DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

CAREER CONCERNS AND FIRM- SPONSORED GENERAL TRAINING

Christos Bilanakos

Discussion Paper 07-2011

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: +357-22893700, Fax: +357-22895028
Web site: http://www.econ.ucy.ac.cy

Career Concerns and Firm – Sponsored General Training

Christos Bilanakos*

Abstract

This paper studies the provision of firm-sponsored general training in the presence of workers' career concerns. We use a model building on the argument that the acquisition of general skills increases the worker's bargaining power vis-à-vis the employer. In this context, we show that the worker's implicit incentives to provide effort increase with the level of acquired general training. The employer takes this reciprocal effect into account and is thus more willing to invest in general human capital in the first place. When the positive effect of training on worker's incentives is strong enough, the equilibrium outcome may even involve overinvestment in general training. It is also shown that a sharper increase in worker's power associated with additional training may strengthen the employer's investment incentives and have beneficial effects on welfare.

Keywords: General Training, Career Concerns, Power.
JEL Classifications: D82, J24, J31.

1. Introduction

This paper studies the provision of firm-sponsored general training in the presence of workers' career concerns[1]. The exploration of this issue is largely motivated by the contrast between the theoretical recognition of efficiency benefits associated with the acquisition of general human capital and the ambiguous empirical evidence related to firm-sponsored investment in general skills. Indeed, there is much recent work arguing in favor of extensive training that should be directed towards general (rather than firm-specific) capacities which improve organizational performance by enabling workers to

* Department of Economics, University of Cyprus, P.O. Box 20537, Nicosia 1678, Cyprus. E-mail: bilanakos.christos@ucy.ac.cy . Telephone: + 35722893692.

[1] General training refers to human capital investment that increases the worker's productivity to the same extent when the latter stays with her current employer as when she moves to another firm (or even forms her own firm).

engage in innovative or entrepreneurial activities[2]. However, evidence reveals a rather contradictory actual pattern of skill development. In fact, there seems to be a process of polarization under way: Firms in countries like the United States or the United Kingdom tend to underinvest in employees' general training relative, for example, to Japanese firms[3]. We use a model of general training incorporating workers' career concerns to provide an explanation for these differences.

In his seminal work on the provision of general human capital, Becker (1964) assumes a perfectly competitive labor market. The current employer initially chooses the level of general training provided to the worker. Then, firms compete with each other by making wage offers to attract the trained worker. Finally, production takes place (with the worker's productivity being positively related to the level of prior investment in general human capital). In this setting, the worker reaps all productivity benefits associated with general training due to competitive wage offers in labor market. In turn, the employer anticipates that he will not be able to recoup the costs of his initial investment and thus optimally chooses to make zero investment in the first place.

In the face of empirical evidence showing that employers do invest in their employees' general skills (see e.g. Krueger, 1993; Autor, 1998; Acemoglu and Pischke 1998, 1999a; Booth and Bryan, 2002), a more recent branch of the related literature has studied the conditions under which firms are indeed willing to pay for their workers' general human capital. A number of studies made by Acemoglu and Pischke (1998, 1999a, 1999b, 2003) has concluded that positive firm-sponsored investment requires that the labor market is frictional (rather than perfectly competitive) and the marginal effect of training on productivity exceeds the marginal effect of training on worker's wage (i.e. there must be some degree of 'wage compression')[4]. When these conditions hold, the firm's investment will be positive but still inefficiently low relative to the first-best. The reason for underinvestment is that a higher level of general training increases the worker's outside

[2] See Pfeffer (1998) for an extensive discussion on benefits associated with this kind of 'people-centered management'.

[3] It is a noteworthy fact that Japanese firms provide 364 hours of training to new workers in the first six months of employment, while US firms firms only provide 42 hours on average (Pfeffer, 1998, p. 85-90).

[4] Similar arguments have been raised by Katz and Ziderman (1990) or Chang and Wang (1996), who assume that investment is not observable by other potential employers. More recent contributions on the subject of general training include Booth and Zoega (2000, 2004), Gersbach and Schmultzer (2003) or Balmaceda (2005, 2008).

wage, implying that the current employer has to pay a higher wage in order to keep the worker from moving to another firm. Furthermore, any increase in the worker's bargaining power implies a lower level of firm-sponsored investment (since it allows the worker to reap a larger share of productivity benefits associated with training) and thus reduces social welfare.

All the above studies seek to account for underinvestment (which is broadly consistent with evidence related to countries like the UK and the USA) but rule out the possibility of overinvestment in general human capital. Therefore, they cannot explain the exceptionally high levels of training observed in other cases (e.g. in Japanese firms)[5]. We develop a model that incorporates workers' career concerns and takes power considerations into account to provide a more unified framework explaining both underinvestment and overinvestment in general training[6]. The model assumes that the provision of general capacities shifts the balance of bargaining power within the organization. In particular, workers who receive more extensive general training strengthen their bargaining position vis-à-vis their employer and thus can extract a higher proportion of the produced surplus in future periods. This assumption usually implies that the employer rationally provides inefficiently low levels of general training in order to secure his power in the long-run[7]. We show, however, that this result does not necessarily hold in the presence of workers' career concerns.

The literature on career concerns starts with Fama (1980), who argued that workers-managers' career concerns provide them with implicit incentives that motivate high levels of effort even in the absence of explicit agency contracts. This idea has been formalized

---

[5] Different patterns of training between Japanese and US firms have also been studied by Owan (2004) in the context of asymmetric learning about workers' characteristics.

[6] The question of power in the context of career concerns has also been studied by Ortega (2003) in a model of team production. However, Ortega's conceptualization of power is entirely different from the one adopted in our model. In particular, he assumes that power is captured by the marginal effect of effort on firm's performance. This means that power does not confer any direct material benefits to workers-managers, contrasting the assumptions made here.

[7] This argument has been formulated by Wright (2004). Similarly, Marglin (1974) has argued that power (rather than efficiency) considerations may explain the evolution of organizational practices within the workplace. In particular, hierarchical work organization and extensive specialization might be viewed as a kind of divide-and-conquer strategy adopted by employers in order to secure their role in the production process as integrators of workers' individual efforts into marketable products. In the same context, Gindin (1998) claims that employers are more willing to make concessions at the level of wages than to pay for employees' general training because the widespread development of workers' general capacities for creative planning and decision-making would generate a far more decisive shift in the balance of class forces than a direct increase of workers' purchasing power.

by Holmstrom (1982; 1999) in a multi-period model where there is symmetric (but imperfect) information about the worker's ability and, at the same time, the worker's effort is nonverifiable (i.e. there is moral hazard). Under the assumption that commitment to a long-term contract is not possible, the employer makes a wage offer to the worker at the beginning of each period. Since the wage offer is contingent on the expected output of the respective period (given the history of past output realizations), the worker has an incentive to work harder in early periods so as to raise market beliefs about her ability, thus increasing the expected output and wage offers to be received in future periods[8].

The two-period model developed below studies the firm's incentives to provide general training in the presence of worker's career concerns. In the first period, the employer initially chooses the level of investment in general human capital. Next, the employer makes a wage offer to the worker according to the specified structure of the labor market. Then, the worker chooses her (nonverifiable) effort contribution and the first-period output is produced. All agents' beliefs about the worker's ability are updated given the realization of first-period output (which is observable by all agents). The acquisition of general skills increases the worker's productivity and, furthermore, increases the worker's bargaining power vis-à-vis the employer in the second period. This shift in the balance of power is reflected in the new wage offer: A higher level of general skills acquired in the first period implies that the worker can extract a higher share of the second-period expected output. After the new wage has been determined, the worker chooses her new effort contribution and the second-period output is realized.

In this framework, it is shown that the worker's implicit incentives to contribute effort in the first period increase with the level of general training provided by the firm. The employer anticipates this positive effect of training on worker's incentives and is thus willing to invest in general human capital with higher intensity. In fact, when this positive impact on worker's incentives is strong enough, the employer may even overinvest (relative to the first-best) in equilibrium. Furthermore, it is shown that a higher level of worker's power (associated with an additional unit of training) may enhance the

[8] Further contributions on the issue of career concerns include Gibbons and Murphy (1992) or Andersson (2002), who study the interaction between career concerns and explicit incentives optimally designed in an agency contract. This question has also been investigated by Auriol, Friebel and Pechlivanos (2002) in the context of teams. The basic career concerns model has also been extended in multiple directions by Dewatripont et al (1999).

employer's investment incentives, due to the anticipated positive impact of this investment on worker's incentives. This result is in contrast with the usual prediction that any increase in the worker's bargaining power has a negative impact on firm's investment incentives. Turning to welfare implications, we study the socially optimal level of worker's power and identify conditions under which initial increases in worker's power are welfare-enhancing. More precisely, it is shown that a zero level of power can be optimal only if the worker's average ability is high enough. In all other cases, there is an inverse-U relationship between power and welfare: Initial increases in worker's power enhance social welfare, whereas further increases above a critical level of power will be detrimental to welfare. A series of numerical examples is used to illustrate these results.

The rest of the paper is organized as follows: In Section 2, the basic model is introduced; Section 3 involves the computation of the first-best outcome, while Section 4 deals with the derivation of equilibrium in the second-best environment. The positive implications of equilibrium are explored in Section 5, whereas welfare implications are investigated in Section 6. Finally, Section 7 discusses possible extensions as well as policy implications and Section 8 provides some concluding remarks.

2. The Model

We consider a two-period model. We assume that there is one worker **A** in the economy, one current employer **P** and many potential employers who are identical with P. On the other hand, there is a consumption good **x** which is produced in each period according to the technology specified below. The worker contributes effort α_t to the production of the consumption good in each period (*t =1, 2*). Finally, we denote by ***I*** the level of investment in general skills. The cost of this investment is paid by the employer in the first period, whereas future benefits are realized in the form of increased second-period worker productivity. In particular, we assume the following production technology:

$x_1 = n + a_1 + \varepsilon_1$: Output in the first period

$x_2 = n + a_2 + \varepsilon_2 + I$: Output in the second period

where n denotes the worker's innate ability, $a_t \in [0, +\infty)$ is the level of effort contributed by the worker in period t and ε_t is an idiosyncratic productivity shock. The level of effort α_t is nonverifiable - i.e. there is moral hazard. The output x_t in each period is observable by all agents in the economy. Furthermore, we assume that n and ε_t are normally distributed:

- $n \sim N(m_0, 1/h_0)$: Distribution of worker's innate ability[9].
- $\varepsilon_t \sim N(0, 1/h_\varepsilon)$: Distribution of the productivity shock in period t.

