Career Concerns and The Provision of General Training

Christos Bilanakos
PhD Candidate, Athens University of Economics and Business
Department of International and European Economic Studies

April 2009

Abstract
This paper builds on the argument that the provision of general training increases the worker’s bargaining power vis-à-vis the employer. In a model incorporating career concerns, it is then shown that the worker’s implicit incentives to provide effort increase with the level of acquired general skills. The employer takes this effect into account and is thus willing to increase the intensity of investment in its employee’s general training. When this positive effect of training on worker’s incentives is strong enough, the equilibrium outcome may even involve overinvestment in general training relative to the first-best. In this framework, it is also shown that sharper increases in worker’s power associated with general training can strengthen the employer’s investment incentives and have beneficial effects on social welfare. The paper also discusses and evaluates alternative policy measures to alleviate the associated market inefficiencies.

1. Introduction

1.1 Motivation
The subject of this paper is the provision of firm-sponsored general training to workers. The process of economic restructuring that began three decades ago in developed countries has generated a number of new requirements: The growth of service industries and occupations implies that firms increasingly need workers with extended and heterogeneous skills, so that the latter may be able to keep up with new communication and information technologies, to engage in problem-solving and information-processing as well as to work efficiently in teams (Green, 2005).

The efficiency benefits associated with the provision of general training have been widely recognized in both the economic and management literature. In the context of the latter, Pfeffer (1998, p.85-90) provides compelling evidence in favor of ‘people-centered management’: This involves (among else) the provision of extensive training
which, however, should not be directed towards the acquisition of specialist skills but rather to ‘general competence and organizational culture’. In other words, Pfeffer recommends – along with other organizational practices – the provision of general (rather than firm-specific) skills to workers. He argues that this kind of general competence allows the efficient decentralization of decision-making within the organization, the formation of self-managed workteams and the establishment of processes which encourage all members of the firm to generate and share information and tacit knowledge. Similarly, the acquisition of general skills allows all members of the organization to engage in participatory planning. The superiority of intra-firm generalized participation in decision-making has been recently suggested by Adaman and Devine (2002): empowerment renders possible the more efficient mobilization of explicit or tacit knowledge and helps to take into account a wider range of criteria with respect to the allocation of resources. The competence theory of the firm (see Tsoukas, 1996) focuses on the social and distributed character of organizational knowledge and conceives the firm as “a discursive practice: a form of life, a community in which individuals come to share an unarticulated background of common understandings”. This context also encourages workers to undertake innovative or entrepreneurial activities, thus raising productivity and firm performance. All these efficiency benefits (as well as the standard productivity benefits usually recognized in the associated literature) presuppose the provision of general training to workers-managers employed by the firm.

This wide recognition of the increasing importance of employees’ general human capital for the performance of an organization seems to contrast the actual pattern of skill provision in many cases, thus indicating the potential existence of significant inefficiencies. In particular, empirical evidence (see e.g. Green, 2005) reveals a rather contradictory pattern of skill development: On the one hand, an increase in the average level of skills is indeed observed in economically developed countries; at the same time, however, a process of polarization seems to be under way, since the increasing demand for highly-skilled jobs is accompanied by an increasing demand for unskilled jobs and, in any case, the supply of skilled jobs increases considerably faster than the demand for qualified workers (this is starkly reflected in the increasing proportion of college graduates working in jobs that do not require college-level skills). The hypothesis of organizational members’ empowerment and generalized
involvement in workplace decisions is clearly not verified, since the discretionary power of a managerial minority has been intensified and the choice of ‘hard human resource management’ or ‘low-road strategies’ has been reinforced in recent years (involving downsizing and reengineering, lack of employment security or implicit contracts etc). Furthermore, field evidence shows that firms in the United States and the United Kingdom underinvest in their employees’ training relatively, for example, to Japanese firms (more precisely, Pfeffer underlines that Japanese firms provide 364 hours of training on average to new workers in the first six months of employment, whereas US firms only provide 42 hours). At the same time, training in US and UK firms is not only inadequate but also tends to focus on the provision job-specific rather than general skills – contrary to the efficiency requirements underlined above.

The inefficiently low levels of firm-sponsored general training have often been attributed to employers’ short-termism. The increasing competitive pressures (associated with the global expansion of trade and the internationalization of manufacturing activities) lead many firms to focus on input cost minimization and makes them unwilling to engage in costly investments that only yield middle-term or long-term benefits. Indeed, the higher level of employment security in Japanese firms (relative to US firms) seems to indicate a longer time horizon which may explain the significantly higher investment in their employees’ human capital. In this context, a recommendation often articulated (see e.g. Streeck, 1997) concerns the imposition of institutional constraints (which may include binding collective agreements) that would function as commitment devices and would help to sustain mutual benefits associated with the provision of training.

