Lecture 3
Shapiro-Stiglitz Model of Efficiency Wages

Leszek Wincenciak, Ph.D.

University of Warsaw

Lecture outline:

Introduction

The model set-up
- Workers
- The effort decision of a worker
- Values of E, U and S
- No-Shirking Condition (NSC)
- Employers
- Market equilibrium

Implications of the model

Simple comparative statics

Alternative methods for the enforcement of discipline
- Performance bonds
- Other costs of dismissal
- Heterogenous workers

Introduction

- Why do not wages fall to clear the market?
  - Informational structure of employer-employee relationship – imperfect monitoring of workers effort on the job
  - Principal-agent problem
  - Higher wages and a pool of unemployment offer incentives for the workers to exert effort
  - This is a central idea of the Shapiro-Stiglitz model (*Equilibrium Unemployment as a Worker Discipline Device*, American Economic Review, Vol. 74, No. 3, 1984)

- Consider a situation where all workers can receive the market wage and there is no unemployment
- In this case, the worst thing that can happen to a worker is that he will be fired and instantaneously rehired
- There is therefore no penalty for not exerting effort ('shirking')
- To induce workers not to shirk, firms pay above-market wages. Therefore, job loss imposes a penalty
- But if one firm pays above-market wages, then presumably they all will
- In this case, the incentive not to shirk disappears, but:
  - Unemployment results since wages are above the natural equilibrium level
  - Unemployment creates its own penalty for shirking

- Hence, the model implies that unemployment and monitoring are substitutes
- Consequently, wages serve two functions: allocating labour and providing incentives for employee effort conditional on employment. As is usually the case when one instrument is used to solve two problems, this is likely to lead to inefficient outcomes

The model set-up

Workers

Assume there are $\overline{L}$ (total labour supply is fixed) identical workers, all of whom dislike putting forth effort, but enjoy consuming goods. The workers lifetime utility is given by:

$$U = \int_{t=0}^{\infty} e^{-\rho t} u(t) dt, \quad \rho > 0, \tag{1}$$

where $u(t)$ is instantaneous utility at time t, and ρ is the discount rate. The instantaneous utility is defined as:

$$u(t) = \begin{cases} w(t) - e(t) & \text{if employed} \\ 0 & \text{if unemployed.} \end{cases} \tag{2}$$

Wages are denoted by w and e denotes workers effort. There are only two possible values for e: workers may choose to shirk, then $e = 0$, or to provide some fixed positive level of effort, $\overline{e} > 0$.

At any moment in time, a worker may be in one of three states:

- employed and not shirking (E)
- employed and shirking (S)
- unemployed (U)

Assume that with probability b per unit of time, jobs naturally end, due to reallocation, etc. If worker begins to work at time t_0, the probability that he is still working at time t is:

$$P(t) = e^{-b(t-t_0)}, \quad b > 0. \tag{3}$$

Equation (3) states that $P(t+\tau)/P(t) = e^{-b\tau}$ which is independent of t. It implies that it doesn't matter for how long the worker worked on the job. This follows from the properties of Poisson process ('no memory'), which simplifies the analysis greatly.

Random variable 'losing job in time t' has the following density function:

$$f(t) = b \cdot e^{-bt}.$$

Distribution of this random variable – i. e. the probability that losing a job occurred before time t is:

$$F(t) = \int_0^t f(y)dy = 1 - e^{-bt},$$

thus the probability that given person is still in job in time t is:

$$1 - F(t) = e^{-bt}.$$

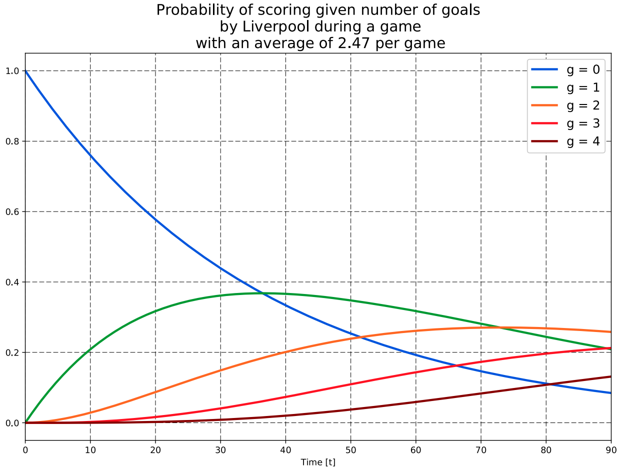
Parameter b is the *hazard rate* of losing a job, which is a conditional probability of losing a job in time t provided that the person worked until the moment t.

$$h(t) \equiv \lim_{\Delta t \to 0} \frac{Pr(t + \Delta t > T > t \mid T > t)}{\Delta t} = \frac{f(t)}{1 - F(t)}.$$

$$\frac{f(t)}{1 - F(t)} = \frac{be^{-bt}}{e^{-bt}} = b.$$

Assuming Poisson process (exponential distribution) means we have a constant hazard rate – independent of time.

Example


Probability of scoring given number of goals
by Liverpool during a game
with an average of 2.47 per game
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
Time [t]
g = 0
g = 1
g = 2
g = 3
g = 4

The effort decision of a worker

The only choice workers make is the selection of an effort level, which is discrete by assumption. If a worker chooses to exert some positive level of effort ($\overline{e}$), he receives the wage (w) and retains the job, until exogenous factors cause a separation to occur (with probability b per unit of time).

If a worker decides to shirk, there is some probability q per unit of time, that he will be caught. We assume that the firms' detection of shirkers also follows Poisson process. The probability that a shirking worker is still employed at time τ later is equal to $e^{-q\tau}$ (prob. that he was not caught shirking) times $e^{-b\tau}$ (prob. that the job did not end naturally).

Workers who are caught shirking are fired and enter the unemployment pool. The probability per unit of time of acquiring new job while in the unemployment pool (the acquisition rate) is a, which is taken by all workers as given. However this transition rate is determined endogenously in the economy as a whole. Firms choose workers at random out of the pool of unemployed workers. Thus a is determined by the rate at which firms are hiring (which is determined by the number of employed workers and the rate at which jobs end) and the number of unemployed workers. Because workers are identical, the probability of finding new job does not depend on how workers became unemployed or how long they are unemployed.

Being fired carries no stigma – the next potential employer knows that the worker is not more immoral than any other worker. He knows that the previous firm must have paid sufficiently low wage that it paid for the worker to shirk.

Values of E, U and S

The worker selects an effort level in order to maximize his discounted utility stream. This involves comparison of the utility from shirking with utility from not shirking.

- V_i – value of being in state i $(i = E, S, U)$
- V_i is a discounted lifetime utility from present moment forward of a worker who is in state i
- Poisson transition processes imply that V_i does not depend on how long the worker has been in current state nor on his prior history
- Focusing on steady-states implies that V_i's are constant
- Use of dynamic programming (Bellman equations)

The central idea of dynamic programming is to look only at a brief interval of time and use V_i themselves to summarize what occurs after the end of the brief interval. In order to find V_E, V_S and V_U it is not then necessary to analyze various paths the worker may follow over infinite time horizon.

Consider a worker who is employed and exerts effort at time $t = 0$. Suppose time is divided into intervals of the length Δt. Then:

$$V_E(\Delta t) = \int_{t=0}^{\Delta t} e^{-bt} e^{-\rho t}(w - \overline{e})dt + e^{-\rho \Delta t}\left[e^{-b\Delta t}V_E(\Delta t) + (1 - e^{-b\Delta t})V_U(\Delta t)\right]. \quad (4)$$

If we compute the integral, we can simplify equation (4) to:

$$\begin{aligned} V_E(\Delta t) =& \frac{1}{\rho + b}(1 - e^{-(\rho+b)\Delta t})(w - \overline{e}) \\ &+ e^{-\rho\Delta t}\left[e^{-b\Delta t}V_E(\Delta t) + (1 - e^{-b\Delta t})V_U(\Delta t)\right]. \end{aligned} \tag{5}$$

