$$J_i^V = k_i. \tag{12.8}$$

Finally, the steady state unemployment rate is given by equating flows out of unemployment to the number of destroyed jobs. Thus:

$$u = \frac{s}{s + \theta q(\theta)}. \tag{12.9}$$

1.3. Characterization of Steady State Equilibria. A steady state equilibrium is defined as a proportion ϕ of bad jobs, tightness of the labor market θ, value functions J_b^V, J_b^F, J_b^E, J_g^V, J_g^F, J_g^E and J^U, prices for the two goods, p_b and p_g such that equations (12.2), (12.3), and (12.4), (12.5), (12.6), (12.7) and (12.8) for both $i = b$ and g are satisfied. The steady state unemployment rate is then given by (12.9).[1]

In steady state, both types of vacancies meet workers at the same rate, and in equilibrium workers accept both types of jobs, therefore $Y_b = (1-u)\phi$ and $Y_g = (1-u)(1-\phi)$. Then, from (12.2), the prices of the two inputs can be written as:

$$\begin{aligned} p_g &= (1-\alpha)(1-\phi)^{\rho-1}\left[\alpha\phi^\rho + (1-\alpha)(1-\phi)^\rho\right]^{\frac{1-\rho}{\rho}} \\ p_b &= \alpha\phi^{\rho-1}\left[\alpha\phi^\rho + (1-\alpha)(1-\phi)^\rho\right]^{\frac{1-\rho}{\rho}}. \end{aligned} \tag{12.10}$$

Simple algebra using (12.4), (12.5), (12.7) and (12.8) gives:

$$w_i = \beta\left(p_i - rk_i\right) + (1-\beta)rJ^U \tag{12.11}$$

as the wage equation. Intuitively, the surplus that the firm gets is equal to the value of output which is p_i minus the flow cost of the equipment, rk_i. The worker gets a share β of this, plus $(1-\beta)$ times his outside option, rJ^U. Using (12.5) and (12.6), the zero-profit condition (12.8) can be rewritten as:

$$\frac{q(\theta)(1-\beta)\left(p_b - rJ^U\right)}{r + s + (1-\beta)q(\theta)} = rk_b \tag{12.12}$$

$$\frac{q(\theta)(1-\beta)\left(p_g - rJ^U\right)}{r + s + (1-\beta)q(\theta)} = rk_g. \tag{12.13}$$

[1] One might wonder at this point whether a different type of equilibrium, with $J^U = J_b^E$ and workers accepting bad jobs with probability $\zeta < 1$, could exist. The answer is no. From equation (8.1), this would imply $J_b^V = J_b^F$, but in this case, firms could never recover their upfront investment costs.

A firm buys equipment that costs k_i, which remains idle for a while due to search frictions (i.e. because $q(\theta) < \infty$). This cost is larger for firms that buy more expensive equipment and open good jobs. They need to recover these costs in the form of a higher net flow profits: i.e. $p_g - rk_g > p_b - rk_b$. From rent-sharing, this immediately implies that $w_g > w_b$. More specifically, combining (12.11), (12.12) and (12.13), we get :

$$w_g - w_b = \frac{(r+s)\beta(rk_g - rk_b)}{(1-\beta)\,q(\theta)} > 0 \tag{12.14}$$

Therefore, wage differences are related to the differences in capital costs and also to the average duration of a vacancy. In particular, when $q(\theta) \to \infty$, the equilibrium converges to the Walrasian limit point, and both w_g and w_b converge to rJ^U, so wage differences disappear. The reason is that in this limit point, capital investments never remain idle, thus good jobs do not need to make higher net flow profits. Also, with equal creation costs, i.e., $k_b = k_g$, wage differentials disappear again.

Finally, (12.3) gives the value of an unemployed worker as

$$rJ^U = G(\theta, \phi) \equiv \frac{(r+s)z + \beta\theta q(\theta)\left[\phi(p_b - rk_b) + (1-\phi)(p_g - rk_g)\right]}{r + s + \beta\theta q(\theta)} \tag{12.15}$$

It can easily be verified that $G(.,.)$ is continuous, strictly increasing in θ, and strictly decreasing in ϕ. Intuitively, as the tightness of the labor market, θ, increases, workers find jobs faster, thus rJ^U is higher. Also as ϕ decreases, the greater fraction of good jobs among vacancies increases the value of being unemployed since $w_g > w_b$ (i.e., $J_g^V > J_b^E$). The dependence of rJ^U on ϕ is the general equilibrium effect mentioned in the introduction: as the composition of jobs changes, the option value of being unemployed also changes.

A steady-state equilibrium is characterized by the intersection of two loci: *bad job locus,* (12.12), and the *good job locus,* (12.13) (both evaluated with (12.10) and (12.15) substituted in).

The next figure draws these two loci in the θ-ϕ plane.

In this figure, the curve for (12.13), along which a firm that opens a good job vacancy makes zero-profits, is upward sloping: a higher value of ϕ increases the

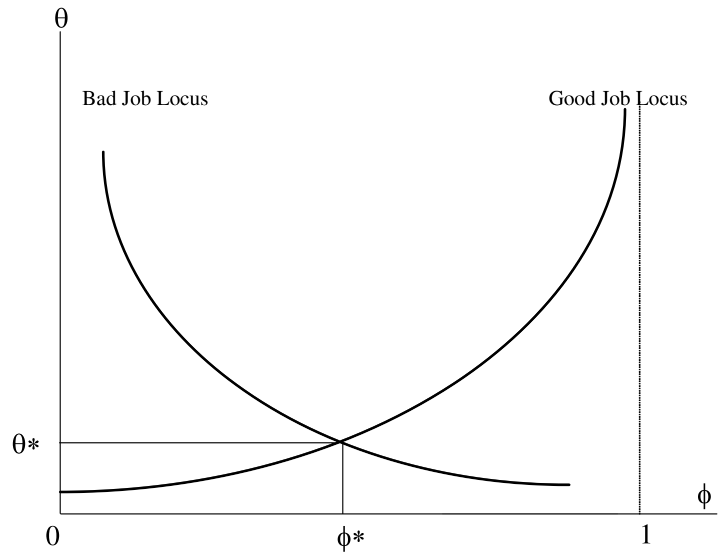


Figure 12.1

left hand side, thus θ needs to change to increase the right-hand side (and reduce the left-hand side through $G(\theta,\phi)$). Intuitively, an increase in ϕ implies a higher p_g (from equation (12.10)). So to ensure zero profits, θ needs to increase to raise the duration of vacancies. In contrast, (12.12) cannot be shown to be decreasing everywhere. Intuitively, an increase in ϕ reduces p_b, thus requires a fall in θ to equilibrate the market, but the general equilibrium effect through J^U (i.e. that a fall in ϕ reduces J^U) counteracts this and may dominate. This issue is discussed further below.

Here, let us start with the case in which $\rho \leq 0$, so that good and bad jobs are gross complements. In this case, it is straightforward to see that as ϕ tends to 1, (12.12) gives $\theta \to \infty$ whereas (12.13) implies $\theta \to 0$. Thus, the bad job locus is above the good job locus. The opposite is the case as ϕ goes to zero. Then by the continuity of the two functions, they must intersect at least once in the range $\phi \in (0,1)$. Therefore, we can conclude that there always exists a steady state equilibrium with $\phi \in (0,1)$ always exists and is characterized by (12.10), (12.11), (12.12), (12.13) and (12.15). In equilibrium, for all $k_g > k_b$, we have $p_g > p_b$ and $w_g > w_b$.

When $\rho > 0$, an equilibrium continues to exist, but does not need to be interior, so one of (12.12) and (12.13) may not hold. We now discuss a particular case of this.

1.4. Multiple equilibria. Since (12.12) can be upward sloping over some range, more than one intersections, hence multiple equilibria, are possible. (12.12) is more likely to be upward sloping when relative prices change little as a result of a change in the composition of jobs. Therefore, to illustrate the possibility of multiple equilibria, let us consider the extreme case where $\rho = 1$, so that goods g and b are perfect substitutes, and there are no relative price effects. Furthermore, we assume that

$$1 - 2\alpha > r(k_g - k_b).$$

In the absence of this assumption, good jobs are not productive enough, and will never exist in equilibrium.