This paper suggests that (apart form short-termism) dynamic power considerations should also be taken into account in order to explain the patterns of underprovision (or overprovision) of general training within firms. In particular, employers may have an additional reason not to pay for their employees’ general training if the acquisition of such general capacities decisively shifts the balance of bargaining power in favor of workers (see the related argument in Wright, 2004). This means that workers equipped with considerable general capacities may receive attractive wage offers from

---

1 See Green (2005, Ch. 1) or Pfeffer (1998) for an inspection of this evidence; see also O’Reilly and Pfeffer (2000).
other potential employers or may even be able to form their own firm. A well-known example of the latter case is the Intel Corporation: Intel was founded in 1968 by two former executive members of the Fairchild Semiconductor Corporation who left their employer to form their own company. A significant number of other Fairchild employees-managers also left to form or to participate in other Silicon Valley companies. The loss of iconic executives resulted in a generalized sense of demoralization within the Fairchild, prompting the entire exodus of employees and a complete drain-out of talent. The Fairchild Corporation suffered huge losses during that period. More generally, it can be said that workers who receive extensive general training strengthen their bargaining position vis-à-vis their current employer and thus can extract a higher proportion of the produced surplus. The employer, in turn, anticipates this effect of general human capital on the balance of class forces and may thus rationally provide inefficiently low levels of general training in order to secure his power in the long-run. The model developed in this paper studies the validity of this argument and establishes conditions under which this prediction about the employer’s incentives to invest in general human capital is verified or not.

In a broader context, Marglin (1974) has similarly argued that power – rather than efficiency – considerations may explain the pattern of technological progress and work organization practices within the workplace. In his influential paper on ‘The Origins of Hierarchy’, Marglin questions Smith’s famous pin-making arguments about the (technical) efficiency of the minute division of labor and extensive specialization. A combination of theoretical arguments and historical evidence is used to support the view that hierarchical work organization (and the associated extensive specialization) was strategically chosen by employers in order to secure an essential role in the production process as integrators of workers’ individual efforts into marketable products (p.70-71) – i.e. a kind of divide-and-conquer strategy. In other words, Marglin argues that excessive specialization reinforces the capital owner’s legitimacy as the ultimate locus of decision-making within the organization. This view is in line with the argument that employers are less willing to pay for the provision of general human capital (than for job-specific skills) to their employees; at the same time, it sheds more light to this argument by claiming that employers may seek to retain these general organizational capacities for themselves to secure their class power and control in the long-run.
In a similar spirit, Gindin’s (1998) discussion on the possibility of transition to socialism focuses on the fact that employers are generally more willing to make concessions at the level of wages than to pay for their employees’ general training. He explains this by arguing that the widespread development of workers’ general capacities for creative planning and decision-making generates a far more decisive shift in the balance of class power than a direct increase of their purchasing power. As Gindin puts it: “What workers give up in selling their labor are precisely the kind of capacities and potentials which are absolutely fundamental to one day building a different kind of society; the capacities for doing, creating, planning, executing” (p.79). To the extent that the employer monopolizes planning and work organization activities, he remains the legitimate and, indeed, the indispensable organizer or integrator of isolated workers’ individual contributions. The lack of general skills and capacities for self-management implies that the costs of transition to the socialist alternative are prohibitively high. In this framework, Gindin’s recommendation is a kind of ‘movement unionism’ going beyond mere wage bargaining to promote the “democratic development of workers’ capacities” through programs of general training in which workers engage in participatory structures of decision-making and get used to identifying themselves as producers and political agents. In fact, this recommendation is much in line with the basic prescription of the orthodox economic literature (to which we turn in the following subsection): workers – facing their employers’ unwillingness to sponsor their general training – will optimally pay themselves to acquire these general skills and finance this kind of investment through wage cuts in early periods of their careers. The basic results of the literature associated with the theory of general training are now briefly presented.

1.2 A Brief Review of the Literature

Training can be formally defined as investment in human capital, typically leading to subsequent increases in productivity. In particular, general training refers to the acquisition of skills that increase the worker’s productivity even if the latter quits her current employer to move to another firm (or form her own firm). The cost of this investment can be paid either by the employer (firm-sponsored general training) or by the worker herself.
In his seminal work on the provision of general human capital, Becker (1964) assumes a competitive labor market (in the sense that there exist many identical employers who can use the general skills acquired by the worker and, furthermore, the worker does not incur any costs when moving to another firm). The current employer initially chooses (and pays for) 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 general skills acquired in the first place). In this setting, Becker’s prediction is that the firm will make zero investment in general training: The current employer cannot recoup the costs of investment later, since the worker reaps all productivity benefits due to the competitive wage offers made in the labor market. In fact, the general conclusion of underinvestment in general training had already been articulated earlier by Pigou. However, Becker goes on to predict that the worker will optimally bear herself the cost of investment in general human capital: Since the worker is the full residual claimant of productivity benefits associated with acquired skills, she faces the right incentives to invest in her own general training. Therefore, the outcome will be efficient (the level of training will be equal to the first-best) and the worker will finance the cost of training investment through a wage cut in early periods of employment (under the assumption that she faces no credit constraints). This prediction fits well the fact that many workers receive lower wages at early stages of their career (during which they are subject to a process of acquisition of general skills).