The random variables n, ε_1 and ε_2 are (identically and) independently distributed:

- $\mathrm{cov}(\varepsilon_1, \varepsilon_2) = \mathrm{cov}(n, \varepsilon_1) = \mathrm{cov}(n, \varepsilon_2) = 0$

The agents' preferences are represented by the following (expected) utility functions:

$EU_A = E\{\sum_{t=1}^{2} \beta^{t-1}[x_t^A - \frac{\rho}{2}\alpha_t^2] = E\{x_A^1 - \frac{\rho}{2}\alpha_1^2 + \beta(x_2^A - \frac{\rho}{2}\alpha_2^2)\}$: A's lifetime discounted expected utility.

$EU_P = E\{x_P^1 - \frac{\theta}{2}\mathrm{I}^2 + \beta x_2^P\}$: P's lifetime discounted expected utility,

where x_t^i denotes agent i's consumption ($i \in \{A, P\}$) in period t and $\beta \in [0,1]$ is the (common) discount factor. The disutility of labor provided in each period is represented by the function $g(a_t) = \rho\alpha_t^2 / 2$ (with $\rho > 0$), while the cost of investment in general training is captured by the cost function $\tau(I) = \theta I^2 / 2$ (with $\theta > 0$).

3. The First-Best Outcome

Since all agents' preferences are represented by quasilinear utility functions, the optimality problem involves the maximization of total surplus (sum of utilities) subject to technological and resource constraints:

$$\max_{\{a_t, I, x_t^i\}_{t=1,2}^{i=A,P}} EU_A + EU_P = E\{x_1^{\mathrm{A}} + x_1^P - \frac{\rho}{2}\alpha_1^2 - \frac{\theta}{2}\mathrm{I}^2 + \beta(x_2^A + x_2^P - \frac{\rho}{2}\alpha_2^2)\}$$

s.t.: $x_1 = n + a_1 + \varepsilon_1$, $x_2 = n + a_2 + \varepsilon_2 + \mathrm{I}$: Technological Constraints

$x_1^{\mathrm{A}} + x_1^P \le x_1$, $x_2^{\mathrm{A}} + x_2^P \le x_2$: Resource Constraints

$a_t, I, x_t^i \ge 0$: Nonnegativity Constraints ($t=1,2$)

Equivalently, the above problem can be written:

[9] The model assumes that beliefs about n (represented by the above distribution) are symmetric – i.e. they are shared by all agents in the economy.

$$\max_{\{a_t,I\}_{t=1,2}} TS = E\{x_1 - \frac{\rho}{2}\alpha_1^2 - \frac{\theta}{2}\mathrm{I}^2 + \beta(x_2 - \frac{\rho}{2}\alpha_2^2)\} = E\{n + a_1 + \varepsilon_1 - \frac{\rho}{2}\alpha_1^2 - \frac{\theta}{2}\mathrm{I}^2 + \beta(n + a_2 + \varepsilon_2 - \frac{\rho}{2}\alpha_2^2)\}$$

$$= (1+\beta)m_0 + a_1 - \frac{\rho}{2}\alpha_1^2 - \frac{\theta}{2}\mathrm{I}^2 + \beta(a_2 - \frac{\rho}{2}\alpha_2^2)\}$$

The solution of this problem is:

$a_1^{FB} = a_2^{FB} = a^{FB} = 1/\rho$, $\mathrm{I}^{FB} = \beta/\theta$

yielding the first-best level of welfare:

$$W^{FB} = (1+\beta)m_0 + (1+\beta)[a^{FB} - \frac{\rho}{2}(\alpha^{FB})^2] + \beta I^{FB} - \frac{\theta}{2}(I^{FB})^2$$

4. The Second-Best Outcome: Perfect Bayesian Equilibrium

In order to find the equilibrium in our second-best environment, we must define the structure of the labor market. Since there is no possibility of commitment to a long-term contract, the employer makes a wage offer at the beginning of each period. The wage offer is contingent on the expected output of the respective period given the history of past output realizations (implying that there exist implicit incentives – i.e. career concerns – for the worker to provide effort). In particular, we assume the following wage structure (where w_t denotes the wage offer at the beginning of period t):

$$w_1 = kE(x_1 / prior),\ k \in [0,1] \qquad (1)$$

The term *1-k* represents the employer's ex-ante monopsony power. Of course, if *k=1* the labor market is perfectly competitive and the worker receives her full (marginal) productivity; this extreme case replicates the assumption in Becker's seminal model. It is already known by the related literature (and it will also be made clear below) that a frictional (imperfect) labor market (*k<1*) is a necessary condition for the employer to have any incentive to make a positive investment in general human capital. For the second period wage offer, we assume the following structure:

$$w_2 = \min\{k + \delta I, 1\} \cdot E(x_2 / x_1)\ ,\ \delta \geq 0 \qquad (2)$$

The implicit assumption here is that the worker's bargaining power vis-à-vis the employer in the second period increases with the level of general training received in the first period. This reflects the argument that the provision of general training shifts the balance of power in favor of workers. Due to the very nature of general training, the worker's (second-period) outside wage opportunity increases with the level of training received in the first period. This implies that the current employer must pay a higher second-period wage to keep the worker. This reasoning is depicted in the wage structure specified above.

The parameter δ represents the increase in worker's second-period bargaining power (or the second-period wage premium) associated with an additional unit of general human capital. The crucial question is how this wage structure affects the worker's incentives to provide effort as well as the employer's investment incentives.

The structure of wage payments implies the following payoffs for the worker and the employer, respectively:

$$EU_A = E\{w_1 - \frac{\rho}{2}\alpha_1^2 + \beta(w_2 - \frac{\rho}{2}\alpha_2^2)\}$$

$$EU_P = E\{x_1 - w_1 - \frac{\theta}{2}\mathrm{I}^2 + \beta(x_2 - w_2)\}$$

The timing of the associated game is the following:

- Period 1 (*t=1*)

1. P chooses the level of investment in general training *I*.

2. P offers $w_1 = kE(x_1 / prior)$ to A according to the wage structure specified in (1).

3. A chooses her first-period effort contribution α_1 and output x_1 is realized.

4. All agents update their beliefs about A's ability given the first-period output realization.

- Period 2 (*t=2*)

1. P offers w_2 according to the wage structure specified in (2).

2. A chooses her second-period effort contribution α_2 and output x_2 is realized.

The equilibrium concept which must be used here is the Perfect Bayesian Equilibrium (PBE). In particular, the equilibrium outcome consists of strategies (α_1^*, a_2^*, I^*), conjectures $(\overline{a}_1, \overline{a}_2)$ and beliefs about the worker's ability such that:

(i) A's strategy (a_1^*, a_2^*) maximizes her expected payoff.

(ii) P's strategy I^* maximizes his expected payoff.

(iii) Conjectures are correct: $(\overline{a}_1, \overline{a}_2) = (a_1^*, a_2^*)$

(iv) Beliefs about the worker's ability are derived from Bayes' rule given equilibrium strategies.

We use backward induction to compute the equilibrium outcome. In the second period, the worker A chooses her effort contribution α_2 (given *I*, w_1, α_1 and w_2) so as to maximize her expected payoff:

$$\max_{\{a_2\}} EU_A = w_1 - \frac{\rho}{2}\alpha_1^2 + \beta(w_2 - \frac{\rho}{2}\alpha_2^2)$$

Of course, the (Kuhn-Tucker) first order conditions imply the solution:

$$a_2^* = 0 \qquad (3)$$

Since the worker is motivated only by her career concerns, it is clear that she has no incentive to provide any effort at all in the last period of the interaction.

Anticipating the worker's rational behavior in the second period ($a_2^* = 0$), the employer makes the second-period wage offer (given *I, w_1, α_1*) according to the wage structure specified in (2):

$$w_2(x_1) = \min\{k + \delta I, 1\} \cdot E(x_2 / x_1) = \min\{k + \delta I, 1\} \cdot [I + E(n / x_1)] \quad (4)$$

Given the output realization x_1, all agents' beliefs about the worker's ability are updated at the end of the first period. We define a new random variable:

$z_1 \equiv x_1 - \overline{a}_1 = n + a_1 + \varepsilon_1 - \overline{a}_1 = n + \varepsilon_1$, which is also normally distributed since n and ε_1 are normally distributed (note that $a_1 = \overline{a}_1$ along the equilibrium path). Therefore, we can use the following normal updating formula associated with two normally distributed random variables y_1, y_2:

$y_1 / y_2 \sim N(\mu_1 + \frac{\sigma_{12}}{\sigma_2^2}(y_2 - \mu_2), (1 - r^2)\sigma_1^2)$, where:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right), \; r = \frac{\sigma_{12}}{\sigma_1 \sigma_2}.$$

For $y_1 = n,\ y_2 = z_1$ the above formula yields:

$$n / z_1 \sim N(\frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon}, \frac{1}{h_0 + h_\varepsilon}) \quad (5)$$

We have: $E(n / z_1) = \frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon} \Rightarrow E(n / x_1) = \frac{h_0 m_0 + h_\varepsilon (x_1 - \overline{a}_1)}{h_0 + h_\varepsilon}$ (6)

Now, we can substitute (6) into (4) to get:

$$w_2(x_1) = \min\{k + \delta I, 1\} \cdot [I + \frac{h_0 m_0 + h_\varepsilon (x_1 - \overline{a}_1)}{h_0 + h_\varepsilon}] \quad (7)$$

Anticipating $w_2(x_1)$ in (7) and $a_2^* = 0$, the worker chooses her effort contribution in the first period (given I, w_1) so as to maximize her expected payoff:

$$\max_{\{a_1\}} EU_A = w_1 - \frac{\rho}{2}\alpha_1^2 + \beta[Ew_2(x_1) - \frac{\rho}{2}(\alpha_2^*)^2] =$$

$$= w_1 - \frac{\rho}{2}\alpha_1^2 + \beta \cdot \min\{k + \delta I, 1\} \cdot \left[I + E\left(\frac{h_0 m_0 + h_\varepsilon (n + a_1 + \varepsilon_1 - \overline{a}_1}{h_0 + h_\varepsilon} \right) \right]$$

The first-order necessary (and sufficient) conditions for maximization are:

$\frac{\partial EU_A}{\partial a_1} = \beta \cdot \min\{k+\delta I, 1\} \cdot \frac{h_\varepsilon}{h_0 + h_\varepsilon} - \rho \alpha_1 \leq 0$, $\frac{\partial EU_A}{\partial a_1}\alpha_1 = 0$. These conditions yield the solution:

$$\alpha_1 = \frac{\beta}{\rho} \cdot \min\{k+\delta I, 1\} \cdot \frac{h_\varepsilon}{h_0 + h_\varepsilon} \qquad (8)$$

Note that (for $I \leq (1-k)/\delta$) the worker's implicit incentives are strengthened when she receives more training by the employer in the first place. This happens because higher training implies that the worker has a larger share in the second-period expected output and thus has stronger incentives to raise market beliefs about her ability by working harder in the first period:

$$\frac{\partial \alpha_1}{\partial I} = \frac{\beta}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} \delta > 0$$

In other words, the worker reciprocates the firm's willingness to provide her with general training by increasing her own willingness to work for the employer in the first period. This kind of reciprocal behavior is confirmed by field evidence[10]. It should also be emphasized that this reciprocal effect of firm-sponsored training on worker's incentives is not the result of any behavioral assumption stipulating reciprocal preferences for the worker (as in Leuven et al, 2002).