The absence of substitution between good and bad jobs immediately implies that

$$p_g = 1 - \alpha > p_b = \alpha.$$

The equilibrium can then be characterized diagrammatically. To do this, totally differentiate (12.12) and (12.13), with $p_g = 1 - \alpha$ and $p_b = \alpha$, which gives

$$\left.\frac{d\theta}{d\phi}\right|_i = \frac{-\frac{\partial G(\theta,\phi)}{\partial \phi}}{\frac{\partial G(\theta,\phi)}{\partial \theta} - k_i \frac{(r+s)(1-\beta)q'(\theta)}{(1-\beta)q(\theta)^2}\frac{\partial G(\theta,\phi)}{\partial \theta}} > 0 \qquad (12.16)$$

where $i = b$ is zero profit condition for bad jobs, (12.12), and $i = g$ is the zero profit condition for good jobs, (12.13). The derivative in (12.16) is positive, irrespective of whether it is for good or bad jobs, because $rJ^U = G(\theta, \phi)$ is decreasing in ϕ and increasing in θ, while $q'(\theta) < 0$. Since $k_b < k_g$, this equation also immediately implies that (12.12) is steeper than (12.13). So (12.12) has to intersect (12.13) from below if at all, in which case there will be three equilibria. This is shown in the next figure.

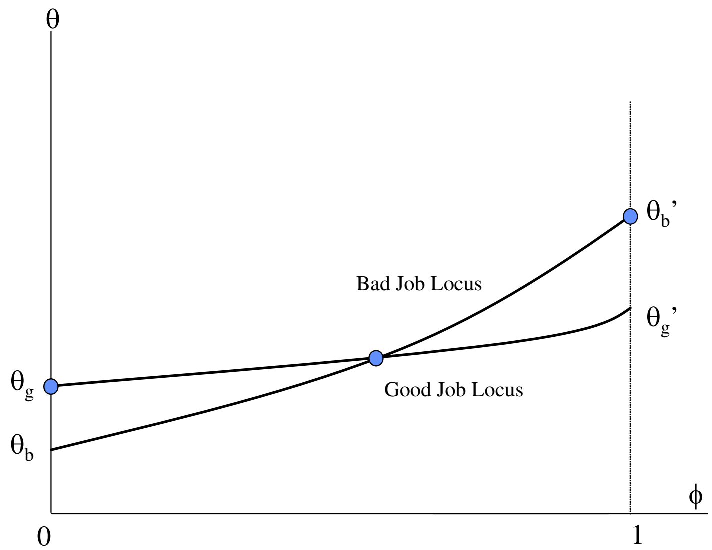


Figure 12.2

The first is a "mixed strategy" equilibrium at the point where the two curves intersect. The other two equilibria are more interesting. When $\phi = 0$, we have $\theta_g > \theta_b$, so that it is more profitable to open a good job. Hence there is an equilibrium in which all firms open good jobs. It is not profitable for firms to open a bad job, because when $\phi = 0$, workers receive high wages and have attractive outside options; so a firm that opens a bad job will be forced to pay a relatively high wage, making a deviation to a bad job unprofitable. In contrast, at $\phi = 1$, we have $\theta_g' < \theta_b'$, so it is an equilibrium for all firms to open bad jobs.

Intuitively, when all firms open bad jobs, the outside option of workers is low, so firms bargain to low wages, making entry relatively profitable. In equilibrium, θ has to be high to ensure zero profits. But a tight labor market (a high θ) hurts good jobs relatively more since they have to make larger upfront investments. The multiplicity of equilibria in this model illustrates the strength of the general equilibrium forces that operate through the impact of job composition on the overall level of wages.

1.5. Welfare. Let us next analyze the welfare properties of equilibrium using the notion of total surplus as in the baseline search model. In this case, total surplus (in steady state) can be written as:

$$TS = (1-u)\left[\phi(p_b - rk_b) + (1-\phi)(p_g - rk_g)\right] - \theta u \left(\phi rk_b + (1-\phi)rk_g\right) \tag{12.17}$$

Total surplus is equal to total flow of net output, which consists of the number of workers in good jobs ($(1-\phi)(1-u)$) times their net output (p_g minus the flow cost of capital rk_g), plus the number of workers in bad jobs ($\phi(1-u)$) times their net product ($p_b - rk_b$), minus the flow costs of job creation for good and bad vacancies (respectively, $\theta u(1-\phi)rk_g$ and $\theta u\phi rk_b$).

It is straightforward to locate the set of allocations that maximize total social surplus. This set would be the solution to the maximization of (12.17) subject to (12.9). Inspecting the first-order conditions of this problem, it can be seen that decentralized equilibria will not in general belong to this set, thus a social planner can improve over the equilibrium allocation. The results regarding the socially optimal amount of job creation are standard: if β is too high, that is $\beta > \eta(\theta)$ where $\eta(\theta)$ is elasticity of the matching function, $q(\theta)$, then there will be too little job creation, and if $\beta < \eta(\theta)$, there will be too much. Since this paper is concerned with the composition of jobs, we will not discuss these issues in detail. Instead, we will show that irrespective of the value of θ, the equilibrium value of ϕ is always too high; that is, there are too many bad jobs relative to the number of good jobs.

To prove this claim, it is sufficient to consider the derivative of TS with respect to ϕ at $z = 0$ (note the constraint, (12.9), does not depend on ϕ):

$$\frac{dTS}{d\phi} = (1-u)\cdot\left[\frac{d(\phi p_b + (1-\phi)p_g)}{d\phi}\right] - (1-u+u\theta)\cdot\{rk_b - rk_g\} \tag{12.18}$$

For the composition of jobs to be efficient at the laissez-faire equilibrium, (12.18) needs to equal zero when evaluated in the equilibrium characterized above. Some simple algebra using (12.9), (12.10), (12.12) and (12.13) to substitute out $u,$ and k_i

gives (details of the algebra available upon request):

$$\left.\frac{dTS}{d\phi}\right|_{dec.\ eq.} = \frac{\theta q(\theta)}{s+\theta q(\theta)} \cdot \left(1 + \frac{(s+q(\theta))(1-\beta)}{r+s+(1-\beta)q(\theta)}\right) \cdot (p_b - p_g) < 0$$

This expression is always negative, irrespective of the value of θ, so starting from laissez-faire equilibrium, a reduction in ϕ will increase social surplus. Therefore, we can conclude that, given the labor market tightness θ, a surplus-maximizing social planner would choose $\phi^s(\theta) < \phi^*(\theta)$, where $\phi^*(\theta)$ is the decentralized equilibrium with $z = 0$. In other words, the equilibrium proportion of bad jobs is too high.

The intuition is simple; in a decentralized equilibrium, it is always the case that $w_g > w_b$. Yet, firms do not take into account the higher utility they provide to workers by creating a good job rather than a bad job, hence there is an uninternalized positive externality, which leads to an excessively high fraction of bad jobs in equilibrium. Search and rent-sharing are crucial for this result. Search ensures that firms have to share the ex post rents with the workers, and they cannot induce competition among workers to bid down wages. Firms would ideally like to contract with their workers on the wage rate before they make the investment decision, but search also implies that they do not know who these workers will be, thus cannot contract with them at the time of investment.

1.6. The Impact of Minimum Wages and Unemployment Benefits. As is usual in models with potential multiple equilibria, only the comparative statics of "extremal" equilibria are of interest. Therefore, let us focus on an economy where in equilibrium (12.13) cuts (12.12) from below (or alternatively, an economy with a unique equilibrium). Now consider an increase in z which corresponds to the UI system becoming more generous. Both the bad job locus, (12.12), and the good job locus, (12.13), will shift down. Hence, θ will definitely fall. It is also straightforward to verify that (12.12) will shift by more, therefore, ϕ is unambiguously reduced. Intuitively, with ϕ unchanged, relative prices and hence wages will be unchanged, but then with the higher unemployment benefits, workers would prefer to wait for good jobs rather than accept bad jobs. This increases w_b and reduces ϕ (the fraction of bad jobs).

Furthermore, a more generous unemployment benefit not only increases the fraction of good jobs, but may also increase the total number of good jobs. Totally differentiating (12.12) and (12.13), we obtain that the total number of good jobs will increase if and only if:

$$w_g - w_b > \left(\frac{1}{\eta(\theta)} - 1\right) u(1-\phi) \left(\frac{d(p_g - p_b)}{d\phi}\right)$$

where recall that $\eta(\theta)$ is the elasticity of $q(\theta)$. This inequality is likely to be satisfied when the two inputs are highly substitutable, i.e. ρ close to 1; when wage differences are large; when $\eta(\theta)$ is close to 1; and/or when unemployment is low to start with. Thus, it is only increases in unemployment benefit starting from moderate levels that increase the number of good jobs.

The impact on welfare depends on how large the effect on θ is relative to the effect on ϕ. We can see this by totally differentiating (12.17) after substituting for u. This gives a relationship between θ and ϕ, drawn as the dashed line in the next figure, along which total surplus is constant.

Shifts of this curve towards North-East give higher surplus. When this curve is steeper than (12.13), a higher z can improve welfare, and this is the case drawn in the figure. For example, if β is very low to start with, then unemployment will be too low relative to the social optimum, and in this case an increase in z will unambiguously increase total welfare.