However, the prediction of zero firm-sponsored general training is in contrast with considerable evidence showing that firms do provide general skills to their employees (see e.g. Krueger, 1993; Autor, 1998; Acemoglu and Pischke 1998, 1999a, Booth and Bryan, 2002). Consequently, a branch of the recent literature has been studying the conditions under which firms are indeed willing to pay for their workers’ general human capital. We rely on Acemoglu and Pischke’s (1998, 1999a, 1999b, 2003)
influential work to state the following widely recognized conditions for firm-sponsored investment in general training:

(i) Frictional (imperfect) labor market: The worker’s productivity must exceed her outside wage (i.e. the wage she receives if she leaves her current employer). This is a deviation from Becker’s assumption of a perfectly competitive labor market which implied that the worker receives her full marginal product when moving to another firm.

Similarly, the bargaining power of the firm vis-à-vis the worker must be nonzero, so that the employer can reap at least a share of the joint surplus and thus have an incentive to invest in general training.

(ii) Wage Compression: The marginal effect of training on productivity must exceed the marginal effect of training on worker’s wage.

Under these assumptions, the basic results of Acemoglu-Pishke’s model can be summarized as follows:

(1) The firm has an incentive to make a positive investment in general training.

(2) However, this investment will still be inefficiently low (i.e. the model predicts underinvestment in general training). In particular, the employer’s incentives to invest in general skills are too weak because:

(a) A higher level of general training increases the worker’s outside wage and thus the current employer has to pay a higher wage in order to keep the worker from moving to another firm.

(b) The worker’s (positive level of) bargaining power allows her to reap a part of productivity benefits associated with training.

(3) Any increase in the worker’s bargaining power implies a lower level of firm-sponsored investment in general skills and reduces social welfare.

(4) As a result, a government subsidy to the firm in order to strengthen the latter’s investment incentives seems to be the appropriate policy measure to mitigate the associated market failure.

---

2 Similar arguments have also been raised by Katz and Ziderman (1990) or Chang and Wang (1996), who rely on the assumption that investment in training is not observable by other potential employers. More recent contributions on the subject of general training include Booth and Zoega (2000, 2001), Gersbach and Schnultzer (2003) or Balmaceda (2005, 2008); see Asplund (2004) for a review of the related theoretical and empirical literature.
The model developed below is also closely related to the literature on career concerns. This literature starts with Fama (1980), who argued that workers-managers’ career concerns provide them with implicit incentives that may be sufficient to motivate efficient levels of effort even in the absence of explicit agency contracts. Building on this idea, Holmstrom (1982; 1999) has provided a formal framework on the subject of career concerns: in Holmstrom’s multi-period model, there is symmetric (but imperfect) information about the worker’s ability and, at the same time, the worker’s effort is nonverifiable by the employer in a court of law (i.e. there is moral hazard). The worker’s performance in each period depends on both her ability and her effort contribution. The labor market is competitive, i.e. there are many identical firms (with symmetric information) which can make competitive wage offers to attract the worker (as in Becker’s seminal paper on general training). The wage is offered at the beginning of each period since (by assumption) the firm is not able to commit to a long-term contract. In this framework, the wage offer made in each period will be contingent on the expected output of that period given the history of past output realizations. The expected output of each period will depend, in turn, on market beliefs about the worker’s ability. This means that the worker has an incentive to work hard in early periods of her career in order to raise market beliefs about her ability and thus increase the expected output and the associated wage offers received in future periods.

Further contributions on the issue of career concerns include Gibbons and Murphy (1992) or Andersson (2002) who study the interaction between career concerns (as a source of implicit incentives) and explicit incentives optimally designed in an agency contract. This issue is also examined by Auriol, Friebel and Pechlivanos (2002) in the context of teams. The basic career concerns model has been extended in multiple directions by Dewatripont et al (1999). Finally, it should be noted that the issue of power in the context of career concerns (which is raised in the model developed below) has also been studied by Ortega (2003) in a model of team production. However, Ortega’s conception 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 assumption made here.
1.3 Basic Results