Then solving for $V_E(\Delta t)$ gives:

$$V_E(\Delta t) = \frac{1}{\rho + b}(w - \overline{e}) + \frac{1}{1 - e^{-(\rho+b)\Delta t}}e^{-\rho\Delta t}(1 - e^{-b\Delta t})V_U(\Delta t) \tag{6}$$

Now we can use the fact that:

$$\lim_{\Delta t \to 0} V_E(\Delta t) = V_E$$
$$\lim_{\Delta t \to 0} V_U(\Delta t) = V_U$$

Applying de l'Hospital rule to (6) we get:

$$V_E = \frac{1}{\rho + b}\left[(w - \overline{e}) + bV_U\right]. \quad (7)$$

Equation (7) can be also derived from so called Bellman equation. Think of an asset which pays dividend at rate $w - \overline{e}$ per unit of time when a worker is employed and no dividends when he is unemployed. The asset is being priced by risk-neutral investors with required rate of return of ρ. Since the expected present value of lifetime dividends of this asset is the same as worker's expected present value of lifetime utility, the asset's price must be V_E when the worker is employed and V_U when unemployed. For the asset to be held, it must provide an expected rate of return equal ρ. Its dividends per unit of time plus any expected capital gains or losses per unit of time must be equal ρV_E. When the worker is employed, the dividends per unit of time are $w - \overline{e}$, and there is a probability b per unit of time of a capital loss of $V_E - V_U$. Thus:

For the worker who is employed and not shirking:

$$\rho V_E = (w - \overline{e}) + b(V_U - V_E), \tag{8}$$

for the worker who is employed and shirking:

$$\rho V_S = w + (b + q)(V_U - V_S), \tag{9}$$

and if worker is unemployed:

$$\rho V_U = a(V_N - V_U), \tag{10}$$

where $V_N = \max\{V_E, V_S\}$.

No-Shirking Condition (NSC)

No-Shirking Condition (NSC)

The firm must pay high enough, that workers choose not to shirk, so that $V_E \geq V_S$. This is called the no-shirking condition (NSC). Solving equations (8) and (9) for V_E and V_S yields:

$$V_E = \frac{(w - \overline{e}) + bV_U}{\rho + b}. \tag{11}$$

$$V_S = \frac{w + (b + q)V_U}{\rho + b + q}. \tag{12}$$

Since NSC requires that $V_E \geqslant V_S$, it implies:

$$(w - \overline{e}) + b(V_U - V_E) \geqslant w + (b + q)(V_U - V_E), \tag{13}$$

which is equivalent to:

$$V_E - V_U \geqslant \frac{\overline{e}}{q}. \tag{14}$$

Equation (14) implies that firms set wages high enough that workers strictly prefer employment to unemployment. Thus workers receive rents. The size of the premium is increasing with the effort and decreasing in firms' efficiency in detecting shirkers, q. Unemployment benefits by raising V_U will require higher wages in equilibrium.
Using equations (8) and (10) we can find the wage, which is needed to induce desirable effort level. This efficiency wage ($\hat{w}$) is given by:

$$\hat{w} \geqslant \overline{e} + (a + b + \rho)\frac{\overline{e}}{q}. \tag{15}$$

$$\hat{w} \geqslant \overline{e} + (a + b + \rho)\frac{\overline{e}}{q}$$

The efficiency wage is:

- increasing in the cost (disutility) of effort, $\overline{e}$
- increasing in the ease of finding new job, a
- increasing in the rate of job exogenous breakup, b – If you are going to lose your job soon anyway, why not cheat?
- increasing in the discount rate, ρ (future is less important)
- decreasing in the probability of shirkers detection, q

It is more convenient to express the efficiency wage as a function of employment per firm, L, rather than in terms of a. In the steady state, the inflows and and outflows from unemployment balance. The number of workers becoming unemployed per unit of time is N (the number of firms) times L (the number of workers per firm) times b (job breakup rate). The number of unemployed workers finding new jobs is $\overline{L} - NL$ times a.