The first-period wage is calculated (anticipating the worker's choice of α_1) according to the wage structure specified in (1):

$$w_1 = kE(x_1 / prior) = k(m_0 + a_1) \qquad (9)$$

where α_1 is given in (8).

At the first stage of the game, the employer P chooses the level of investment in general training I (anticipating w_1, α_1, w_2, α_2 as given in (9), (8), (7), (3) respectively) so as to maximize his lifetime expected payoff:

$$\max_{\{I\}} EU_P = E\{x_1 - w_1 - \frac{\theta}{2}\mathrm{I}^2 + \beta(x_2 - w_2)\} =$$
$$= (1-k)(m_0 + \min\{k+\delta I, 1\} \cdot \frac{\beta}{\rho} \cdot \frac{h_\varepsilon}{h_0 + h_\varepsilon}) - \frac{\theta}{2} I^2 + \beta(m_0 + I)(1 - \min\{k + \delta I, 1\})$$

[10] For example, Pfeffer (1998, p.89) refers to the case of Taco Inc (a private manufacturer with 450 employees) offering "astonishing educational opportunities" to its employees in an on-site learning center. When the company's chief executive John Haze White was asked to put a monetary value on the returns from operating the center (which had a high cost for the firm to build and, furthermore, implies significant direct expenses and lost production costs every year), he said: "*It comes back in the form of attitude*. People feel they are playing in the game – not being kicked around in it" (emphasis added).

For $I \geq (1-k)/\delta$, we have: $\partial EU_P / \partial I = -\theta I < 0$. This means that the employer's investment will never be higher than $(1-k)/\delta$. Consequently, we focus on the interval $0 \leq I \leq (1-k)/\delta$ and solve the problem:

$$\max_{\{I\}} EU_P = (1-k)(m_0 + a_1) - \frac{\theta}{2}\mathrm{I}^2 + \beta(m_0 + I)(1-k-\delta I)$$

s.t. $0 \leq I \leq (1-k)/\delta$

We write down the Lagrangian and the Kuhn-Tucker necessary (and sufficient) conditions for maximization:

$$L = (1-k)(m_0 + a_1) - \frac{\theta}{2}\mathrm{I}^2 + \beta(m_0 + I)(1-k-\delta I) + \lambda(\frac{1-k}{\delta} - I)$$

The *FOCs* are:

$$\frac{\partial L}{\partial I} = \frac{\partial EU_P}{\partial I} - \lambda = (1-k)\frac{\partial a_1}{\partial I} - \theta I + \beta(1-k-\delta I) - \beta\delta(m_0 + I) - \lambda \leq 0,\ \frac{\partial L}{\partial I} I = 0$$

The first term represents the employer's marginal benefit of investment in terms of worker's increased first-period effort. The second term is the direct marginal cost of investment. The third term is the marginal benefit in terms of increased second-period output and the fourth term is the marginal cost in terms of worker's increased second-period wage.

$\frac{\partial L}{\partial \lambda} = \frac{1-k}{\delta} - I \geq 0$, $\frac{\partial L}{\partial \lambda}\lambda = 0$. These conditions yield the following solution:

$$I^* = \frac{1-k}{\delta} = I_H \qquad \text{, if } m_0 \leq R - \frac{2\beta\delta + \theta}{\beta\delta} I_H$$

$$I^* = \frac{\beta\delta}{2\beta\delta + \theta}(R - m_0) \ \text{, if } R - \frac{2\beta\delta + \theta}{\beta\delta} I_H \leq m_0 \leq R$$

$$I^* = 0 \qquad \text{, if } m_0 \geq R$$

where: $R = (1-k)(\frac{1}{\delta} + \frac{1}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon})$

The employer anticipates the positive impact of general training on worker's first-period incentives and takes this effect into account when choosing *I* in the first place. Therefore, the employer is willing to invest in general human capital with higher intensity because he predicts the positive effect of investment on worker's effort contribution α_1. This interaction between the worker's incentives to provide effort and the employer's investment incentives provides an additional rationale (which has not been taken into account so far in the associated literature) for firm-sponsored general training.

Note that $I^*=0$ when $k=1$: This case corresponds to Becker's prediction that the firm will make zero investment if there are no frictions in the labor market (i.e. if the labor market is perfectly competitive).

The equilibrium outcome is summarized in the following proposition.

Proposition 1. The Perfect Bayesian Equilibrium involves:

$$I^* = \begin{cases} \dfrac{1-k}{\delta} = I_H & \text{, if } m_0 \le R - \dfrac{2\beta\delta+\theta}{\beta\delta} I_H \\ \dfrac{\beta\delta}{2\beta\delta+\theta}(R-m_0) & \text{, if } R - \dfrac{2\beta\delta+\theta}{\beta\delta} I_H \le m_0 \le R \\ 0 & \text{, if } m_0 \ge R = (1-k)(\dfrac{1}{\delta}+\dfrac{1}{\rho}\dfrac{h_\varepsilon}{h_0+h_\varepsilon}) \end{cases}$$

$$a_1^* = \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(k+\delta I^*)\ ,\ a_2^* = 0$$

The equilibrium welfare is:

$$TS = W^* = (1+\beta)m_0 + a_1^* - \frac{\rho}{2}(\alpha_1^*)^2 + \beta I^* - \frac{\theta}{2}(I^*)^2$$

5. Implications

5.1 The Impact of Power on Investment and Effort

The negative relationship between m_o and I^* (which is depicted in Figure 1) is contingent on the specific production technology assumed here – i.e. on perfect substitutability between innate ability n and investment in general human capital. Some other comparative statics results are easy to understand. Equilibrium investment decreases with k (i.e. investment decreases as the labor market becomes more competitive) and increases with the discount factor β. Furthermore, a decrease in ρ or h_o as well as an increase in h_ε increase investment, because they imply a stronger positive impact of I on worker's first-period effort. The employer takes this effect into account and thus has an incentive to invest with higher intensity.

We proceed to study the effect of worker's power (captured by the parameter δ) on employer's investment incentives. To this end, we rewrite the equilibrium level of training as a function of δ. After the appropriate calculations (which can be found in the Appendix A), we find:

- **Case 1.** For $m_0 \le \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$:

$$I^* = \begin{cases} \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0) & \text{, if } \delta \le \delta_2 \\ \frac{1-k}{\delta} = I_H & \text{, if } \delta \ge \delta_2 \end{cases}$$

where: $\delta_2 = \dfrac{\beta(1-k) + \sqrt{\beta^2(1-k)^2 + 4\beta\theta(1-k)(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0)}}{2\beta(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0)}$

- **Case 2.** For $m_0 \ge \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$:

$$I^* = \begin{cases} \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0) & \text{, if } \delta \le \delta_0 \\ 0 & \text{, if } \delta \ge \delta_0 \end{cases}$$

where: $\delta_0 = \dfrac{1-k}{m_0 - \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}}$

For the parameter interval where $I^*=I_H$, it is obvious that $\partial I^*/\partial\delta < 0$. For the intervals where $I^* = \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0)$, we have:

$$\frac{\partial I^*}{\partial\delta} = \frac{-2\beta}{(2\beta\delta+\theta)^2}[(1-k)\frac{\partial a_1}{\partial I} + \beta(1-k-\delta m_0)] + \frac{1}{2\beta\delta+\theta}[(1-k)\frac{\partial^2 a_1}{\partial I\partial\delta} - \beta m_0] \qquad (10)$$

The first term in (10) is the direct negative effect of worker's power on employer's incentives to invest in general training, since a higher value of δ reduces the employer's share in the second-period output. But the second term represents an indirect positive effect (since $\frac{\partial^2 a_1}{\partial I\partial\delta} = \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} > 0$) associated with the fact that a higher value of δ implies a stronger positive effect of training on worker's incentives and thus makes the firm more willing to invest in general human capital. If this positive effect is stronger than the direct negative effect, then the overall impact of worker's power on employer's investment incentives will be positive. In particular, we can find the results summarized in the next proposition:

Proposition 2.

- **Case 1a:** If $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I}$ then $\frac{\partial I^*}{\partial \delta} > 0$ (and $\frac{\partial^2 I^*}{\partial \delta^2} < 0$) for all $\delta < \delta_2$,

  where: $\hat{I} = \frac{\beta(1-k)}{\theta}$ (see Figure 2)

- **Case 1b:** If $\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I} < m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then $\frac{\partial I^*}{\partial \delta} < 0$ (and $\frac{\partial^2 I^*}{\partial \delta^2} > 0$) for all $\delta < \delta_2$

  (see Figure 3)

- **Case 2:** If $m_0 > \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then $\frac{\partial I^*}{\partial \delta} < 0$ (and $\frac{\partial^2 I^*}{\partial \delta^2} > 0$) for all $\delta < \delta_0$

  (see Figure 4)

Proof: See Appendix B.