The two-period model developed in the next section studies the firm’s incentives to provide general training while taking into account the worker’s implicit incentives to provide effort (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 based on 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 the worker, the current employer and the market). 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 in training (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 employer’s incentives to invest in general skills – due to the anticipated positive impact of such investment on worker’s incentives. This is in sharp contrast with the usual prediction that any increase in the worker’s power has a negative impact on firm’s investment incentives. Turning to welfare implications, the focus is on the derivation of a socially optimal level of worker’s power. In this respect, we identify parameter intervals in which increases of 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 up to a critical value enhance social welfare, whereas further increases above that critical level are 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 the equilibrium in our second-best environment. The positive and normative implications of the equilibrium outcome are discussed in Section 5. Finally, Section 6 discusses possible extensions and policy implications and Section 7 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 a technology specified below. The worker contributes effort $a_t$ in each period ($t = 1, 2$) to the production of the consumption good. Finally, we denote by I the level of investment in general skills. The cost of this investment is borne by the employer in the first period, whereas the benefits appear 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 $\varepsilon_t$ is an idiosyncratic productivity shock (noise). The level of effort $a_t$ is nonverifiable - i.e. there is moral hazard in the model. The output $x_t$ in each period is observable by all agents in the economy. Furthermore, we assume that $n$ and $\varepsilon_t$ are normally distributed:

- $n \sim N(m_0, \frac{1}{h_0})$: Prior Beliefs about the worker’s ability. These beliefs are symmetric – i.e. they are shared by all agents in the economy.
- $\varepsilon_t \sim N(0, \frac{1}{h_\varepsilon})$: Distribution of the productivity shock in period $t$. 

The random variables $n$, $\varepsilon_1$ and $\varepsilon_2$ are (identically and) independently distributed:

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

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

$EU_A = E\{\sum_{t=1}^{2} \beta^{t-1}[x_A^t - \frac{\rho}{2} \alpha_1^t]\} = E\{x_A^1 - \frac{\rho}{2} \alpha_1^1 + \beta(x_A^2 - \frac{\rho}{2} \alpha_2^2)\}$: Worker A’s lifetime discounted expected utility,

$EU_P = E\{x_P^1 - \frac{\theta}{2} l^2 + \beta x_P^p\}$: Employer P’s lifetime expected utility,

where $x_i^t$ denotes agent $i$’s consumption in period $t$, $i \in \{A, P\}$ and $\beta \in [0,1]$ is the (common) discount factor. The disutility of labor provided in each period is represented by the function $g(a_i) = \frac{\rho}{2} \alpha_i^2$, while the cost of investment in general training is captured by the cost function $\tau(l) = \frac{\theta}{2} l^2$.

### 3. The First-Best Problem

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:

$$\max_{\{a_i, A, P\}_i} EU_A + EU_P = E\{x_A^1 + x_P^p - \frac{\rho}{2} \alpha_1^1 - \frac{\theta}{2} l^2 + \beta(x_A^2 + x_P^p - \frac{\rho}{2} \alpha_2^2)\}$$

s.t: $x_A^1 = n + a_1 + \varepsilon_1$

$x_A^2 = n + a_2 + \varepsilon_2 + l$ : Technological Constraints (P)

$x_A^1 + x_P^p \leq x_1$

$x_A^2 + x_P^p \leq x_2$ : Resource Constraints

$a, l, x_i^t \geq 0$ : Nonnegativity Constraints

Equivalently, the above problem (P) can be written:

$$\max_{\{a_i, A, P\}_i} TS = E\{x_1 - \frac{\rho}{2} \alpha_1^1 - \frac{\theta}{2} l^2 + \beta(x_2 - \frac{\rho}{2} \alpha_2^2)\} = E\{n + a_1 + \varepsilon_1 - \frac{\rho}{2} \alpha_1^1 - \frac{\theta}{2} l^2 + \beta(n + a_2 + \varepsilon_2 - \frac{\rho}{2} \alpha_2^2)\}$$

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

The solution of this optimality problem is:
\[ a_i^{FB} = a_2^{FB} = a_1^{FB} = \frac{1}{\rho}, \quad I^{FB} = \frac{\beta}{\theta} \]  
\hspace{1cm} (1)

yielding the optimal level of welfare:

\[ W^{FB} = (1 + \beta)m_0 + (1 + \beta)[a_1^{FB} \frac{\rho}{2} (a_1^{FB})^2] + \beta I^{FB} \frac{\theta}{2} (I^{FB})^2 \]

4. Second-Best Environment: The Perfect Bayesian Equilibrium

In order to find the equilibrium in our second-best environment, we define the following 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. This wage offer is contingent on the expected output of the same period given the history of past output realizations (this implies that there exist only implicit incentives – i.e. career concerns – for the worker to provide effort, as in Holmstrom’s model). In particular, we assume the following wage structure (\(w_t\) denotes the wage offered at the beginning of period \(t\)):

\[ w_t = kE(x_t / prior), \quad k \in [0,1]. \]  

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 made in Becker’s seminal paper on general training. Of course, as it is already known by the related literature (and as it will also be made clear below) 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:

\[ w_2 = (k + \delta I)E(x_2 / x_1) \]  
\hspace{1cm} (2)
\[ w_2 = E(x_2 / x_1) \quad , \quad \text{if} \quad I \leq \frac{1-k}{\delta} \]

The assumption implicit in the second-period wage offer 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 (of course, the wage \(w_2\) cannot exceed the second-period expected output). 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, any increase in worker’s productivity is preserved even if the worker leaves her initial employer to work for a different firm. As a result, the worker’s (second-period) outside wage opportunity increases with the level of training received in the first period and thus 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 crucial question is, then, how this structure affects the worker’s incentives to provide effort and the employer’s incentives to invest in general human capital in the first place.