More generally, irrespective of whether total surplus increases, a more generous unemployment benefit raises average labor productivity, $\phi p_b + (1-\phi)p_g$, which is unambiguously decreasing in ϕ. Therefore, when unemployment benefits increase, the composition of jobs shifts towards more capital intensive good jobs, and labor productivity increases.

A minimum wage has a similar effect on job composition. Consider a minimum wage $\underline{w}$ such that $w_b < \underline{w} < w_g$, so it is only binding for bad jobs. The equation for

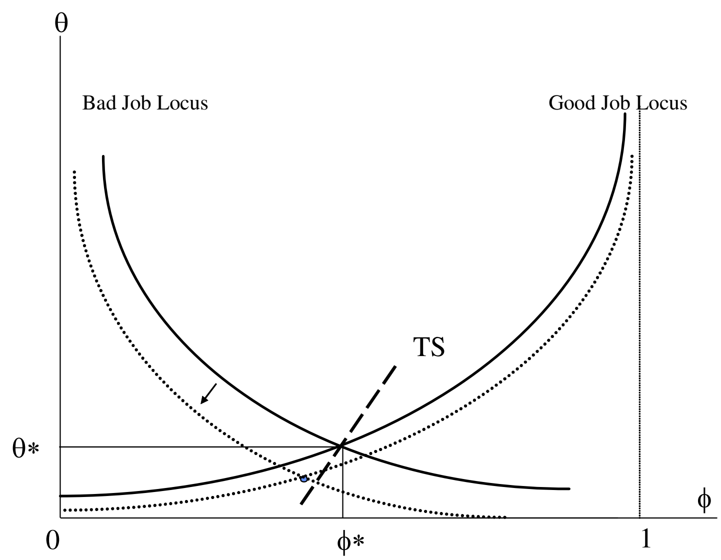


Figure 12.3

J_b^F now becomes:

$$J_b^F = \frac{p_b - \underline{w} + sk_b}{r+s}.$$

Then, (12.12) changes to:

$$q(\theta)\frac{p_b - \underline{w}}{r+s+q(\theta)} = rk_b. \tag{12.19}$$

Since at a given θ, the left-hand side of (12.19) is less than that of (12.12), the impact of higher minimum wages is to shift the bad job locus, curve (12.12), down. The good job locus is still given by (12.13), but now, combining (12.3) and (12.4),

$$rJ^U = G(\theta,\phi) \equiv \frac{(r+s)z + \beta\theta q(\theta)\left[\phi\underline{w} + (1-\phi)(p_g - rk_g)\right]}{r+s+\theta q(\theta)(1-(1-\beta)(1-\phi))}$$

Since $\underline{w} > w_b$, both curves shift down, but as in the case of unemployment benefits, (12.12) shifts down by more, so both ϕ and θ fall. Again, the rise in minimum wages can increase the number, not just the proportion, of good jobs and total welfare. Moreover, for the same decline in θ, an increase in minimum wages reduces ϕ more than an increase in z, therefore, minimum wages appear to be more powerful in shifting the composition of employment away from bad towards good jobs.

Overall, we can conclude that both the introduction of a minimum wage $\underline{w}$ and an increase in unemployment benefit z decrease θ and ϕ. Therefore, they improve the composition of jobs and average labor productivity, but increase unemployment. The impact on overall surplus is ambiguous.

2. Endogenous Composition of Jobs with Heterogeneous Workers

Now consider a somewhat more realistic environment in which workers are also of heterogeneous skills. In particular, consider a world in which workers may have high or low skills and they have to match with firms. Firms will choose the level of their capital stock before matching with the workers. The basic idea that will be highlighted by the model is that when either the productivity gap between skilled and unskilled workers is limited or when the number of skilled workers in the labor force is small, it will be profitable for firms to create jobs that to employ both skilled and unskilled workers. But when the productivity gap is large or that are a sufficient number of skilled workers, it may become profitable for (some) firms to target skilled workers, designing the jobs specifically for these workers. Then these firms will wait for the skilled workers, and will try to screen the more skill once among the applicants. In the meantime, there will be lower-quality (low capital) jobs specifically targeted at the unskilled.

Suppose that there are two types of workers. The unskilled have human capital (productivity) 1, while the skilled have human capital $\eta > 1$. Denote the fraction of skilled workers in the labor force by ϕ.

Firms choose the capital stock k before they meet a worker, and matching is assumed to be random, in the sense that each firm, irrespective of its physical capital, has exactly the same probability of meeting different types of workers. Once the firm and the worker match, separating is costly, so there is a quasi-rent to be divided between the pair. Here, the economy is assumed to last for one period, so if the firm and worker do not agree they lose all of the output (see Acemoglu, 1999, for the model where the economy is infinite-horizon and agents who do not agree with their partners can resample). Therefore, bargaining will result in workers receiving a certain fraction of output, which is again denoted by β.

The production function of a pair of worker and firm is

$$y = k^{1-\alpha} h^{\alpha},$$

where k is the physical capital of the firm and h is the human capital of the worker.

Firms choose their capital stock to maximize profits, before knowing which type of worker will apply to their job. For simplicity, we assume that firms do not bear the cost of capital if they decides not to produce with the worker who has applied to the job. We also denote the cost of capital by c.

Their expected profits are therefore given by

$$\phi x^H (1-\beta) \left(k^{1-\alpha}\eta - ck\right) + (1-\phi) x^L (1-\beta) \left(k^{1-\alpha} - ck\right),$$

where x^j is the probability, chosen by the firm, that it will produce with a worker of type j conditional on matching that type of worker. Therefore, the first term is profits conditional on matching with a skilled worker, and the second term gives the profits from matching with an unskilled worker.

There can be to different types of equilibria in this economy:

(1) A pooling equilibrium in which firms choose a level of capital and use it both of skilled and unskilled workers. We will see that in the pooling equilibrium inequality is limited.
(2) A separating equilibrium in which firms target the skilled and choose a higher level of capital. In this equilibrium inequality will be greater.

In this one-period economy, firms never specifically target the unskilled, but that outcome arises in the dynamic version of this economy.

Now it is straightforward to characterize the firms profit maximizing capital choice and the resulting organization of production (whether firms will employ both skilled and unskilled workers). It turns out that first choose the pooling strategy as long as

$$\eta < \left(\frac{1-\phi}{\phi^{\alpha} - \phi}\right)^{1/\alpha}$$

Therefore, a sufficiently large increase in η (in the relative productivity of skilled workers) and/or in ϕ (the fraction of skilled workers in the labor force) switches the economy from pooling to separating).

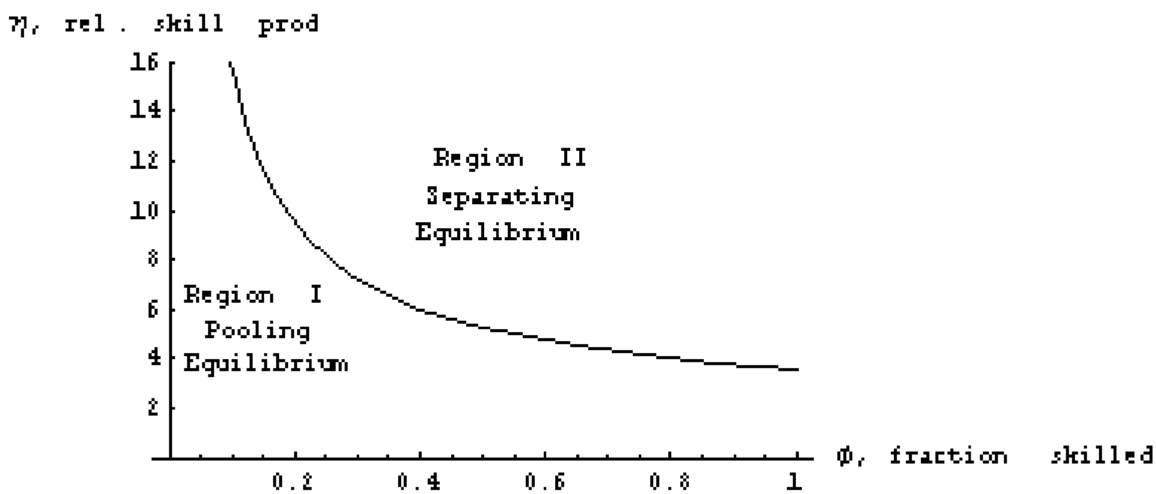


FIGURE 12.4

Such a switch will be associated with important changes in the organization of production, an increase in inequality, and a decline in the wages of low-skill workers.