The comparative statics analysis for case 1a shows that a higher level of worker's power (up to a threshold value δ_2) increases the employer's incentives to invest in general training. In particular, this happens when the worker's expected ability is low enough. This result contrasts the predictions of previous studies, according to which any increase in worker's power weakens the employer's investment incentives.

We can also derive the effect of worker's power δ on first-period equilibrium effort. First, we write down a_1^* derived above as a function of δ:

- **Case 1.** For $m_0 \le \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$:

$$a_1^* = \begin{cases} \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(k+\delta I^*) & \text{, if } \delta \le \delta_2 \ \ (\text{where } I^* = \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0)) \\ \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} & \text{, if } \delta \ge \delta_2 \end{cases}$$

- **Case 2.** For $m_0 \ge \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$:

$$a_1^* = \begin{cases} \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(k+\delta I^*) & \text{, if } \delta \le \delta_0 \\ \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}k & \text{, if } \delta \ge \delta_0 \end{cases}$$

For $\delta \le \delta_2$ in case 1 and for $\delta \le \delta_0$ in case 2, we have[11]:

$$\frac{\partial a_1^*}{\partial \delta} = \frac{\beta}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} (I^* + \delta \frac{\partial I^*}{\partial \delta}) \qquad \textbf{(14)}$$

The first term in the parenthesis is the direct positive effect of power on worker's incentives associated with the higher share extracted by the worker in the second period. However, the second term can be either positive or negative depending on the overall impact of δ on employer's investment incentives (as described in Proposition 2).

Equation (14) implies the following result:

Lemma 1: If $\frac{\partial I^*}{\partial \delta} > 0$ then $\frac{\partial a_1^*}{\partial \delta} > 0$, too.

Clearly, if an increase in worker's power has a positive impact on employer's investment incentives, then it surely increases the worker's first-period implicit incentives to provide effort, too. On the other hand, if the overall impact of power on equilibrium training is negative, then the sign of $\partial a_1^* / \partial \delta$ is ambiguous. The appropriate calculations yield the following result:

Proposition 3.

- **Case 1a:** If $m_0 < \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - 2\hat{I}$ then: $\frac{\partial a_1^*}{\partial \delta} > 0$ (and $\frac{\partial^2 a_1^*}{\partial \delta^2} > 0$) for $\delta < \delta_2$

(see Figure 5)

- **Case 1b:** If $\frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - 2\hat{I} < m_0 < \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ then: $\frac{\partial a_1^*}{\partial \delta} > 0$ (and $\frac{\partial^2 a_1^*}{\partial \delta^2} < 0$) for $\delta < \delta_2$

(see Figure 6)

- **Case 2:** If $m_0 > \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ then: $\frac{\partial a_1^*}{\partial \delta} > 0$ for $\delta < \delta_4$

$\frac{\partial a_1^*}{\partial \delta} < 0$ for $\delta_4 < \delta < \delta_0$ (and $\frac{\partial^2 a_1^*}{\partial \delta^2} < 0$ for $\delta < \delta_0$)

where: $$\delta_4 = \frac{-2\theta(m_0 - \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}) + \sqrt{4\theta^2 (m_0 - \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon})^2 + 8\beta\theta(1-k)(m_0 - \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon})}}{4\beta(m_0 - \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon})}$$

(see Figure 7)

[11] Clearly, we have $\partial a_1^* / \partial \delta = 0$ for all other parameter intervals.

Proof: See Appendix C.

Note that when the worker's expected ability is high enough (case 2), any increase in worker's power above a critical value δ_4 implies a negative impact on employer's investment incentives ($\partial I^* / \partial \delta < 0$) so strong that it also pushes worker's first-period incentives downward ($\partial a_1^* / \partial \delta < 0$).

5.2 The Possibility of Overinvestment in General Training

In previous models studying the issue of general training, a common prediction is that the employer underinvests in general human capital (relative to the first-best) in equilibrium. In this subsection, we show that equilibrium is compatible with overinvestment under certain conditions. From the comparative statics analysis conducted above, it is clear that overinvestment can arise only in case 1a. To see this, note first that $I^* = \hat{I} < I^{FB}$ when $\delta=0$. Furthermore, in cases 1b and 2 any increase of worker's power δ reduces equilibrium investment further below the first-best level, implying that underinvestment is always the case. Therefore, we focus on case 1a (where investment in training increases with worker's power in the interval $\delta < \delta_2$) to examine the possibility of overinvestment. In particular, we can state the following proposition:

Proposition 4. If $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon} - \frac{2-k}{1-k} \cdot I^{FB}$

then equilibrium involves overinvestment in general training ($I^* > I^{FB}$) for $\delta_5 < \delta < \delta_6$,

where: $\delta_5 = \dfrac{k}{\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon} - m_0 - 2I^{FB}}$, $\delta_6 = \dfrac{\theta(1-k)}{\beta}$

Proof: See Appendix D.

The case of overinvestment is depicted in Figure 8.

The driving force of overinvestment is precisely the positive effect of general training on worker's first-period incentives to provide effort. As already noted, the employer takes this effect into account and is thus willing to provide general skills with higher intensity. If the positive effect on worker's incentives is strong enough, then the employer's investment incentives can be excessively strong.

Note that overinvestment by the employer requires that the worker's expected innate ability m_0 is sufficiently small. This is fairly intuitive in our framework, since training and innate ability are perfect substitutes in the production function. Of course, the opposite (i.e. overinvestment for high values of expected innate ability) might be the case if the production technology assumed that general human capital and innate ability are complements. Furthermore, it should be noted that the case of overinvestment arises for intermediate values of worker's power ($\delta_5 < \delta < \delta_6$). To the contrary, underinvestment is always the case for both too low and too high values of δ. If the worker's power is very small ($\delta < \delta_5$), then the employer anticipates that his investment will not have a very strong positive effect on worker's first-period incentives; as a result, the employer does not invest much. On the other hand, if the worker is too strong ($\delta > \delta_6$) then the firm's investment incentives are again pushed below the optimum (note that, in this case, even a small level of investment implies that the worker enjoys the full second-period expected output). Therefore, the model predicts that overinvestment can take place only for intermediate values of worker's power.

It is well-known that long-term employment is a distinct characteristic of the Japanese employment system. In terms of our model, this means that workers in Japanese firms have relatively strong implicit incentives to provide effort in early periods of their career. In terms of the model developed here, the positive impact of these strong career concerns on employers' investment incentives may explain the high levels of training typically observed in Japanese firms.

6. Welfare Analysis

In Proposition 1, we have stated the equilibrium level of welfare:

$$TS = W^* = (1+\beta)m_0 + a_1^* - \frac{\rho}{2}(\alpha_1^*)^2 + \beta I^* - \frac{\theta}{2}(I^*)^2$$

We focus here on the effect of worker's power δ on social welfare:

$$\frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha_1^*}{\partial \delta}(1-\rho a_1^*) + \frac{\partial I^*}{\partial \delta}(\beta - \theta I^*)$$

The above equation shows that the overall impact of power on welfare is the sum of two separate effects: First, the effect of power on worker's first-period incentives; and second, the effect of power on employer's investment incentives. These two effects can work either in the same or in opposite directions (in the latter case, the total impact of worker's power on welfare is ambiguous). We are interested in finding the optimal level of power (in the sense of welfare maximization). It will be made clear that a zero level of power is socially optimal only for a limited range of parameters, while in all other cases the optimization problem has an interior solution. This also contrasts previous models of general training, where any increase in worker's power is detrimental to welfare. We study each parameter interval separately first:

- **<u>Case 1a:</u>** For $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I}$

(i) If $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{2-k}{1-k}I^{FB}$, then we have:

- For $\delta<\delta_5$: $\frac{\partial W^*}{\partial \delta} = \underset{(+)}{\frac{\partial \alpha_1^*}{\partial \delta}}\underset{(+)}{(1-\rho a_1^*)} + \underset{(+)}{\frac{\partial I^*}{\partial \delta}}\underset{(+)}{(\beta-\theta I^*)} > 0$

In words: Initial increases in worker's power enhance welfare, since they imply a more efficient (higher) first-period effort choice by the worker as well as a more efficient (higher) investment choice by the employer.

- For $\delta_5<\delta<\delta_2$: $\frac{\partial W^*}{\partial \delta} = \underset{(+)}{\frac{\partial \alpha_1^*}{\partial \delta}}\underset{(+)}{(1-\rho a_1^*)} + \underset{(+)}{\frac{\partial I^*}{\partial \delta}}\underset{(-)}{(\beta-\theta I^*)}$, which can be either positive or negative. In this interval, equilibrium involves overinvestment and any increase in worker's power (up to δ_2) keeps pushing the level of training further above the optimum. This effect is now detrimental to welfare but coexists with the positive effect of power on worker's implicit incentives. As a result, the overall impact on welfare is ambiguous.

- For $\delta_2<\delta<\delta_6$: $\frac{\partial W^*}{\partial \delta} = \underset{(=0)}{\frac{\partial \alpha_1^*}{\partial \delta}}\underset{(+)}{(1-\rho a_1^*)} + \underset{(-)}{\frac{\partial I^*}{\partial \delta}}\underset{(-)}{(\beta-\theta I^*)} > 0$

In this interval, a further increase in worker's power has again an unambiguously positive effect on welfare, because it implies a lower (i.e. closer to the first-best) level of investment in training, while the worker's incentives remain unaffected.

- For $\delta>\delta_6$: $\frac{\partial W^*}{\partial \delta} = \underset{(=0)}{\frac{\partial \alpha_1^*}{\partial \delta}}\underset{(+)}{(1-\rho a_1^*)} + \underset{(-)}{\frac{\partial I^*}{\partial \delta}}\underset{(+)}{(\beta-\theta I^*)} < 0$

As the worker's power increases very much, the overall impact on welfare becomes negative, since the employer's investment falls again below the optimum and keeps falling to more and more inefficiently low levels (while the worker's incentives remain unaffected).

Note that δ=δ_6 is a local (and potentially global) welfare maximizer in this case.

We proceed in the same way to examine the other parameter intervals (in which equilibrium always involves underinvestment).

(ii) If $\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}-\frac{2-k}{1-k}I^{FB}<m_0<\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}-2\hat{I}$, then we have:

- For $\delta<\delta_2$: $\frac{\partial W^*}{\partial\delta}=\underset{(+)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(+)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}>0$

- For $\delta>\delta_2$: $\frac{\partial W^*}{\partial\delta}=\underset{(=0)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(-)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}<0$

Note that δ=δ_2 is the socially optimal level of worker's power in this case.