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

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

\[ EU_p = E\{x_1 - w_1 - \frac{\theta}{2} I^2 + \beta(x_2 - w_2)\} \]

The timing of the associated game is the following:

- **Period 1 (t=1)**
  1. The employer P chooses the level of investment in general training I.
  2. The employer P offers \( w_1 = kE(x_1 / prior) \) to the worker A.
  3. The worker A chooses her first-period effort contribution \( \alpha_1 \), and output \( x_1 \) is realized
  4. All agents update their beliefs about A’s ability given the first-period output realization \( x_1 \).

- **Period 2 (t=2)**
  1. The employer P offers \( w_2 \) according to the wage structure specified in (2)
  2. The worker A chooses her second-period effort contribution \( \alpha_2 \), 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 \((\alpha_1^*, \alpha_2^*, I^*)\), conjectures \((\tilde{\alpha}_1, \tilde{\alpha}_2)\) and beliefs about the worker’s ability such that:

(i) The worker A’s strategy \((\alpha_1^*, \alpha_2^*)\) maximizes her expected payoff

(ii) The employer P’s strategy \(I^*\) maximizes his expected payoff
(iii) Conjectures are correct: \((\tilde{a}_1, \tilde{a}_2) = (a'_1, a'_2)\)

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

We use backward induction to derive the equilibrium outcome. In the second period, the worker chooses her effort contribution \(a_2\) (given \(I, w_1, a_1, w_2\)) so as to maximize her expected payoff:

\[
\max_{a_2} EU_A = w_1 - \frac{\rho}{2} a_2^2 + \beta(w_2 - \frac{\rho}{2} a_2^2)
\]

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

\[a_2^* = 0\]  \hspace{1cm} (3)

Since the worker is motivated only by her career concerns (implicit incentives), 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 \(I, w_1, a_1\)) according to the wage structure specified in (2):

\[
w_2(x_i) = (k + \delta I)E(x_2 / x_i) = (k + \delta I)E(n + a_2^* + \varepsilon_2 + I / x_i) = (k + \delta I)[I + E(n / x_i)] ,
\]

if \(I \leq \frac{1-k}{\delta}\)

\[
w_2(x_i) = E(x_2 / x_i) = E(n + a_2^* + \varepsilon_2 + I / x_i) = I + E(n / x_i) , \text{ if } I \geq \frac{1-k}{\delta} \]  \hspace{1cm} (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 can define a new random variable:

\[z_i \equiv x_i - \tilde{a}_i = n + a_i + \varepsilon_i - \tilde{a}_i = n + \varepsilon_i,\] which is also normally distributed since \(n\) and \(\varepsilon_i\) are normally distributed by assumption (note that \(a_i = \tilde{a}_i\) along the equilibrium path).

Therefore, we can use the following normal updating formula associated with two normally distributed random variables \(y_1, y_2\):

\[
y_1 / y_2 \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}\right), \text{ where:} \]

\[
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}\right), \text{ and } r = \frac{\sigma_{12}}{\sigma_1 \sigma_2}.
\]
For \( y_1 = n, y_2 = z \) the above formula yields:

\[
n / z_1 \sim N \left( \frac{h_y m_0 + h_z z_1}{h_0 + h_z}, \frac{1}{h_0 + h_z} \right)
\]  

(5)

We have: 

\[
E(n / z_1) = \frac{h_y m_0 + h_z z_1}{h_0 + h_z} \Rightarrow E(n / x_i) = \frac{h_y m_0 + h_z (x_i - a_i)}{h_0 + h_z}
\]  

(6)

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

\[
w_2(x_i) = (k + \delta I) \left[ I + \frac{h_y m_0 + h_z (x_i - a_i)}{h_0 + h_z} \right], \text{ if } I \leq \frac{1-k}{\delta}
\]

\[
w_2(x_i) = I + \frac{h_y m_0 + h_z (x_i - a_i)}{h_0 + h_z}, \text{ if } I \geq \frac{1-k}{\delta}
\]

(7)

Anticipating \( w_2(x_i) \) 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:

\[
\max_{\{a_1\}} \text{EU}_A = w_1 - \frac{\rho}{2} \alpha_1^2 + \beta [Ew_2(x_i) - \frac{\rho}{2} (\alpha_2^*)^2] =
\]

\[
= w_1 - \frac{\rho}{2} \alpha_1^2 + \beta (k + \delta I) \left[ I + E \left( \frac{h_y m_0 + h_z (n + a_i + e_i - a_i)}{h_0 + h_z} \right) \right], \text{ if } I \leq \frac{1-k}{\delta}
\]

\[
= w_1 - \frac{\rho}{2} \alpha_1^2 + \beta \left[ I + E \left( \frac{h_y m_0 + h_z (n + a_i + e_i - a_i)}{h_0 + h_z} \right) \right], \text{ if } I \geq \frac{1-k}{\delta}
\]