Is there any evidence that there has been such a change in the organization of production? This is difficult to ascertain, but some evidence suggests that there may have been some important changes in how jobs are designed and organized now.

First, firms spend much more on recruiting, screening, and are now much less happy to hire low-skill workers for jobs that they can fill with high skill workers.

Second, as already mentioned above, the distribution of capital to labor across industries has become much more unequal over the past 25 years. This is consistent with a change in the organization of production where rather than choosing the same (or a similar) level of capital with both skilled and unskilled workers, now some firms target the skilled workers with high-capital jobs, while other firms go after unskilled workers with jobs with lower capital intensity.Third, evidence from

Figure 12.5

the CPS suggests that the distribution of jobs has changed significantly since the early 1980s, with job categories that used to pay "average wages" have declined in importance, and more jobs at the bottom and top of the wage distribution. In particular, if we classify industry-occupation cells into high-wage the middle-wage and low-wage ones (based either on wages or residual wages), there are many fewer workers employed in the middle-wage cells today as compared to the early 1980s, or the weight-at-the-tales of the vob quality distribution has increased substantially as the next figure shows.

This framework also suggests that there should be better "matching" between firms and workers now, since firms are targeting high skilled workers. Therefore,

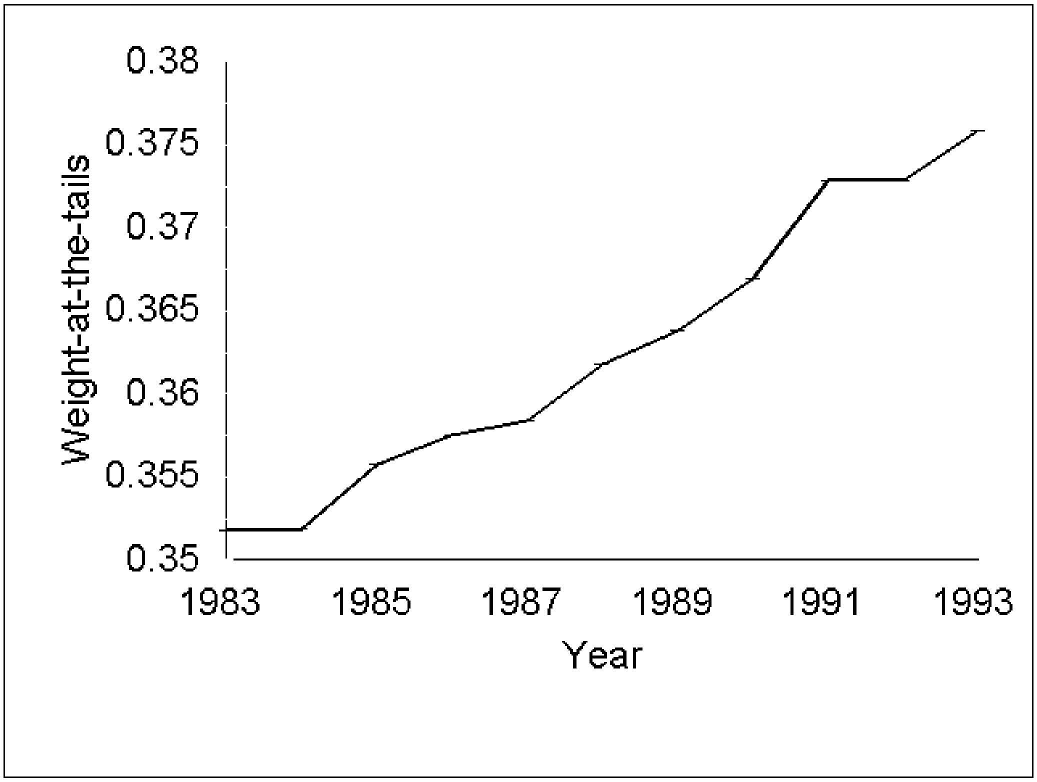


FIGURE 12.6. The evolution of the percentage of employment in the top and bottom 25 percentile industry-occupation cells (weight-at-the-tails of the job quality distribution).

measures of mismatch should have declined over the past 25 or so years. Consistent with this prediction, evidence from the PSID suggests that there is much less over- or under-education today than in the 1970s.

CHAPTER 13

Wage Posting and Directed Search

1. Inefficiency of Search Equilibria with Investments

Before turning to wage posting and directed search, let us highlight a more severe (and more fundamental) source of inefficiency in search models than the bargaining power not satisfying the Hosios condition. This results in the presence of *investments.*

Production still requires 1 firm - 1 worker, but now there is the intensive margin of capital per worker. In particular, this pair produces $f(k)$, where k is capital per worker. We assume

$$f' > 0, \quad f'' < 0$$

The most important feature is that k is to be chosen ex ante and is irreversible. The important economic implications of this are two:

(1) If there is bargaining, at this stage of bargaining, the capital is already sunk and the capital to labor ratio is irreversibly determined.
(2) While looking for a worker, the firm incurs an opportunity cost equal to be user cost of capital times the amount of capital that has, i.e., $u_k \times k$, where u_k is the user cost which will be determined below.

Trading frictions will be modeled in a way similar to before, but since my interest here is with "inefficiency," which is easily possible with increasing or decreasing returns to scale in the matching technology, I will assume constant returns to scale from the beginning. I will also develop the notation that will be useful when we look at wage posting and directed search.

First note that if $M = M(U, V)$ exhibits constant returns to scale, then exploiting the standard linear homogeneity properties, we can write

$$\begin{aligned} q &= \frac{M}{V} = M\left(\frac{U}{V}, 1\right) \\ &= q(\theta) \end{aligned}$$

where $\theta \equiv V/U$ is the tightness of the labor market (the vacancy to unemployment ratio), and the function $q(\theta)$ is decreasing in θ given our assumptions above. This means that vacancies have a harder time finding matches in a tighter labor market.

This is the standard notation in the Diamond-Mortensen-Pissarides macro search models.

Moreover,

$$\begin{aligned} p &= \frac{M}{U} = \frac{V}{U} M\left(\frac{U}{V}, 1\right) \\ &= \theta q(\theta) \end{aligned}$$

where $\theta q(\theta)$ is increasing in θ. This means that unemployed workers have an easier time finding matches in a tighter labor market.

Now let us develop a slightly different notation. Assume that if there are Q workers searching for 1 job (think of the analogy to *queues*), Q is equivalent to $1/\theta$ in the above notation.

Then with constant returns to scale, we have

$\mu(Q)$: flow rate of match for workers, assumed it is continuously differentiable and $\mu' < 0$

$\eta(Q) \equiv Q\mu(Q)$: flow rate of match for vacancy, with $\eta' > 0$

The fact that μ, η are simply functions of Q is equivalent to assuming Constant Returns to Scale.

As before let r be the rate of time preference, and s be the separation rate due to destruction of capital

Here let us change the order a little, and start with the efficient allocation, which is again a solution to the planner's problem subject to the search constraints.

The objective function of the planner can be written as:

$$\int_0^\infty e^{-rt}\left[\underbrace{\left(\mu(Q_t)\frac{f(k_t)-(r+s)k_t}{r+s}\right)u_t}_{\text{net output of a matched worker}} - \underbrace{(r+s)k_t\frac{u_t}{Q_t}}_{\text{cost of unfilled vacancies}}\right]dt$$

where u_t is the measure of unemployed workers, or alternatively the unemployment rate, at time t.

Here it is easy to see that $(r+s)k$ is the flow cost of investment, or user cost of capital, k. (k paid up front and rk opportunity cost, sk cost of destruction). The planner incurs this cost for $V_t = u_t/Q_t$ vacancies

Less obvious at first, but equally intuitive is that the value of an unemployed worker is that with probability $\mu(Q_t)$ he will find a job, in which case he will produce a net output of $f(k_t) - (r+s)k_t$, until the job is destroyed, which has discounted value $\frac{f(k_t)-(r+s)k_t}{r+s}$, thus the value of an unemployed worker is

$$\mu(Q_t)\frac{f(k_t)-(r+s)k_t}{r+s}.$$

This expression already imposes that all firms will choose the same capital level, and no segmentation in the market (Homework exercise: set up and solve this problem when the planner allows firms to choose different levels of capital).

The constraint that the planner faces is very similar to the flow constraints we saw above:

$$\dot{u}_t = s(1-u_t) - \mu(Q_t)u_t$$

This equation says that the evolution of unemployment is given by the flows into unemployment, $s(1-u_t)$, and exits from unemployment, i.e., job creation, $\mu(Q_t)u_t$.