$$a = \frac{NLb}{\overline{L} - NL}. \tag{16}$$

Substituting this into (15) yields:

$$\hat{w} \geqslant \overline{e} + \left(\rho + \frac{\overline{L}}{\overline{L} - NL} b\right) \frac{\overline{e}}{q}. \tag{17}$$

$$\hat{w} \geqslant \overline{e} + \left(\rho + \frac{\overline{L}}{\overline{L} - NL} b\right) \frac{\overline{e}}{q}. \tag{17}$$

Expression (17) is the final no-shirking condition. It shows, as a function of employment, the wage which has to be paid in order to make incentives for the workers not to shirk. When more workers are employed, the pool of unemployment is smaller, and it is easier for the unemployed to find new job. At full employment, the unemployed (job leavers at rate b) find next work instantly, so there is no cost of being fired, and thus no wage can deter shirking. Also note that: $u = \frac{\overline{L} - NL}{\overline{L}}$, so we can write (17) as:

$$\hat{w} \geqslant \overline{e} + \left(\rho + \frac{b}{u}\right) \frac{\overline{e}}{q}.$$

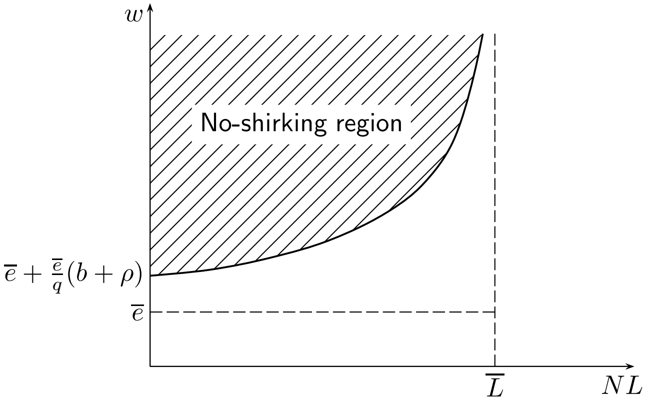


Figure 1. The aggregate no-shirking constraint

Employers

There are N identical firms. Each firms maximizes profits at time t:

$$\pi(t) = AF(L(t)) - w(t)[L(t) + S(t)], \quad F'(\cdot) > 0, \quad F''(\cdot) < 0, \tag{18}$$

where L is the number of workers who exert effort, S is the number of those who shirk. The firm's problem is to set wage high enough to prevent shirking, and then to choose L. The firm chooses w and L at each moment to maximize the instantaneous flow of profits.

$$AF'(\overline{L}/N) > \overline{e}. \tag{19}$$

This condition states that if there were perfect monitoring, there would be full employment.

Market equilibrium

Firms hire workers up to the point where the marginal product of labour equals the wage:

$$AF'(NL) = \hat{w} \tag{20}$$

The set of points that satisfy (20) is simply the aggregate labour demand curve.

The equilibrium wage and employment are now easy to identify. Each firm (small) taking a as given, finds that it must offer at least the wage equal $\hat{w}$. The firm's demand for labour then determines how many workers are hired at that wage. Equilibrium occurs where the aggregate demand for labour intersects the aggregate NSC.

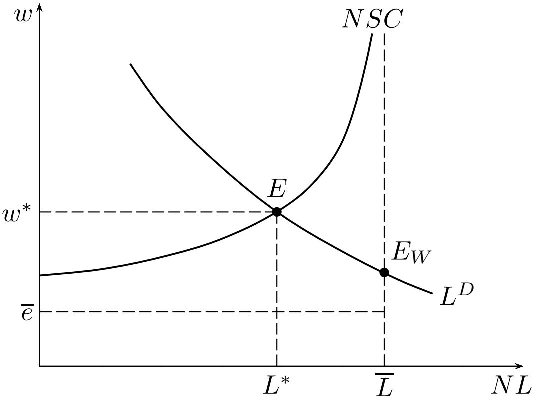


Figure 2. Equilibrium wage and employment in the Shapiro-Stiglitz model

Implications of the model

Implications of the model: welfare analysis

If firms are owned by workers and the ownership is equally distributed, then the central planning problem is to maximize the expected utility of workers subject to the NSC and the resource constraint:

$$\max_{w,NL} = (w - \overline{e})NL \tag{21}$$

$$\text{subject to: } w \geqslant \overline{e} + \left(\rho + \frac{\overline{L}}{\overline{L} - NL} b\right) \frac{\overline{e}}{q} \tag{NSC}$$

$$\text{subject to: } wNL \leqslant F(NL) \tag{Feasibility}$$

Implications of the model: welfare analysis

The set of points that satisfy constraints are shaded on the Figure 3. Social welfare is therefore attained at point A, where the line of average product $F(NL)/NL$ intersects with NSC line. Employment in social optimum is greater than in the equilibrium in the model.

Implications of the model: welfare analysis

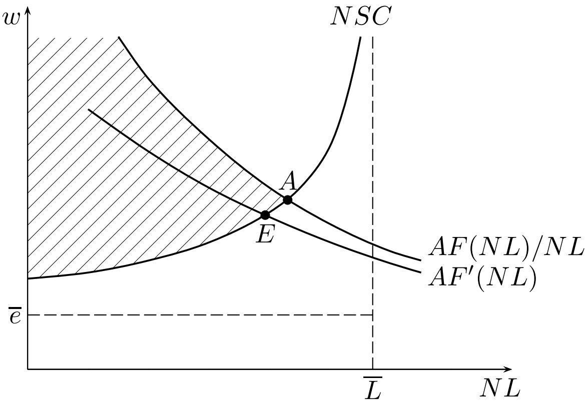


Figure 3. Social optimum

Implications of the model: welfare analysis

- Equilibrium level of employment is inefficient
- Each firm tends to employ too few workers, since it sees the private cost of hiring an additional worker as w, while social cost is only $\overline{e}$, which is lower
- On the other hand, when firm hires one more worker, it fails to take account of the effect that this has on V_U (by reducing the size of unemployment pool) and forces other firms to pay higher wages – this leads to negative externality (overemployment)
- But the first effect dominates and we have too high unemployment
- Wages could be subsidized using taxes on profits

Implications of the model

- There is unemployment in equilibrium
- Equilibrium wage w^* does not clear the market
- The unemployment is involuntary: all unemployed would prefer to work at prevailing wage or lower, but cannot make a credible promise not to shirk at such wage
- Thus wages do not fall and unemployment remains
- The unemployment results because of firms inability to monitor the activities of their employees costlessly and perfectly
- Source of inefficiency – costly monitoring

Implications of the model

Model implies downward wage rigidity

- Consider a negative aggregate productivity shock ($A \downarrow$):
  - In a Walrasian labour market L remains at $\overline{L}$ and wages fall
  - Here: L goes down and wages fall a little, but less than in classical case
  - If wage adjustment is costly this may lead to wage rigidity
- Now consider a positive aggregate productivity shock ($A \uparrow$):
  - In a Walrasian labour market L remains at $\overline{L}$ and wages rise
  - Here: L goes up and wages increase a little, but less than in classical case
  - Can there be a (menu) cost that prevents this wage adjustment?
  - No! All workers would start to shirk!

Simple comparative statics

Simple comparative statics

Consider an exogenous rise in q, that is in the probability per unit of time that a shirking worker is detected.