For \( I \leq \frac{1-k}{\delta} \), the first-order conditions of this maximization problem are:

\[
\frac{\partial \text{EU}_A}{\partial a_1} = \beta (k + \delta I) \frac{h_z}{h_0 + h_z} - \rho \alpha_1 \leq 0 , \quad \frac{\partial \text{EU}_A}{\partial a_1} a_1 = 0 . \text{ These conditions yield the solution:}
\]

\[
\alpha_1 = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (k + \delta I). \text{ We solve the problem in the same way for } I \geq \frac{1-k}{\delta} \text{ and thus get the complete solution:}
\]

\[
\alpha_1 = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (k + \delta I), \text{ if } I \leq \frac{1-k}{\delta}
\]

\[
\alpha_1 = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z}, \text{ if } I \geq \frac{1-k}{\delta}
\]

(8)
Note that (for $I \leq \frac{1-k}{\delta}$) the worker’s implicit incentives (career concerns) 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:

\[ \frac{\partial \alpha_i}{\partial I} = \frac{\beta}{\rho} \frac{h_{x}}{h_{0} + h_{\epsilon}} \delta \]

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. 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 (more than six dozen courses in all). 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). It should also be emphasized that this reciprocal effect of firm-sponsored training on worker’s incentives does not stem from 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 rational choice of $\alpha_i$) according to the wage structure specified above:

\[ w_1 = kE(x_i / prior) = k(m_0 + a_i) \]  \hspace{1cm} (9)

where $\alpha_i$ is given in (8).

At the first stage of the game, the employer $P$ chooses the level of investment in general training $I$ (anticipating $w_1, a_1, w_2, a_2$ as given in (9), (8), (7), (3) respectively) so as to maximize her lifetime expected payoff:

\[ \max_{[I]} EU_p = E\{x_i - w_i - \frac{\theta}{2}I^2 + \beta(x_2 - w_2)\} \]

\[ = (1-k)(m_0 + a_i) - \frac{\theta}{2}I^2 + \beta(m_0 + I)(1-k - \delta I), \text{ if } 0 \leq I \leq \frac{1-k}{\delta} \]
\[= (1-k)(m_0 + a_i) - \frac{\theta}{2} I^2, \text{ if } I \geq \frac{1-k}{\delta} \]

We separately solve each of these two sub-problems:

For \( I \geq \frac{1-k}{\delta} \), we have: \( \frac{\partial EU_p}{\partial I} = -\theta I < 0 \). In other words, the employer’s investment will never be higher than \((1-k)/\delta\). Consequently, we focus on the interval \( 0 \leq I \leq \frac{1-k}{\delta} \):

\[
\max_{(I)} EU_p = (1-k)(m_0 + a_i) - \frac{\theta}{2} I^2 + \beta(m_0 + I)(1-k - \delta I)
\]

s.t. \( 0 \leq I \leq \frac{1-k}{\delta} \)

The objective function \( EU_p \) and the constraints are concave with respect to \( I \), implying that the Kuhn-Tucker first-order conditions are (necessary and) sufficient to find the solution of this maximization problem. We write down the Lagrangian:

\[ L = (1-k)(m_0 + a_i) - \frac{\theta}{2} I^2 + \beta(m_0 + I)(1-k - \delta I) + \lambda \left( \frac{1-k}{\delta} - I \right) \]

The FOCs are:

\[
\frac{\partial L}{\partial I} = \frac{\partial EU_p}{\partial I} - \lambda = (1-k) \frac{\partial a_i}{\partial I} - \theta I + \beta[1-k - \delta I - \delta(m_0 + I)] - \lambda \leq 0 , \quad \frac{\partial L}{\partial I} I = 0
\]

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

\[
I^* = \frac{1-k}{\delta} = I_{\mu}, \text{ if } m_0 \leq R - \frac{2\beta \delta + \theta}{\beta \delta} I_{\mu}
\]

\[
I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0), \text{ if } R - \frac{2\beta \delta + \theta}{\beta \delta} I_{\mu} \leq m_0 \leq R
\]

\[
I^* = 0 , \text{ if } m_0 \geq R ,
\]

where \( R = (1-k)\left(\frac{1}{\delta} + \frac{1}{\rho} \frac{h_x}{h_y + h_z}\right) \)

It should be emphasized here that the employer \( P \) 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. In other words, the employer is willing to invest in general human capital with higher intensity because he predicts the positive effect of this investment on worker’s attitude (i.e. on her effort contribution \( a_i \)). This important kind of interaction between the worker’s incentives to provide effort and the
employer’s incentives to invest in general human capital provides an additional rationale for firm-sponsored general training and has largely been ignored as yet in the associated literature.