- **Case 1b**: If $\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}-2\hat{I}<m_0<\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$, then we have:

- For $\delta<\delta_2$: $\frac{\partial W^*}{\partial\delta}=\underset{(+)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(-)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}$ which can be either positive or negative.

- For $\delta>\delta_2$: $\frac{\partial W^*}{\partial\delta}=\underset{(=0)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(-)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}<0$

- **Case 2**: If $m_0>\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$, then we have:

- For $\delta<\delta_4$: $\frac{\partial W^*}{\partial\delta}=\underset{(+)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(-)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}$ which can be either positive or negative.

- For $\delta_4<\delta<\delta_0$: $\frac{\partial W^*}{\partial\delta}=\underset{(-)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(-)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}<0$

- For $\delta>\delta_0$: $\frac{\partial W^*}{\partial\delta}=\underset{(=0)}{\frac{\partial\alpha_1^*}{\partial\delta}}\underset{(+)}{(1-\rho a_1^*)}+\underset{(=0)}{\frac{\partial I^*}{\partial\delta}}\underset{(+)}{(\beta-\theta I^*)}=0$

The above results show that a zero level of worker's power is a possible welfare maximizer only in cases 1b and 2. On the contrary, the welfare maximizing level of power is always positive (and bounded) in case 1a.

In order to derive more precise results, we write down the general welfare maximization problem for the parameter intervals where $I^* = \beta\delta(R - m_0)/(2\beta\delta + \theta)$ (i.e. for $\delta<\delta_2$ in case 1 and for $\delta<\delta_0$ in case 2)[12]. Define $\bar{\delta}$ as follows:

$$\bar{\delta} \equiv \delta_2 \text{ for } m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon} \text{ (case 1)}$$

$$\bar{\delta} \equiv \delta_0 \text{ for } m_0 > \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon} \text{ (case 2).}$$

Then, we must solve the following maximization problem:

$$\max_{\{\delta\}} W^*(\delta) = (1+\beta)m_0 + a_1^* - \frac{\rho}{2}(a_1^*)^2 + \beta I^* - \frac{\theta}{2}(I^*)^2$$

s.t: $0 \le \delta \le \bar{\delta}$

The solution of this problem allows us to state the following proposition:

Proposition 5. The socially optimal level of worker's power can be zero only if:

$$m_0 \ge \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0 + h_\varepsilon}\left[1 + \frac{\beta(1-\beta k\frac{h_\varepsilon}{h_0 + h_\varepsilon})}{\beta - \theta \hat{I}}\right] - 2\hat{I} = m^*$$

i.e. only if the worker's expected innate ability is high enough.

Proof: See Appendix E.

This is a necessary condition for $\delta=0$ to be a welfare maximizer. When this condition holds, there is always underinvestment in general human capital and, furthermore, any initial increase in worker's power weakens the employer's investment incentives – thus pushing I^* further below the optimum. When this negative impact on welfare dominates the positive impact associated with the worker's stronger incentives to provide effort in the first period, then any initial increase in worker's power above $\delta=0$ will be detrimental to welfare.

Finally, Proposition 5 along with the overall welfare analysis implies the following result:

Proposition 6. If $m_0 < m^*$, then there is a positive and bounded level of power δ^* which maximizes social welfare.

[12] Note that we have already found the sign of $\partial W^*/\partial\delta$ for $\delta>\delta_2$ (case 1) or $\delta>\delta_0$ (case 2).

In this case, initial increases in worker's power up to a critical value δ^* enhance welfare, whereas further increases above that critical value will be detrimental to welfare[13].

We close the subsection related to the welfare analysis with some numerical examples, illustrating each of the different cases described above.

6.1 Numerical Examples

(i) Let $\beta = \rho = h_\varepsilon = h_0 = 1$, $k = 1/2$, $m_0 = 0$, $\theta = 1$.

These parameter values correspond to *case 1b* (involving underinvestment for all values of δ). The equilibrium outcome is:

$$I^* = \frac{\delta + 2}{4(2\delta + 1)}, \text{ if } \delta \le \delta_2 = \sqrt{3} + 1$$

$$I^* = 1/2\delta, \text{ if } \delta \ge \sqrt{3} + 1$$

$$\alpha_1^* = \frac{1}{2}(\frac{1}{2} + \delta I^*), \text{ if } \delta \le \delta_2 = \sqrt{3} + 1$$

$$\alpha_1^* = 1/2, \text{ if } \delta \ge \sqrt{3} + 1 \quad \text{, and: } I^{FB} = a^{FB} = 1.$$

The welfare maximization problem becomes:

$$\max_{\{\delta\}} W^*(\delta) = a_1^* - \frac{1}{2}(a_1^*)^2 + I^* - \frac{1}{2}(I^*)^2$$

s.t. $0 \le \delta \le \sqrt{3} + 1$

The solution is $\delta = 0$. Since $\partial W^* / \partial \delta < 0$ for all $\delta > \delta_2 = \sqrt{3} + 1$, a zero level of worker's power maximizes equilibrium welfare in this case (i.e. $\delta^*=0$). Indeed, for these parameter values it can be directly verified that the condition stated in Proposition 5 holds.

(ii) Let $\beta = \rho = h_\varepsilon = h_0 = 1, k = 1/2$, $m_0 = 0$, $\theta = 8$.

These parameter values correspond to *case 1a-ii* (again with underinvestment for all values of δ). The equilibrium outcome in this case involves:

$$I^* = \frac{\delta + 2}{8(\delta + 4)}, \text{ if } \delta \le \delta_2 = \sqrt{17} + 1$$

$$I^* = 1/2\delta, \text{ if } \delta \ge \sqrt{17} + 1$$

$$\alpha_1^* = \frac{1}{2}(\frac{1}{2} + \delta I^*), \text{ if } \delta \le \sqrt{17} + 1$$

[13] It should be noted that a similar inverted U-shaped relationship between workers' power and efficiency has also been pointed out in a broader context by Wright (2000).

$\alpha_1^* = 1/2$, if $\delta \geq \sqrt{17}+1$, and: $I^{FB} = 1/8$, $a^{FB} = 1$.

The welfare maximization problem becomes:

$$\max_{\{\delta\}} W^*(\delta) = a_1^* - \frac{1}{2}(a_1^*)^2 + I^* - 4(I^*)^2$$

s.t. $0 \leq \delta \leq \sqrt{17}+1$

The solution now is $\delta = \delta_2 = \sqrt{17}+1$. Since $\partial W^*/\partial\delta < 0$ for all $\delta > \delta_2$, the optimal level of power is $\delta^* = \delta_2 = \sqrt{17}+1$ as already predicted above for the general case.

(iii) Let $\beta = \rho = h_\varepsilon = h_0 = 1$, $k = 1/2$, $m_0 = 0$, $\theta = 16$.

These parameter values correspond to *case 1a-i*, with overinvestment for $\delta_5 < \delta < \delta_6$. The equilibrium outcome is:

$$I^* = \frac{\delta+2}{8(\delta+8)}, \text{ if } \delta \leq \delta_2 = \sqrt{33}+1$$

$I^* = 1/2\delta$, if $\delta \geq \sqrt{33}+1$

$$\alpha_1^* = \frac{1}{2}(\frac{1}{2} + \delta I^*), \text{ if } \delta \leq \sqrt{33}+1$$

$\alpha_1^* = 1/2$, if $\delta \geq \sqrt{33}+1$ and: $I^{FB} = 1/16$, $a^{FB} = 1$, $\delta_5 = 4$, $\delta_6 = 8$

In this case, there is overinvestment in equilibrium for $\delta_5 < \delta < \delta_6$, i.e. for *4<δ<8*.

The welfare maximization problem becomes:

$$\max_{\{\delta\}} W^*(\delta) = a_1^* - \frac{1}{2}(a_1^*)^2 + I^* - 8(I^*)^2$$

s.t. $0 \leq \delta \leq \sqrt{33}+1$

The solution now is $\delta = \delta_2 = \sqrt{33}+1$. Since $\partial W^*/\partial\delta > 0$ for $\delta_2 < \delta < \delta_6 = 8$ and $\partial W^*/\partial\delta < 0$ for $\delta > \delta_6 = 8$, the social welfare is maximized for $\delta^* = \delta_6 = 8$.

Note that the condition (16) is not satisfied in examples (ii) and (iii), implying that *δ=0* cannot be a welfare maximizer in either of these cases.

7. Policy Implications and Possible Extensions

A standard question often addressed in the literature concerns the appropriate policy instrument or the optimal degree of public intervention in order to alleviate inefficiencies associated with incentives to invest in general training. In this respect, the prediction of underinvestment (which is common in all previous studies) by the employer implies that a government subsidy to the firm is the appropriate policy measure to strengthen the firm's investment incentives and enhance social welfare[14]. However, the framework adopted here shows that this is not necessarily the appropriate policy recommendation. The possibility of overinvestment in human capital implies that a subsidy to the firm may push the level of training further above the optimum and thus have a detrimental effect on welfare. On the other hand, such an additional increase in general training may also motivate the employee to work harder in the first period. The overall impact of the subsidy on welfare will depend on the relative strength of these two opposite effects. Of course, if there is underinvestment in equilibrium then a subsidy to the firm is indeed recommended because it pushes the employer's investment closer to the first-best and thus also strengthens the worker's first-period incentives.

In his seminal paper on general training, Becker suggested (in the context of a perfectly competitive labour market) that the worker herself should pay for general training because she faces the right incentives to invest efficiently in her own human capital. A possible extension of our paper is to check the validity of this claim within the framework adopted here. We can assume that the worker buys general human capital from a training company which sells general skills at an endogenously determined price. A welfare comparison between that regime and the one described here should take into account not only the amount of training but also the worker's incentives to provide effort in each case. Furthermore, in the light of these two regimes one might also examine whether a government subsidy to the firm or a subsidy to the worker (a public system of training or education, for example) is a superior policy instrument in terms of welfare. Alternative policy measures such as marginal cost pricing can also be studied and evaluated.