Now we can write the Current Value Hamiltonian as

$$H(k,Q,u,\lambda) = u\left[\mu(Q)\left(\frac{f(k)}{r+s} - k\right) - \frac{(r+s)k}{Q}\right] + \lambda\left[s(1-u) - \mu(Q)u\right]$$

The necessary conditions are

$$
\begin{aligned}
H_k &= u\left(\mu(Q)\left(\frac{f'(k)}{r+s}-1\right)-\frac{(r+s)}{Q}\right)=0 \\
H_Q &= u\left(\mu'(Q)\left(\frac{f(k)}{r+s}-k-\lambda\right)+\frac{(r+s)}{Q^2}k\right)=0 \\
H_u &= \mu(Q)\left(\frac{f(k)}{r+s}-k\right)-\frac{(r+s)}{Q}k-\lambda(s+\mu(Q))=r\lambda-\dot{\lambda}
\end{aligned}
$$

Again, focusing on steady state, we impose

$$\dot{\lambda}=0$$

$$H_u=r\lambda\Longrightarrow\lambda=\frac{\mu(Q)\left(\frac{f(k)}{r+s}-k\right)-\frac{(r+s)}{Q}k}{r+s+\mu(Q)}$$

which is the shadow value of an unemployed worker. This equation has a very intuitive interpretation. The shadow value of a worker is given by the probability (flow rate) that he will create a job, which is $\mu(Q)$, and the value of the job is

$$\left(\frac{f(k)}{r+s}-k\right).$$

While unemployed, the worker induces the planner to have more vacancies open (so as to keep Q constant), hence the term

$$-\frac{(r+s)}{Q}k.$$

Finally, once the job is destroyed, which happens at the rate s, a new cycle begins, at the rate $\mu(Q)$, which gives the denominator for discounting.

The condition that $H_k=0$ gives

$$\Longrightarrow\frac{Q^S\mu(Q^S)f'(k^s)}{(r+s)(r+s+Q^S\mu(Q^S))}=1 \tag{13.1}$$

Now combining this and the value of λ obtained about with $H_u=0\Longrightarrow$

$$f(Q^S)\frac{\mu'(Q^S)}{r+s}+\frac{r+s+\mu(Q^S)+Q^S\mu'(Q^S)-(Q^S)^2\mu'(Q^S)}{(Q^S)^2}k=0 \tag{13.2}$$

Conditions (13.1) and (13.2) characterize the constrained efficient allocation.

Next, consider the equilibrium allocation. With bargaining this corresponds to:

$$
\begin{aligned}
rJ^F(k) &= f(k)-w(k)-sJ^F(k) \\
rJ^V(k) &= \eta(Q)(J^F(k)-J^V(k))-sJ^V(k)
\end{aligned}
$$

Recall that there is random matching, so Q workers for each vacancy. Then I can write

$$\begin{aligned} rJ^E(k) &= w(k) + s(J^U - J^E(k)) \\ rJ^U &= \mu(Q) \int a(k)(J^E(k) - J^U)dF(k) \end{aligned}$$

where $a(k)$ is the decision rule of the worker on whether to match with a firm with capital k, and $F(k)$ is the endogenous distribution of capital (please do not confuse this with f which is the production function).

Nash Bargaining again implies:

$$(1-\beta)(J^E(k) - J^U) = \beta(J^F(k) - J^V(k))$$

Now we will impose free entry as in the basic Mortensen-Pissarides models, so

$$J^V(k) - k = 0$$

That is, opening a job costs k (the sunk investment), and has a return of $J^V(k)$.

$$\Longrightarrow w(k) = \beta\left(f(k) - (r+s)k\right) + (1-\beta)rJ^U$$

Now use this wage rule with J^V and J^F

$$J^V(k) = \frac{\eta(Q)\left((1-\beta)f(k) + \beta\,(r+s)\,k - (1-\beta)rJ^U\right)}{(r+s)(r+s+\eta(Q))} \tag{13.3}$$

Also recall that $\eta(Q) = Q\mu\,(Q)$.

How is the capital-labor ratio chosen? Firms will clearly choose it to maximize profits: that is,

$$k \text{ maximizes } J^V(k) - k.$$

Since this is a strictly concave problem, this implies that all firms will choose the same level of capital, k^B

$\Longrightarrow$

$$F(k) \text{ is a degenerate distribution with all of its mass at } k^B$$

where

$$\frac{\eta(Q^B)(1-\beta)f'(k^B)}{(r+s)(r+s+(1-\beta)\eta(Q^B))} = 1 \tag{13.4}$$

with Q^B as the equilibrium queue length in the economy.

Now use (13.3) with J^V and J^E to obtain an equation determining Q^B.

$$\frac{\eta(Q^B)(1-\beta)f(k^B)}{r+s} = \left(r+s+(1-\beta)\eta(Q^B)+\beta\mu(Q^B)\right)k^B \tag{13.5}$$

The equations (13.4) and (13.5) characterize the equilibrium, and can be directly compared to the conditions (13.1) and (13.2) for the efficient allocation.

First, compare k^S to k^B: we can see that for all $\beta > 0$, $k^B < k^S$. In other words, there will be underinvestment as long as workers have ex post bargaining power. This is a form of **holdup**, in the sense that the firm makes an investment and the returns from the investments are shared between the worker and the firm. Because the investment is made before there is a match, there is no feasible way of contracting between the worker and the firm in order to avoid this holdup problem.

Thus the only way of obtaining efficiency is to set $\beta = 0$.

What about Q^S versus Q^B?

To compare Q^S versus Q^B, let $f(k^B) = f(k^S)$, then we obtain

$$\beta = \beta^*(Q) \equiv \frac{\eta'(Q)Q}{\eta(Q)} \equiv 1 + \frac{\mu'(Q)Q}{\mu(Q)},$$

is necessary and sufficient for $Q^S = Q^B$.

In other words, with $f(k^B) = f(k^S)$, we are back to the model without investment, so all we need is the Hosios condition for efficiency.

$$M = \mu \cdot U \Longrightarrow M_U = \mu' Q + \mu,$$
$$\Longrightarrow \frac{M_U U}{M} = 1 + \frac{\mu' Q}{\mu},$$

which can be verified as the Hosios condition in this case.

Thus when $f(k^B) = f(k^S)$, the Hosios condition is necessary and sufficient for efficiency.

This is not surprising, since with $f(k^B) = f(k^S)$, the economy is identical to the one with fixed capital.

The key question is whether it is possible to ensure both $f(k^B) = f(k^S)$ and $Q^S = Q^B$ simultaneously.

Of course, from the analysis the answer is no.
If $\beta > 0$, hold-up problem and $k^S > k^B$
If $\beta = 0$, the excessive entry of firms $Q^B < Q^S$.

THEOREM 13.1. *Constrained efficiency is impossible with ex ante investments and ex post bargaining.*

The intuition is quite straightforward: as long as $\beta > 0$, there is rent sharing on the marginal increase in productivity, thus hold-up. But $\beta = 0$ is inconsistent with optimal entry.

2. The Basic Model of Directed Search

Workers do not randomly search among all possible jobs, but apply for jobs that are more likely to be appropriate for their skills and interests. How do we model this? And how does this changed the positive and normative implications of search models?

One way is to construct the general equilibrium model with a non-degenerate wage distribution and then allow workers to search, perhaps in a smart way, among these jobs.. These models have the potential of leading to a coherent general equilibrium model with sequential search. But they are rather difficult to work with.

However, when all workers are assumed to observe all possible wage offers and can direct their search to one of these potential offers, then these models become quite tractable. At some level, this modeling assumption removes the actual "search" problem, but something akin to this, the coordination problem among the application decisions of workers is present in place the same role.

These models are sometimes referred to *competitive search* models, but is more useful to emphasize the two underlying assumptions: wage posting and directed search, so we will refer to them as *directed search* models.

To bring out the most important points, let us start from the economic environment of the search and investment model. Recall that in this model there are ex ante investments by firms, and bilateral search to form productive partnerships. In particular, recall that production requires 1 firm - 1 worker, with access to the production function $f(k)$, where k is capital for worker chosen before the matching stage by the firm. Recall that

$$f' > 0, \quad f'' < 0$$

The rate of time preference is r, and the rate of separation due to the destruction of capital by s.

We will now think of search frictions as equivalent to "coordination frictions". In particular, if there are an average of q workers per vacancy of a certain type then the flow rate of match for workers is $\mu(q)$, which is assumed to be continuously differentiable with $\mu' < 0$. Similarly, the flow rate of matching for a vacancy is $\eta(q) \equiv q\mu(q)$, where I am purposefully using the notation little q to distinguish this from the capital Q before which referred to the economy-wide queue length, whereas q it's specific to a type of job.