- NSC line shifts down – greater q means that the firm needs not to pay as high wage as before
- Labour demand line is not affected
- Wages in equilibrium fall and employment increases
- If $q \to \infty$ the probability that a shirking worker is caught in any finite length of time approaches 1. The non-shirking wage approaches $\overline{e}$ for any level of employment and we have Walrasian equilibrium with full employment

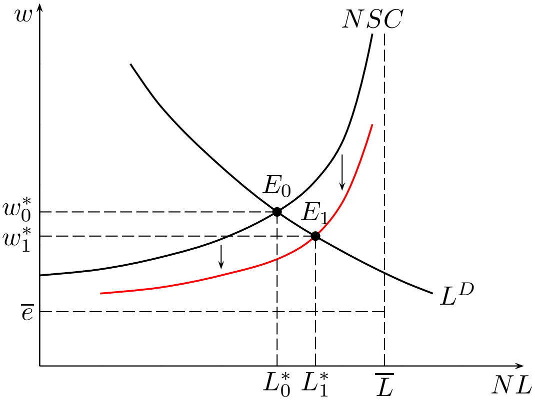


Figure 4. The effect of a rise in q in the Shapiro-Stiglitz model

Consider the case with no turnover, $b = 0$. In this case the unemployed workers are never hired, so the unemployment spell lasts forever. The punishment for shirking is then very serious! As a result, in equilibrium the no-shirking wage is independent of the level of employment.

Equation (17) with $b = 0$ reduces to:

$$w = \overline{e} + \rho \frac{\overline{e}}{q}. \tag{22}$$

Intuition: the gain from shirking relative to exerting effort is $\overline{e}$ per unit of time. The cost is that there is a probability q per unit of time of becoming permanently unemployed and thereby losing the discounted surplus from the job, which is $(w - \overline{e})/\rho$. Equating costs with benefits yields: $w = \overline{e} + \rho \frac{\overline{e}}{q}$.

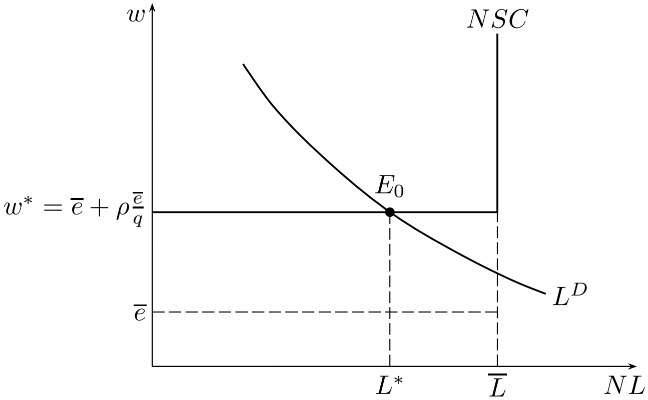


Figure 5. The Shapiro-Stiglitz model without turnover

Alternative methods for the enforcement of discipline

Alternative methods for the enforcement of discipline
Performance bonds

- In this model there is a particular mechanism of enforcement of discipline: workers who are caught shirking are fired, and equilibrium unemployment is sufficiently large that it serves as an effective deterrent to shirking
- Another method: workers post performance bonds
- This raises several problems:
  - Workers may not have wealth to post bond (if q is low the effective bond has to be large)
  - Firms face moral hazard problem (especially if effort is hardly observable) – they may claim that workers shirk in order to appropriate the bond
  - Firms reputation loss may solve the problem (but not in 100%)

Other costs of dismissal

- Unemployment in this model is a cost of dismissal
- If other costs of dismissal are sufficiently high, workers may have an incentive to exert effort even under full employment
- Example of such costs:
  - Search costs (search and matching theory)
  - Moving expenses
  - Loss of job-specific human capital
- When effort is continuous variable, the firm will see the effort as an increasing function of wages, so there would probably exist involuntary and frictional unemployment
- Unemployment will be higher for groups with lower job switching costs

Heterogenous workers

- What if workers are not homogenous?
- Then fired workers could carry stigma, which could serve as a discipline device even with full employment
- In practice firms offer wages which are dependent on worker's history
- Workers concern on their reputation depends on the cost of reputation loss
- Workers already labeled as below average in quality have less to lose from being labeled as 'shirkers'
- Even if reputation matters, equilibrium will entail some unemployment as a discipline device, at least for the lower-quality workers