8. Concluding Remarks

In this paper we have studied the provision of firm-sponsored general training when worker's career concerns are taken into account. A crucial assumption has been that a higher level of general training in the first period shifts the balance of power in favour of workers and thus increases the latter's bargaining power vis-à-vis the employer in the next period (in this sense, power can be interpreted here as the wage premium associated with an additional unit of general human capital). In this framework, it has been shown that a

[14] A notable exception here is Moen and Rosen (2002).

higher investment in training raises the worker's incentives to contribute effort in the first period. The employer takes this positive effect into account and is thus willing to invest in general training with higher intensity. If the positive effect of training on worker's incentives is strong enough, then equilibrium can involve overinvestment in general human capital. In particular, the case of overinvestment can arise for intermediate values of worker's bargaining power. In this context, initial increases in worker's power up to a threshold value may strengthen the employer's investment incentives and enhance welfare. These results show that predictions commonly found in previous studies (involving underinvestment in general human capital along with a negative impact of worker's power on investment and welfare) might not hold in an enriched framework that takes workers' implicit incentives into account.

References

- Acemoglu, Daron and Jörn-Steffen Pischke 1998, Why Do Firms Train? Theory and Evidence, Quarterly Journal of Economics 113, 1, 79-119.

- Acemoglu, Daron and Jörn-Steffen Pischke 1999a, Beyond Becker: Training in Imperfect Labor Markets, The Economic Journal 109 (453), F112-F142.

- Acemoglu, Daron and Jörn-Steffen Pischke 1999b, The Structure of Wages and Investment in General Training, Journal of Political Economy 107, 3, 539-572.

- Acemoglu, Daron and Jörn-Steffen Pischke 2003, Minimum Wages and On-the-Job Training, in Polachek S.W. (ed.), Worker Well-Being and Public Policy, Research in Labor Economics, Volume 22, 159-202, JAI Press.

- Andersson, Fredrik 2002, Career Concerns, Contracts and Effort Distortions, Journal of Labor Economics, 20, 1, 42-58.

- Auriol, Emmanuelle, Guido Friebel and Lambros Pechlivanos 2002, Career Concerns in Teams, Journal of Labor Economics 20, 1, 289-307.

- Autor, David 1998, Why do temporary Help Firms Provide Free General Skills Training?, Mimeo, MIT.

- Balmaceda, Felipe 2005, Firm-Sponsored General Training, Journal of Labor Economics 23, 115-133.

- Balmaceda, Felipe 2008, Firm-Provided Training and Labor Market Policies, Centro de Economia Applicada, Discussion paper No 252.

- Becker, Gary 1964, Human Capital, University of Chicago Press.

- Booth, Alison and Gylfi Zoega 2000, Why Do Firms Invest in General Training? 'Good' Firms and 'Bad' Firms as a Source of Monopsony Power, CEPR Discussion Paper No 2536.

- Booth, Alison and Gylfi Zoega 2004, Is Wage Compression a Necessary Condition for Firm-Financed General Training?, Oxford Economic Papers, 56, 1, 88-97.

- Booth, Alison and Mark Bryan 2002, Who Pays for General Training? New Evidence for British Men and Women, IZA Discussion Paper No 486.

- Chang, Chun and Yijiang Wang 1996, Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited, Journal of Labor Economics 14, 505-519.

- Dewatripont, Mathias, Ian Jewitt and Jean Tirole 1999, The Economics of Career Concerns, Part I: Comparing Information Structures, Review of economic Studies 66, 183-198.

- Fama, Eugene 1980, Agency problems and the Theory of the Firm, Journal of Political Economy 88, 288-307.

- Gersbach, Hans and Armin Schmultzer 2003, A Product Market Theory of Worker Training, CEPR Discussion Paper No 3940.

- Gibbons, Robert and Kevin Murphy 1992, Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence, Journal of Political Economy 100, 3, 468-505.

- Gindin, Sam 1998, Socialism 'with Sober Senses': Developing Workers' Capacities, The Socialist Register, 75-101.

- Holmstrom, Bengt 1982, Managerial Incentive Problems: A Dynamic Perspective, in Essays in Economics and Management in Honour of Lars Wahlbeck, Helsinki, Finland, Swedish School of Economics. Reprinted in Review of Economic Studies 96 (1999), 169-182.

- Katz, Eliakim and Andrian Ziderman 1990, Investment in General Training: The Role of Information and Labor Mobility, Economic Journal 100, 1147-1158.

- Krueger, Alan 1993, How Computers Have Changed the Wage Structure: Evidence from Micodata 1984-1989, Quarterly Journal of Economics 108, 33-60.

- Leuven, Edwin, Hessel Oosterbeek, Randolph Sloof and Chris van Klaveren 2005, Worker Reciprocity and Employer Investment in Training, Economica, 72, 285, 137-149.

- Marglin, Stephen 1974, The Origins and Functions of Hierarchy in Capitalist Production, Review of Radical Political Economics 6, 2, 60-112.

- Moen, Espen and Asa Rosen 2004, Does Poaching Distort Training?, Review of Economic Studies 71, 1143-1162.

- Ortega, Jaime 2003, Power in the Firm and Managerial Career Concerns, Journal of Economics and Management Strategy, 12, 1, 1-29.

- Owan, Hideo 2004, Promotion, Turnover, Earnings and Firm-Sponsored Training, Journal of Labor Economics, 22, 4, 955-978.

- Pfeffer, Jeffrey 1998, The Human Equation: Building Profits by Putting People First, Harvard University Press.

- Wright, Erik Olin 2000, Workers Power, Capitalist Interests and Class Compromise, American Journal of Sociology, 105, 4, 957-1002.

- Wright, Erik Olin 2004, Beneficial Constraints: Beneficial for Whom?, Socio-Economic Review, 2, 407-414.

Appendix A.

We have: $I^* = I_H \Leftrightarrow m_0 \leq R - \frac{2\beta\delta+\theta}{\beta\delta} I_H \Leftrightarrow \frac{1-k}{\delta}\frac{\beta\delta+\theta}{\beta\delta} \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0$ (A1)

The condition (A1) can hold only if $m_0 \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$. In this case:

$(A1) \Leftrightarrow \beta(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0)\delta^2 - \beta(1-k)\delta - \theta(1-k) \geq 0$ or, equivalently: $\delta \geq \delta_2$

Also, we have:

$$I^* = \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0) \Leftrightarrow R - \frac{2\beta\delta+\theta}{\beta\delta} I_H \leq m_0 \leq R \Leftrightarrow$$

$$\Leftrightarrow \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{1-k}{\delta}\frac{\beta\delta+\theta}{\beta\delta} \leq m_0 \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} + \frac{1-k}{\delta} \quad (A2)$$

(i) $m_0 \geq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{1-k}{\delta}\frac{\beta\delta+\theta}{\beta\delta} \Leftrightarrow \frac{1-k}{\delta}\frac{\beta\delta+\theta}{\beta\delta} \geq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0$ (A3)

- If $m_0 \geq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then (A3) holds for all $\delta \geq 0$

- If $m_0 \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then: (A3) $\Leftrightarrow \delta \leq \delta_2$

(ii) $m_0 \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} + \frac{1-k}{\delta} \Leftrightarrow \frac{1-k}{\delta} \geq m_0 - \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ (A4)

- If $m_0 \geq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then: (A4) $\Leftrightarrow \delta \leq \delta_0$

- If $m_0 \leq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ then (A4) holds for all $\delta \geq 0$

Finally: $I^*=0 \Leftrightarrow m_0 \geq R \Leftrightarrow \frac{1-k}{\delta} \leq m_0 - \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ (A5)

- If $m_0 \geq \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ then: (A5) $\Leftrightarrow \delta \geq \delta_0$

- If $m_0 \leq \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ then (A5) cannot hold.

The above results are summarized in the main text.

Appendix B.

For $m_0 \leq \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ we have: $I^* = \frac{\beta\delta}{2\beta\delta + \theta}(R - m_0)$ for $\delta \leq \delta_2$

$$\frac{\partial I^*}{\partial \delta} = \frac{\beta}{(2\beta\delta + \theta)^2}[\theta(\frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - m_0) - 2\beta(1-k)] > 0 \Leftrightarrow m_0 < \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - 2\hat{I} \quad \text{(B1)}$$

$$\frac{\partial^2 I^*}{\partial \delta^2} = \frac{-4\beta^2}{(2\beta\delta + \theta)^3}[\theta(\frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - m_0) - 2\beta(1-k)] > 0 \Leftrightarrow m_0 > \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - 2\hat{I} \quad \text{(B2)}$$

The conditions (B1) and (B2) yield cases 1a, 1b in Proposition 2.

For $m_0 \geq \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ we have:

$$I^* = \frac{\beta\delta}{2\beta\delta + \theta}(R - m_0) \text{ for } \delta \leq \delta_0$$

$$\frac{\partial I^*}{\partial \delta} = \frac{\beta}{(2\beta\delta + \theta)^2}[\theta(\frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - m_0) - 2\beta(1-k)] < 0 \quad \text{(B3)}$$

$$\frac{\partial^2 I^*}{\partial \delta^2} = \frac{-4\beta^2}{(2\beta\delta + \theta)^3}[\theta(\frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - m_0) - 2\beta(1-k)] > 0 \quad \text{(B4)}$$

The conditions (B3) and (B4) yield case 2 in Proposition 3.

Appendix C.

For $m_0 \leq \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}$ we have:

$$a_1^* = \frac{\beta}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}(k + \delta I^*) \text{ with } I^* = \frac{\beta\delta}{2\beta\delta + \theta}(R - m_0) \text{ for } \delta \leq \delta_2$$

$$\frac{\partial a_1^*}{\partial \delta} = \frac{\beta}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}(I^* + \delta \frac{\partial I^*}{\partial \delta}) > 0 \Leftrightarrow \frac{1-k}{\delta} \frac{\theta}{2(\beta\delta + \theta)} > m_0 - \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} \text{ which holds for all } \delta \leq \delta_2. \quad \text{(C1)}$$

Also: $$\frac{\partial^2 a_1^*}{\partial \delta^2} = \frac{\beta}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon}(2\frac{\partial I^*}{\partial \delta} + \delta \frac{\partial^2 I^*}{\partial \delta^2}) > 0 \Leftrightarrow m_0 < \frac{1-k}{\rho} \frac{h_\varepsilon}{h_0 + h_\varepsilon} - 2\hat{I} \quad \text{(C2)}$$

The conditions (C1) and (C2) yield cases 1a, 1b in Proposition 3.