Note that $I^* = 0$ when $k = 1$: This extreme case simply replicates Becker’s argument that the firm will make zero investment in general training if there are no frictions in the labor market (i.e. if the labor market is perfectly competitive and thus the worker reaps all benefits associated with the provision of training).

We can summarize the equilibrium outcome derived above in the following proposition.

**Proposition 1.** The equilibrium outcome involves:

\[
I^* = \frac{1-k}{\delta} I_u, \text{ if } m_0 \leq R - \frac{2\beta \delta + \theta}{\beta \delta} I_u
\]

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

\[
I^* = 0, \text{ if } m_0 \geq R = (1-k)\left(\frac{1}{\delta} + \frac{1}{\rho h_0 + h_\varepsilon}\right)
\]

\[
\alpha^* = \frac{\beta}{\rho h_0 + h_\varepsilon} (k + \delta I^*), \quad \alpha^*_2 = 0
\]

\[
Ew^*_1 = k(m_0 + \alpha^*_1), \quad Ew^*_2 = (k + \delta I^*)(m_0 + I^*)
\]

and the equilibrium welfare is:

\[
TS = W^* = (1 + \beta)m_0 + \alpha^*_1 - \frac{\rho}{2}(\alpha^*_1)^2 + \beta I^* - \frac{\theta}{2}(I^*)^2
\]

5. Implications

5.1 Substitutability Between General Training and Ability

It is clear from the equilibrium described above that a higher level of worker’s average innate ability implies lower investment in general human capital by the employer – i.e. ability and training are substitutes: $\frac{\partial I^*}{\partial m_0} \leq 0$. This negative relationship is depicted in Figure 1 (under the assumption that $R - \frac{2\beta \delta + \theta}{\beta \delta} I_u \geq 0$).
However, it should be kept in mind that this relation of substitution is contingent on the specific form of the production technology adopted here (note that ability and general skills acquired through training I are perfect substitutes in the production function of the second period). This means that the relationship between ability and equilibrium training could be different if, for example, we assumed a Leontief technology (to take the other extreme case – i.e. that of perfect complementarity). In that case, it should be verified that the other implications analyzed below are still valid.

![Figure 1](image)

**5.2 Comparative Statics**

In this subsection we study the effect of worker’s power captured by the parameter $\delta$ (which shows the increase in worker’s bargaining power associated with an additional unit of training) on the employer’s equilibrium investment in general human capital. To this end, we rewrite the equilibrium level of training in terms of $\delta$. After the appropriate calculations (which can be found in the Appendix A), we find:
• **Case 1.** For \( m_0 \leq \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} \):

\[
I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0), \text{ if } \delta \leq \delta_2 = \frac{\beta (1-k) + \sqrt{\beta^2 (1-k)^2 + 4 \beta (1-k) \frac{h_z}{\rho} (h_0 + h_z) -(m_0 + \theta (1-k))}}{2 \beta \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - m_0}
\]

\[
I^* = \frac{1-k}{\delta} = I_H, \text{ if } \delta \geq \delta_2
\]

\[
(11a)
\]

• **Case 2.** For \( m_0 \geq \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} \):

\[
I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0), \text{ if } \delta \leq \delta_0 = \frac{1-k}{m_0 - \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z}}
\]

\[
I^* = 0, \text{ if } \delta \geq \delta_0
\]

For the parameter interval where \( I^* = I_H \), it is clear that: \( \frac{\partial I^*}{\partial \delta} < 0, \frac{\partial^2 I^*}{\partial \delta^2} > 0 \).

For the parameter intervals where \( I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0) \), we have:

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

\[
(12)
\]

The first term in (12) is a direct negative effect of worker’s power on employer’s incentives to invest in general training, since a higher value of \( \delta \) reduces the employer’s share in the second-period output. But the second term represents an indirect positive effect (since \( \frac{\partial^2 a_i}{\partial I^* \partial \delta} = \frac{\beta h_z}{\rho h_0 + h_z} > 0 \)) associated with the fact that a higher value of \( \delta \) implies a stronger positive effect of training on worker’s incentives and thus makes the firm (which takes the latter effect into account) 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 can be positive. In particular, we can immediately find the results summarized in the next proposition:

**Proposition 2**

**Case 1a:** If \( m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2 I \) then \( \frac{\partial I^*}{\partial \delta} > 0 \) (and \( \frac{\partial^2 I^*}{\partial \delta^2} < 0 \)) for all \( \delta \leq \delta_2 \).
where $I^* = \frac{\beta(1-k)}{\theta}$ (see Figure 2)

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

(see Figure 3)

**Case 2:** If $m_0 > \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e}$ then $\frac{\partial I^*}{\partial \delta} < 0$ (and $\frac{\partial^2 I^*}{\partial \delta^2} > 0$) for all $\delta \leq \delta_0$ (see Figure 4)
Figure 3 (Case 1b)

Figure 4 (Case 2)
The comparative statics analysis for case 1a show that a higher level of worker’s power (up to a threshold value $\delta_2$) may in fact increase the employer’s incentives to invest in general training. In particular, this happens when the worker’s expected ability is low enough. This result is in sharp contrast with the predictions of both the orthodox and the radical economic literature briefly reviewed in section 1, according to which any increase in worker’s power weakens the employer’s incentives to invest in general human capital.