So this might seem somewhat strange; workers know what the various wages are, but conditional on applying to a job they may not get it; but this is sensible when there is no (centralized) coordination in the economy, because too many other people may be applying specifically to that job. The urn ball technology captured is in a very specific way, and in particular, we had

$$\eta(q) = 1 - \exp(-q) \text{ and } \mu(q) = \frac{1 - \exp(-q)}{q}$$

The technology here generalizes that.

As explained above, first all firms post wages w and also choose their capital k.

Workers observe *all* wages and then choose which job to seek. (they do not care about capital stocks).

Now more specifically let $q(w)$ be the ratio of workers seeking wage w to firms offering w. then $\mu(q(w))$ is flow rate of workers getting a job with wage w and $\eta(q(w))$ is flow rate of firms filling their jobs.

What equilibrium concept should we use here? Thinking about it intuitively, it is clear that we should ensure that workers apply to jobs that maximize utility and anticipate queue lengths at various wages rationally. This is straightforward.

The harder part is for firms. Firms should choose wages and investment to maximize profits, anticipating queue lengths at wages not offered in equilibrium. The last part is very important and corresponds to *Subgame perfection.* This is obviously important, since we have a dynamic economy, and you can see what will go wrong if we didn't impose subgame perfection.

Before we go further, let us first write the Bellman Equations, which are intuitive and standard for the firm (again imposing steady state throughout):

$$\begin{aligned} rJ^V(w,k) &= \eta(q(w))(J^F(w,k) - J^V(w,k)) - sJ^V(w,k) \\ rJ^F(w,k) &= f(k) - w - sJ^F(w,k) \end{aligned}$$

implying a simple equation for the value of firm

$$J^V(w,k) = \frac{\eta(f(k) - w)}{(r+s)(r+s+\eta)}$$

which we will use below.

The value of an employed worker is also simple:

$$rJ^E(w) = w + s(J^U - J^E(w))$$

What is slightly more involved is the value for unemployed worker.

Recall that unemployed workers take an important action: they decide which job to seek. Let $J^U(w)$ be the value of an unemployed worker when seeking wage w.

$$\underset{\text{utility of applying to wage } w}{rJ^U(w)} = \mu(q(w))[J^E(w) - \underset{\substack{\text{maximal utility} \\ \text{of unemployment}}}{J^U}]$$

where I have suppressed unemployment benefits without loss of any generality.

So what is J^U? Clearly:

$$J^U = \max_{w \in \mathcal{W}} J^U(w)$$

where $\mathcal{W}$ is the support of the equilibrium wage distribution.

Now this already builds in the requirement that w maximizes $J^U(w)$.

Also it is clear that w, k should maximize $J^V(w,k)$.

But what are the $q(w)$'s?

If we did not impose subgame perfection, then we could have crazy $q(w)$'s. Instead, firms would have to anticipate what workers would do if they deviate and create a new wage distribution.

So off-the-equilibrium path $q(w)$ should satisfy

$$\mu(q(w))\left[J^E(w) - J^U\right] = rJ^U$$

or if $J^E(w) - J^U < rJ^U$, then $q(w) = 0$.

To define an equilibrium more formally, let an allocation be a tuple $\left\langle \mathcal{W}, Q, K, J^U \right\rangle$, where $\mathcal{W}$ is the support of the wage distribution, $Q : \mathcal{W} \to \mathbb{R}$ is a queue length function, $K : \mathcal{W} \rightrightarrows \mathbb{R}$ is a capital choice correspondence, and $J^U \in \mathbb{R}$ is the equilibrium utility of unemployed workers.

Definition 13.1. *A directed search equilibrium satisfies*

1. *For all $w \in \mathcal{W}$ and $k \in K(w)$, $J^V(w,k) = 0$.*
2. *For all k and for all w, $J^V(w,k) \leq 0$.*
3. $J^U = \sup_{w \in \mathcal{W}} J^U(w)$.
4. *$Q(w)$ s.t. $\forall w$, $J^U \geq J^U(w)$, and $Q(w) \geq 0$, with complementary slackness.*

In words, the first condition requires firms to makes zero profits when they choose equilibrium wages and corresponding capital stocks. The second requires that for all other capital stock and wage combinations, profits are nonpositive. The third condition defines J^U as the maximal utility that an unemployed worker can get. The fourth condition is the most important one. It defines queue lengths to be such that workers are indifferent between applying to available jobs, or if they cannot be made indifferent, nobody applies to a particular job (thus the *complementary slackness* part is very important). This builds in the notion of *subgame perfection.*

Now we have

THEOREM 13.2. ***(Acemoglu and Shimer)*** *Equilibrium k, w, q maximize $\frac{\mu(q)w}{r+s+\mu}$ $(= rJ^U)$ subject to $\eta(q)\frac{(f(k)-w)}{r+s+\eta(q)} = (r+s)k$. And conversely, any solution to this maximization problem can be supported as an equilibrium.*

Basically what this theorem says is that the equilibrium will be such that the utility of an unemployed worker is maximized subject to zero profit.

PROOF. (sketch) Suppose not. Take k', w', q' which fails to maximize the above program. Then another firm can offer k'', w'' where (k^*, w^*, q^*) is the solution and $w'' = w^* - \varepsilon$. For ε small enough workers prefer k'', w'' to k', w', so $q'' > q^*$, which implies that k'', w'' makes positive profits, proving that (k', w', q) can't be an equilibrium. □

This theorem is very useful because it tells us that all we have to do is to solve the program:

$$\max \frac{\mu(q)w}{r+s+\mu(q)}$$
$$\text{s.t.} \qquad \frac{\eta(q)(f(k)-w)}{r+s+\eta(q)} = (r+s)k$$

Is this a convex problem?

No, but let's assume differentiablity (which we have so far), then first order conditions are necessary.

Forming the Lagrangian with multiplier λ

$$\frac{\eta(q)f'(k)}{r+s+\eta(q)} = r+s \tag{13.6}$$

$$\frac{\mu(q)}{r+s+\mu(q)} - \frac{\lambda\eta(q)}{r+s+\eta(q)} = 0 \tag{13.7}$$

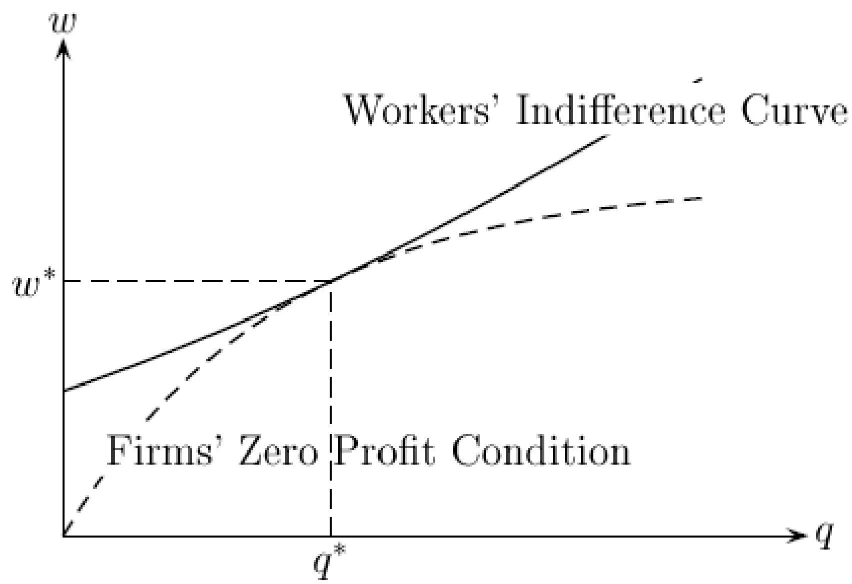


FIGURE 13.1

and

$$\frac{(r+s)\mu'(q)}{(r+s+\mu(q))^2} + \lambda\left(\frac{(r+s)\eta'(q)(f(k)-w)}{(r+s+\eta(q))^2}\right) = 0 \tag{13.8}$$

Now (13.6) is identical to (13.1) above, which was

$$\frac{Q^S\mu(Q^S)f'(k^S)}{r+s+Q^S\mu(Q^S)} = r+s$$

implies that, denoting the capital labor ratio in the wage posting equilibrium by k^{wp},

$$k^{wp} = k^S$$

Therefore, with wage posting, capital investments are always efficient.

Why is this? You might think this is because there is no more holdup problem, and this is essentially true, but the intuition is a bit more subtle. In fact, there is something like hold-up because firms that invest more in equilibrium prefer to pay higher wages, but despite this the efficient level of investment results. The reason is that the higher wages that they pay is exactly offset with the higher probability that they will attract workers, so net returns are not subject to hold-up.