Similarly, for $m_0 \geq \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}$ we have:

$$a_1^* = \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(k+\delta I^*) \text{ with } I^* = \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0) \text{ for } \delta < \delta_0$$

$$\frac{\partial a_1^*}{\partial\delta} = \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(I^* + \delta\frac{\partial I^*}{\partial\delta}) > 0 \Leftrightarrow \frac{1-k}{\delta}\frac{\theta}{2(\beta\delta+\theta)} > m_0 - \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} \Leftrightarrow \delta \leq \delta_4 \quad \text{(C3)}$$

$$\frac{\partial^2 a_1^*}{\partial\delta^2} < 0 \Leftrightarrow m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I} \text{ which holds for all } \delta < \delta_0. \quad \text{(C4)}$$

The conditions (C3) and (C4) yield case 2 in Proposition 3.

Appendix D.

We focus on case 1a to examine the possibility of overinvestment.

For $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I}$ (case 1a) we have:

$$\text{If } \delta < \delta_2 \text{ then: } I^* > I^{FB} \Leftrightarrow \frac{\beta\delta}{2\beta\delta+\theta}(R-m_0) > \frac{\beta}{\theta} \Leftrightarrow \frac{k}{\delta} < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0 - 2I^{FB} \quad \text{(D1)}$$

(i) For $m_0 > \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2I^{FB} \Leftrightarrow \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2I^{FB} < m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2\hat{I}$

(C1) cannot hold, i.e. there is underinvestment for all $\delta < \delta_2$ (and thus for all $\delta \geq 0$, since *I*(δ)* is decreasing in δ for δ>δ$_2$).

(ii) For $m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2I^{FB}$:

$$\text{(C1)} \Leftrightarrow \delta > \frac{k}{\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0 - 2I^{FB}} = \delta_5 \Leftrightarrow \delta_5 < \delta < \delta_2 \quad \text{(D2)}$$

with necessary condition: $\delta_5 < \delta_2$

$$\text{If } \delta \geq \delta_2\text{, then: } I^* > I^{FB} \Leftrightarrow \frac{1-k}{\delta} > \frac{\beta}{\theta} \Leftrightarrow \delta < \frac{\theta(1-k)}{\beta} = \delta_6 \Leftrightarrow \delta_2 \leq \delta < \delta_6 \quad \text{(D3)}$$

with necessary condition:

$$\delta_2 < \delta_6 \Leftrightarrow \frac{\beta(1-k) + \sqrt{\beta^2(1-k)^2 + 4\beta\theta(1-k)(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0)}}{2\beta(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0)} < \frac{\theta(1-k)}{\beta}$$

$$m_0 < \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{2-k}{1-k}I^{FB}$$, i.e:

$$m_0 < \min\{\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - 2I^{FB}, \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{2-k}{1-k}I^{FB}\} = \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - \frac{2-k}{1-k}I^{FB} \qquad (D4)$$

Finally, note that $\delta_2<\delta_6$ implies $\delta_5 < \delta_2$.

The conditions (D1) to (D4) yield the results summarized in Proposition 4.

Appendix E.

Note, first, that the welfare maximization problem is not (globally) concave with respect to δ. Therefore, we must find all (interior and boundary) solutions of first-order conditions and compare the value of the objective function at different solutions to find the maximum.

The Lagrangian is:

$$L = W^*(\delta) + \lambda(\delta^* - \delta) = (1+\beta)m_0 + a_1^* - \frac{\rho}{2}(a_1^*)^2 + \beta I^* - \frac{\theta}{2}(I^*)^2 + \lambda(\delta^* - \delta)$$

The Kuhn-Tucker first-order conditions are:

$$\frac{\partial L}{\partial \delta} = \frac{\partial W^*}{\partial \delta} - \lambda = \frac{\partial \alpha_1^*}{\partial \delta}(1-\rho a_1^*) + \frac{\partial I^*}{\partial \delta}(\beta - \theta I^*) - \lambda =$$

$$= \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}(I^* + \delta\frac{\partial I^*}{\partial \delta})\left[1 - \beta\frac{h_\varepsilon}{h_0+h_\varepsilon}(k+\delta I^*)\right] + \frac{\partial I^*}{\partial \delta}(\beta - \theta I^*) - \lambda \le 0, \; \frac{\partial L}{\partial \delta}\delta = 0 \qquad (E1)$$

$$\frac{\partial L}{\partial \lambda} = \delta^* - \delta \ge 0, \; \frac{\partial L}{\partial \lambda}\lambda = 0 \qquad (E2)$$

We check if $\delta=0$ can be a solution of FOCs. For $\delta=0$ the conditions (E2) imply $\lambda=0$. In this case, (E1) requires:

$$(E1) \Leftrightarrow \frac{\partial L}{\partial \delta} = \frac{\partial W^*}{\partial \delta} \le 0 \overset{\delta=0}{\Leftrightarrow} \frac{\beta}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}\hat{I}(1 - \beta k\frac{h_\varepsilon}{h_0+h_\varepsilon}) + (\beta - \theta\hat{I})\frac{\partial I^*}{\partial \delta}/_{\delta=o} \le 0 \qquad (E3)$$

We have: $\frac{\partial I^*}{\partial \delta} = \frac{\beta}{(2\beta\delta+\theta)^2}\left[\theta(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0) - 2\beta(1-k)\right]$

$$\Rightarrow \frac{\partial I^*}{\partial \delta}/_{\delta=o} = I^{FB}(\frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon} - m_0 - 2\hat{I}) \qquad (E4)$$

$$(E3) \overset{(E4)}{\Leftrightarrow} m_0 \ge \frac{1-k}{\rho}\frac{h_\varepsilon}{h_0+h_\varepsilon}\left[1 + \frac{\beta(1-\beta k\frac{h_\varepsilon}{h_0+h_\varepsilon})}{\beta - \theta\hat{I}}\right] - 2\hat{I}.$$