Next we have

$$\lambda = \frac{r+s+\eta(q)}{(r+s+\mu(q))q}$$

and substitute this into (iii), and used at zero profit constraints to solve for

$$w = f(k) - \frac{(r+s)(r+s+\eta(q))}{\eta(q)}k$$

Then we have:

$$\eta' \frac{q^2 f(k)}{r+s} + \left[r+s+\mu+\mu' q - q^2\mu'\right] k = 0$$

which is identical to (13.2). We have therefore established:

THEOREM 13.3. *The directed search equilibrium of the search and investment model is constrained efficient.*

Therefore, the equilibrium is constrained efficient! (note uniqueness is not guaranteed, but neither was it in the social optimum)

Thus, wage posting decentralizes the efficient allocation as the unique equilibrium.

How can we understand this efficiency better?

Acemoglu-Shimer consider a number of different economies

(1) Wage posting but no directed search. Clearly, in this case things are very bad, and we get the Diamond paradox.
(2) An economy where firms choose their own capital level, and then "post a bargaining parameter β" and upon matching, the firm and the worker Nash bargain with this parameter. It can be shown that if there is no capital choice, this economy will lead to an equilibrium in which all firms post the Hosios β, and constrained efficiency is achieved. But if there is a capital choice, and the only thing workers observe are the posted β's, then in equilibrium all firms offer the Hosios β, but there is under investment because of the hold-up problem.

(3) An economy where firms choose their own capital level and workers apply to firms observing these capital levels, and then they bargain according to some exogenously given parameter β. In this case, the equilibrium is inefficient and may have under or overinvestment. If the value of β is at the Hosios value, then the equilibrium will be constrained efficienct.
(4) An economy where firms choose their own capital level and post β, and workers observe both k and β, then always constrained efficiency.

So what do we learn? What is important is directed search, and especially the ability to direct search towards higher capital intensity firms. With wage posting, those are the high-wage firms, hence the objective is achieved. But the same outcome is also obtained if β is at the Hosios level, and workers observe capital levels.

Next, one might wonder whether an economy in which workers know/observe all of the wages offered in equilibrium is too extreme (especially given our motivation of doing away with a Walrasian auctioneer). A more plausible economy may be one where workers observe a finite number of wages.

Interestingly, we do not need all workers to observe all the wages as the model with a non-degenerate wage distribution in the last lecture illustrated.

THEOREM 13.4. *Suppose each worker observes (can apply to) at least two of the firms among the continuum of active firms, then the efficient allocation is an equilibrium of the search and investment model with directed search and wage posting.*

PROOF. (sketch) Suppose all firms are offering (q^{wp}, w^{wp}, k^{wp}). Now consider a deviation to some other (w', k'). Any worker who observes (w', k') has also observed another firm offering (w^{wp}, k^{wp}). Since (w^{wp}, k^{wp}) maximizes worker utility, he will apply to this in preference of

$$(w', k') \Longrightarrow q(w') = 0.$$

Consequently, all firms will be happy to offer (w^{wp}, k^{wp}) and they will each be tracked the queue length of q^{wp}. $\square$

What is the intuition? *Effectively Bertrand Competition.* Each firm knows that it will effectively be competing with another firm offering the best possible deal to the worker, even though differently from the standard Bertrand model, it does not know which particular firm this will be. Nevertheless, the Bertrand reasoning forces each firm to go to the allocation that is best for the workers.

Note that this theorem is not stated as an "if and only if" theorem. In particular, when each worker only observes two wages, there can be other "non-efficient" equilibria. In particular, it can be proved that: *When each worker observes two wages, there can exist non-efficient equilibria.* This last theorem not withstanding, the conclusion of this analysis is that relatively little information is required for wage posting to decentralize the efficient allocation.

3. Risk Aversion in Search Equilibrium

The tools we developed so far can also be used to analyze general equilibrium search with risk aversion. Let us focus on the one-period model with wage posting. This can again be extended to the dynamic version, but explicit form solutions are possible only under constant absolute risk aversion (see Acemoglu-Shimer, JPE 1999)

Measure 1 workers; and they all have utility $u(c)$ where the consumption of individual i is

$$C_i = A_i + y_i - \tau_i$$

where A_i is the non-labor income of individual, y_i is his labor income, equal to the wage w that he applies it obtains if he's employed, and equal to the unemployment benefit z when unemployed. Finally, τ_i is equal to the taxes paid by this individual. u is increasing, concave and differentiable.

Let us start with a homogeneous economy where $A_i = A_0$ and $\tau_i = \tau$ for all i.

We also assume that firms are risk-neutral, which is not chill for example because workers may hold a balanced mutual fund. I will onlypresent the analysis for the static economy here.

Timing of events:

- Firms decide to enter, buy capital $k > 0$ (as before irreversible,) and post a wage w
- Workers observe all wage offers and decide which wage to seek (apply to).

As before, if on average there are q times as many workers seeking wage w as firms offering w, then workers get a job with prob. $\mu(q)$.

Firms fill their vacancies with prob. $\eta(q) \equiv q\mu(q)$, with our standard assumptions, $\mu'(q) < 0$ and $\eta'(q) > 0$

As before, let an allocation be $\langle \mathcal{W}, Q, K, U \rangle$, where $\mathcal{W}$ is the support of the wage distribution, $Q : \mathcal{W} \to \mathbb{R}$ is a queue length function, $K : \mathcal{W} \rightrightarrows \mathbb{R}$ is a capital choice correspondence, and $U \in \mathbb{R}$ is the equilibrium utility of unemployed workers.

DEFINITION 13.2. *An allocation is an equilibrium iff*

(1) $\forall w \in \mathcal{W}$ *and* $k \in K(w)$, $\eta(Q(w))(f(k) - w) - k = 0$.
(2) $\forall w, k$, $\eta(Q(w))(f(k) - w) - k \leq 0$.
(3) $U = \sup_{w \in \mathcal{W}} \mu(Q(w))u(A + w) + (1 - \mu(Q(w))u(A + z)$
(4) $Q(w)$ *s.t.* $\forall w$, $U \geq \mu(Q(w))u(A+w)+(1-\mu(Q(w)))u(A+z)$ *and* $Q(w) \geq 0$, *with complementary slackness.*

- $\Longrightarrow$ As before type of subgame perfection on beliefs about queue lengths after a deviation.

Characterization of equilibrium is similar to before

THEOREM 13.5. $(\mathcal{W}, Q, K, U)$ *an equilibrium if and only if* $\forall\ w^* \in \mathcal{W}$, $q^* \in Q(w^*), k^* \in K(w^*)$

$$(w^*,\ q^*, k^*) \in \arg\max \mu(q)u(A + w) + (1 - \mu(q))u(A + z)$$

s.t.

$$\eta(q)(f(k) - w) \geq 0.$$

In words, every equilibrium maximizes worker utility subject to zero profits, as proved before in the context of the risk-neutral model.

The analysis is similar to before. Profit maximization implies an even simpler condition (because the environment is static)

$$\eta(q^*)f'(k^*) = 1$$

Zero profits gives

$$\eta(q^*)(f(k^*) - w^*) = k^*$$

Now combining these two:

$$w^* = f(k^*) - k^* f'(k^*),$$

which you will notice is exactly the neoclassical wages equal to marginal product condition. Why is that?

Finally, combining this with, $\eta(q^*)f'(k^*) = 1$, we can derive a relation in the (q, w) space which corresponds to the zero-profits and profit maximization constraints that an equilibrium has to satisfy.

An equilibrium is then a tangency point between the indifference curves of homogeneous workers and this profit-maximization constraint, as we had in the risk-neutral model of Acemoglu-Shimer (IER, 1999):

The equilibrium can be depicted and analyzed diagrammatically.

Notice again that uniqueness **not** guaranteed.

What makes this attractive is that comparative statics can also be done in a simple way, exploiting "revealed preference" or single crossing.

For example, we have a change such that all workers become more risk-averse, i.e., and the utility function becomes more concave, what happens to equilibrium?

We can show that as risk-aversion increases, then we have $w \downarrow, q \downarrow, k \downarrow$.

Why? Indifference curves become everywhere steeper, the causing the tangency point to shift to the left. Unambiguous despite the fact that equilibrium may not be unique.