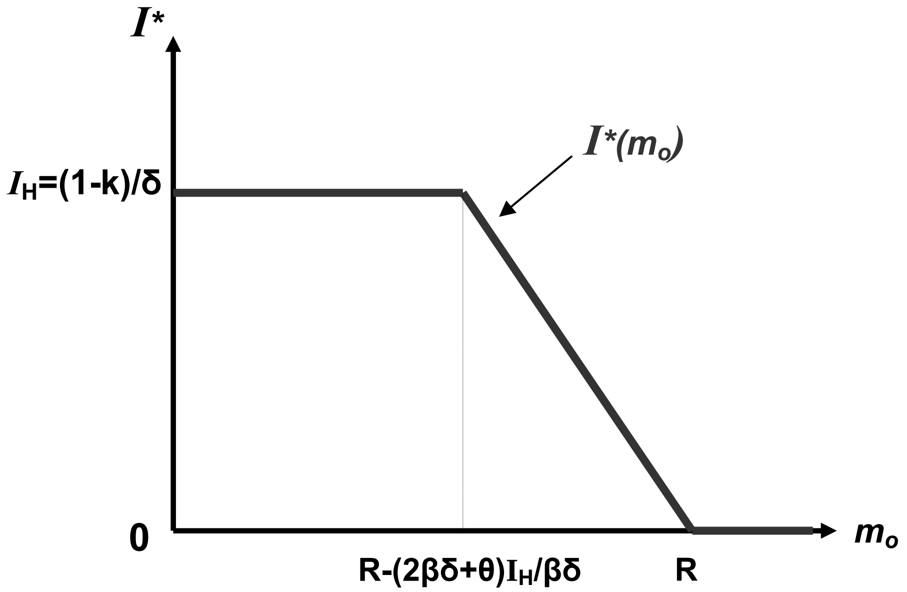

Figure 1

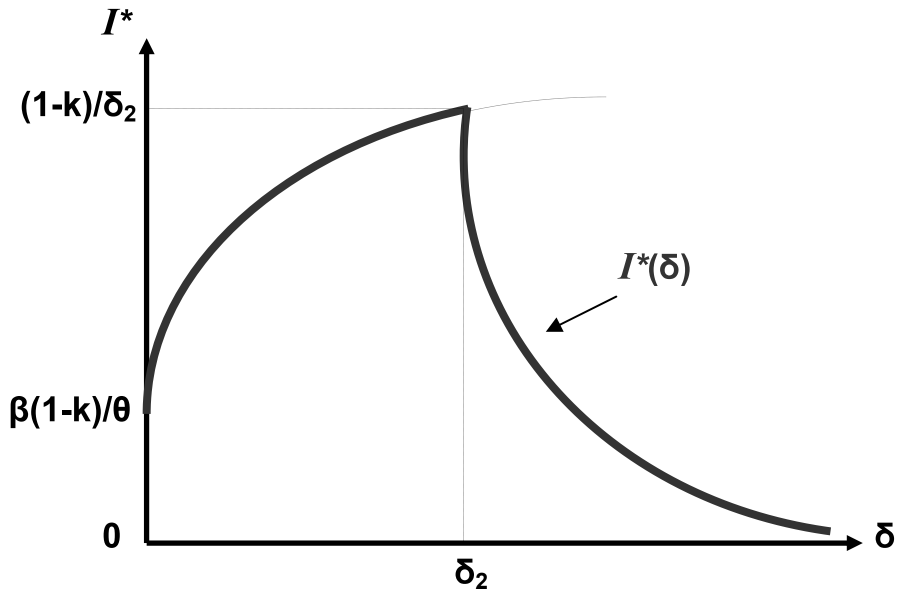

Figure 2: Case 1a

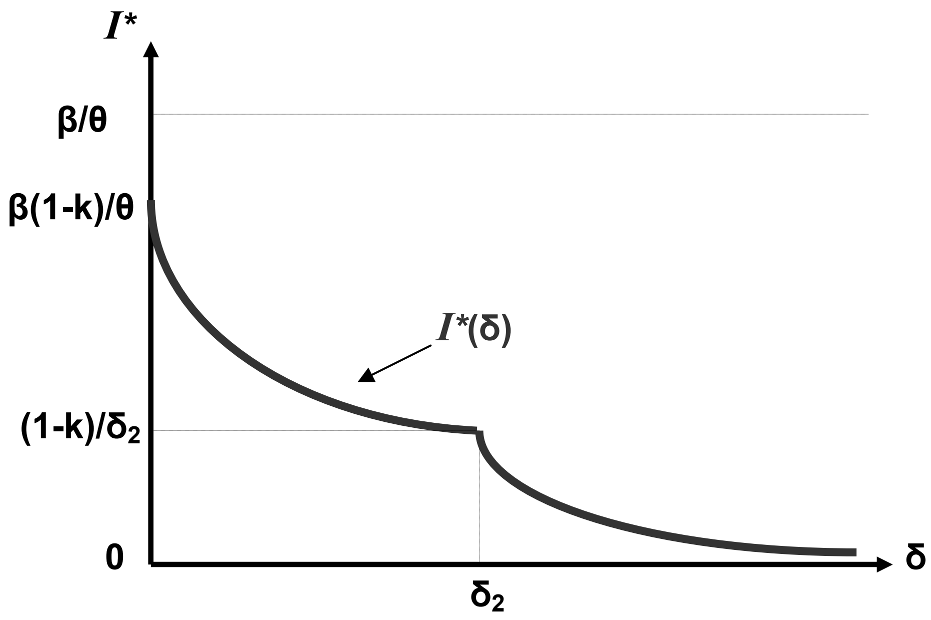

Figure 3: Case 1b

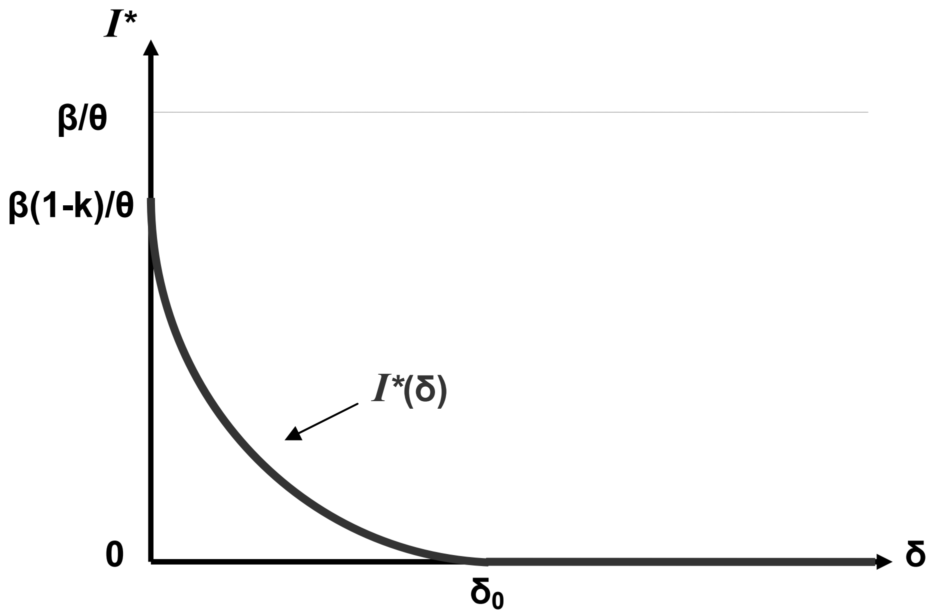

Figure 4: Case 2

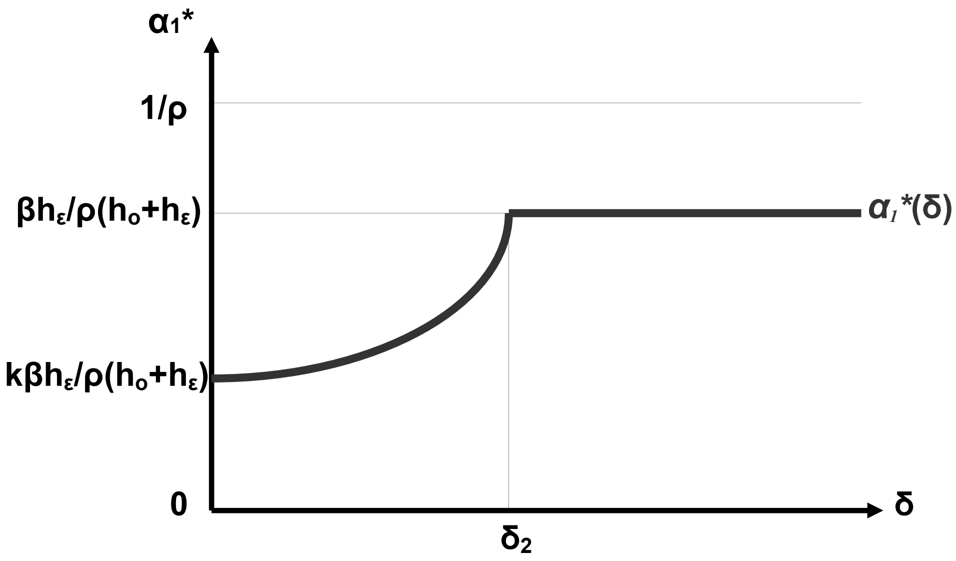

Figure 5: Case 1a

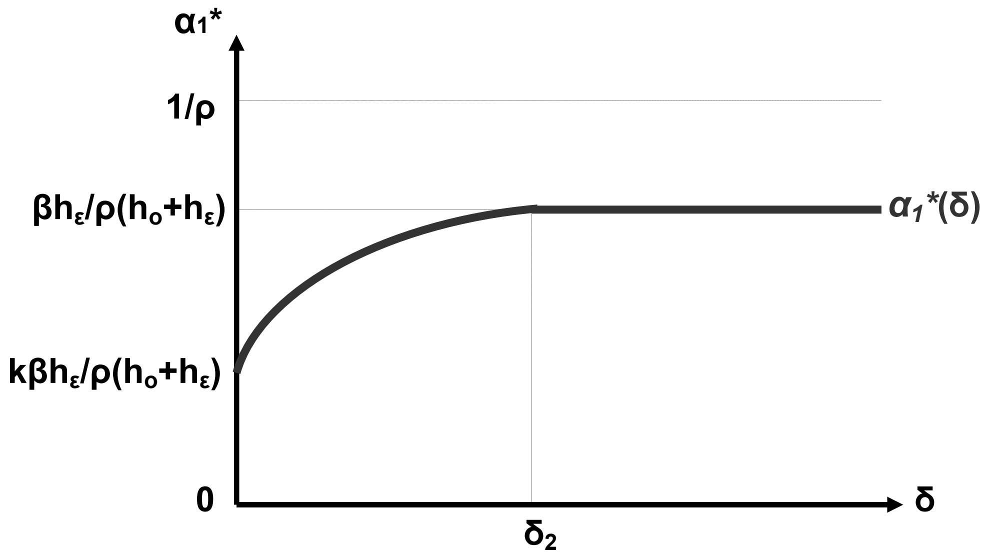

Figure 6: Case 1b

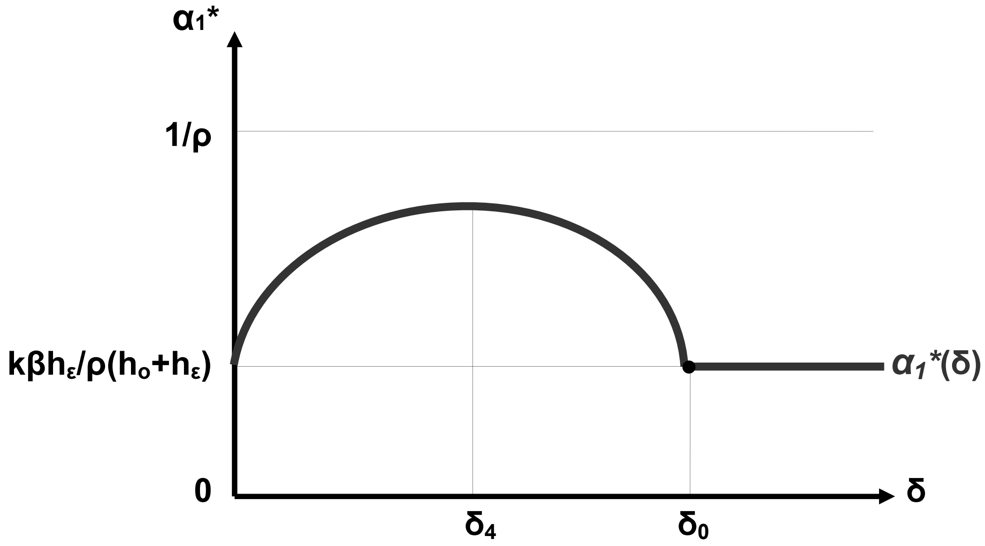

Figure 7: Case 2

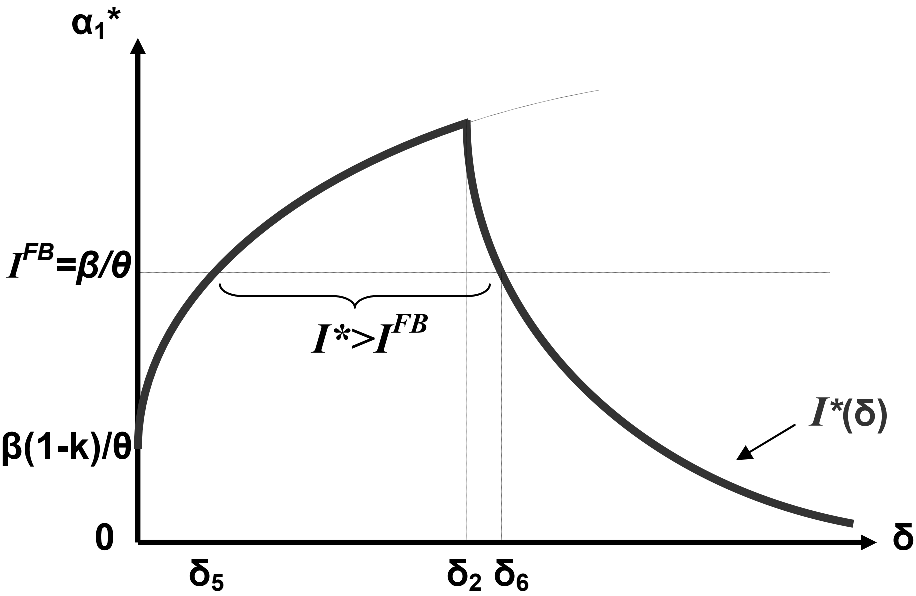

Figure 8: Overinvestment in General Training