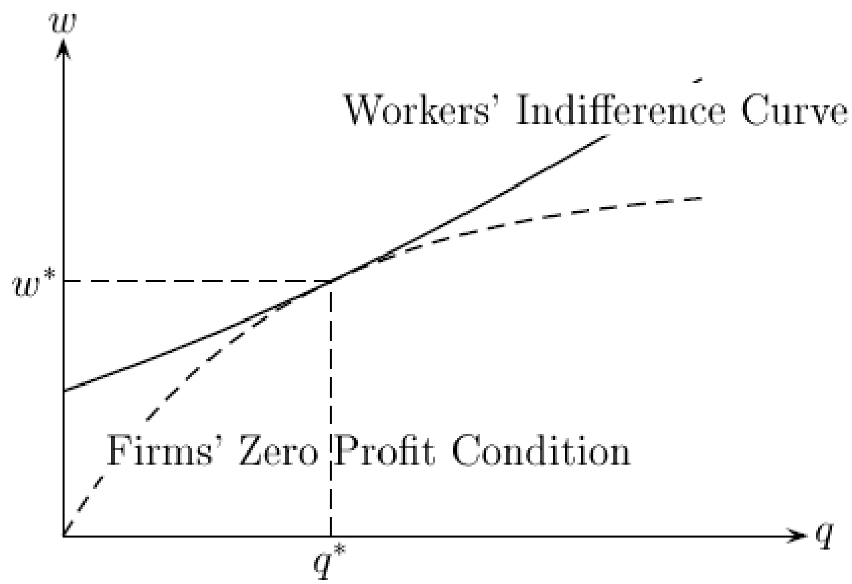


Figure 13.2

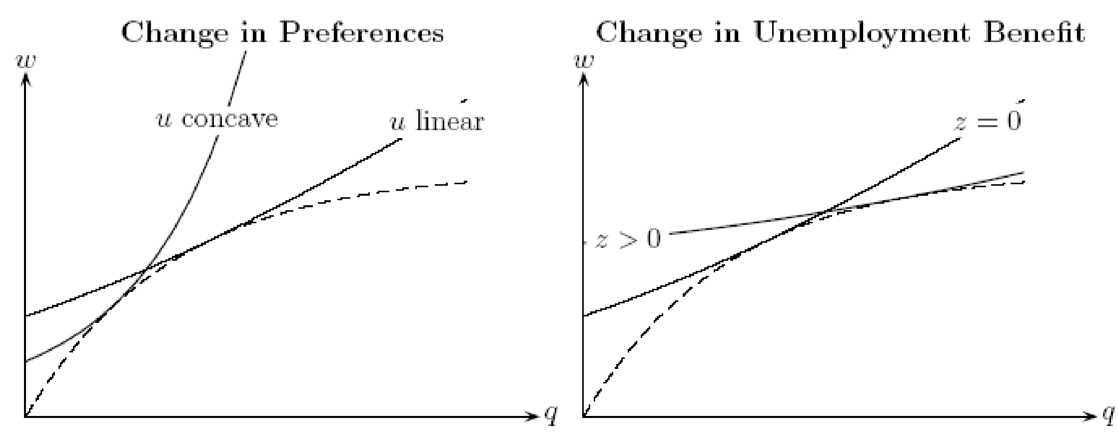


Figure 13.3

Essentially, comparative static result unambiguous because u_1-curve single-crosses u_2-curve.

Intuition: "Market Insurance." Workers are more risk-averse, so firms offer insurance by creating low-wage but easier to get jobs. Capital falls because once jobs are easier to get for workers, vacancies remain open for longer (with higher probability), so capital is unused for longer, reducing investment. Summarizing this:

THEOREM 13.6. *Consider a change from utility function u_1 to u_2 where u_2 is a strictly concave transformation of u_1. Then if (k_1, w_1, q_1) is any equilibrium with preferences u_1 and (k_2, w_2, q_2) is any equilibrium with preferences u_2, then $k_2 < k_1$, $w_2 < w_1$ and $q_2 < q_1$.*

Similarly, what happens when the unemployment benefits z increases from z_1 to z_2?

THEOREM 13.7. *Consider a change from unemployment benefits z_1 to $z_2 > z_1$. Then if (k_1, w_1, q_1) is any equilibrium with benefits z_1 and (k_2, w_2, q_2) is any equilibrium with benefits z_2, then $k_2 > k_1$, $w_2 > w_1$ and $q_2 > q_1$.*

PROOF. (sketch) By revealed preference

$$\begin{aligned} \mu(q_1)(u(A+w_1) - u(A+z_1)) &\geq \mu(q_2)(u(A+w_2) - u(A+z_1)) \\ \mu(q_2)(u(A+w_2) - u(A+z_2)) &\geq \mu(q_1)(u(A+w_1) - u(A+z_2) \end{aligned}$$

Multiply through and simplify

$$\begin{aligned} & (u(A+z_1) - u(A+z_2))(u(A+w_1) - u(A+w_2)) \geq 0 \\ \Longrightarrow \quad & z_1 \leq z_2 \Longleftrightarrow w_1 \leq w_2. \end{aligned}$$

All inequalities strict since all curves smooth. □

What happens when there is heterogeneity?

Suppose that there are $s = 1, 2, \ldots, S$ types of workers, where type s has utility function u_s, after-tax asset level A_s, and unemployment benefit z_s. Let U now be a vector in $\mathbb{R}^S$, and assume, for simplicity. Then:

THEOREM 13.8. *There always exists an equilibrium. If* $\{\mathcal{K}, \mathcal{W}, Q, U\}$ *is an equilibrium, then any* $k_s^* \in \mathcal{K}$,$w_s^* \in \mathcal{W}$, *and* $q_s^* = Q(w_s^*)$, *solves*

$$U_s = \max_{k,w,q} \mu(q)u_s(A_s + w) + (1 - \mu(q))\, u(A_s + z_s)$$

subject to $\eta(q)(f(k) - w) - k = 0$ *for some* $s = 1, 2, .., S$. *If* $\{k_s^*, w_s^*, q_s^*\}$ *solves the above program for some* s, *then there exists an equilibrium* $\{\mathcal{K}, \mathcal{W}, Q, U\}$ *such that* $k_s^* \in \mathcal{K}$, $w_s^* \in \mathcal{W}$, *and* $q_s^* = Q(w_s^*)$.

The important result here is that any triple $\{k_s^*, w_s^*, q_s^*\}$ that is part of an equilibrium maximizes the utility of one group of workers, subject to firms making zero profits. The market *endogenously* segments into S different submarkets, each catering to the preferences of one type of worker, and receiving applications only from that type.

The efficiency and output-maximization implications of this model are also interesting. First, supposed that $u(\cdot)$ is linear. Then $z = \tau = 0$ maximizes output. In particular, we have

THEOREM 13.9. *Suppose that* u *is linear, then* $z = \tau = 0$ *maximizes output.*

PROOF. (sketch) The equilibrium solves $\max \mu(q)w$ subject to $q\mu(q)(f(k)-w) = k$. Substituting for w we obtain:

$$\mu(q)f(k) - k/q \equiv y(k, q),$$

which is net output, thus is maximized by equilibrium choices. □

But an immediate corollary is that if $u(\cdot)$ is strictly concave, than the equilbrium with $z = \tau = 0$ does *not* maximize output.

THEOREM 13.10. *Suppose that* u *is strictly concave, then* $z = \tau = 0$ *does not maximize output.*

This is an immediate corollary of the previous theorems.

THEOREM 13.11. *Let* u *be an arbitrary concave utility function,* q^e *be the output-maximizing level of queue length and let*

$$z^e \equiv \frac{u(A_0 - \tau^e + w^e) - u(A_0 - \tau^e + z^e)}{u'(A_0 - \tau^e + w^e)}$$

and the balanced-budget condition

$$\tau^e = (1 - \mu(q^e))z^e$$

then the economy with unemployment benefit z^e *achieves an equilibrium with* q^e *and the maximum output.*

The following figure gives the intuition:

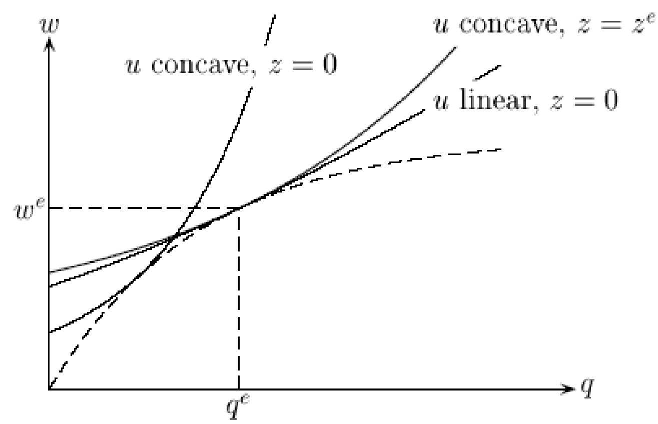


FIGURE 13.4

But this is not "optimal," since when workers are risk averse, maximizing output is not necessarily the right objective. Optimal unemployment benefits, z^o, should maximize ex ante utility. Interestingly, this could be greater or less than the efficient level of unemployment benefits, z^e, which maximizes output. What is the intuition for this?