We can work in a similar manner to derive the effect of worker’s power $\delta$ on first-period equilibrium effort $a_i^*$. First, we write down the equilibrium effort derived above in terms of $\delta$:

- **Case 1.** For $m_0 \leq 1 - \frac{k}{\rho} \frac{h_z}{h_0 + h_z}$:
  
  \[
  a_i^* = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (k + \delta I^*) , \text{ if } \delta \leq \delta_2 \quad (\text{where } I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0))
  \]
  
  \[
  a_i^* = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} k , \text{ if } \delta \geq \delta_2
  \]

- **Case 2.** For $m_0 \geq 1 - \frac{k}{\rho} \frac{h_z}{h_0 + h_z}$:

  \[
  a_i^* = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (k + \delta I^*) , \text{ if } \delta \leq \delta_0 \quad (\text{where } I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0))
  \]
  
  \[
  a_i^* = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} k , \text{ if } \delta \geq \delta_0
  \]

For the intervals where $a_i^* = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (k + \delta I^*)$, with $I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0)$ we have:

\[
\frac{\partial a_i^*}{\partial \delta} = \frac{\beta}{\rho} \frac{h_z}{h_0 + h_z} (I^* + \delta \frac{\partial I^*}{\partial \delta})
\]

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

---

3 Clearly, we have $\frac{\partial a_i^*}{\partial \delta} = 0$ for all other parameter intervals.
overall impact of $\delta$ on employer’s incentives to provide general human capital (which is described in Proposition 2).

Equation (14) directly 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 $\frac{\partial a_1^*}{\partial \delta}$ is ambiguous. The appropriate calculations yield the following result:

**Proposition 3.**

| Case 1a | If $m_0 < \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x} - 2I$ 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 $\frac{1-k}{\rho} \frac{h_x}{h_0 + h_x} - 2I < m_0 < \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x}$ 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 $m_0 > \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x}$ 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 |
| | $\delta_4 = \frac{-2\theta(m_0 - \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x}) + \sqrt{4\theta^2(m_0 - \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x})^2 + 8\theta(1-k)(m_0 - \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x})}}{4\beta(m_0 - \frac{1-k}{\rho} \frac{h_x}{h_0 + h_x})}$ |
| | (see Figure 7) |

Proof: See Appendix B.
Figure 5 (Case 1a)
Note that when the worker’s expected ability is high enough (case 2), any increase in worker’s power above a critical value \( \delta_4 \) implies a negative impact on employer’s investment incentives \( \left( \frac{\partial I^*}{\partial \delta} < 0 \right) \) so strong that it also pushes worker’s first-period incentives downward \( \left( \frac{\partial a^*_1}{\partial \delta} < 0 \right) \).

### 5.3 The Possibility of Overinvestment in General Training

In all previous models studying the issue of general training, a standard prediction is that the employer underinvests in general human capital (relative to the first best) in equilibrium. In this subsection, we show the possibility of overinvestment in training under certain conditions. First of all, it is clear (from the comparative statics analysis conducted above) that overinvestment can arise only in case 1a: For \( \delta = 0 \), we know that \( I^* = I = \frac{\beta(1-k)}{\theta} < I^{FB} = \frac{\beta}{\theta} \); furthermore, any increase of worker’s power \( \delta \) in cases 1b and 2 reduces equilibrium investment in training 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:

(a) \( \beta < \frac{\theta(1-k)}{2} \left( \frac{1}{\rho} \frac{h_k}{h_0 + h_e} + \theta k \right) \) and \( m_0 < \frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \frac{1-k}{\rho} \frac{h_k}{h_0 + h_e} \),

or

(b) \( \beta > \frac{\theta(1-k)}{2} \left( \frac{1}{\rho} \frac{h_k}{h_0 + h_e} + \theta k \right) \) and \( m_0 < \frac{1-k}{\rho} \frac{h_k}{h_0 + h_e} - 2I^{FB} \),

then equilibrium involves overinvestment in general training for \( \delta_5 < \delta < \delta_6 \), where:

\[
\delta_5 = \frac{k}{\rho} \frac{h_k}{h_0 + h_e} - m_0 - 2I^{FB}, \quad \delta_6 = \frac{\theta(1-k)}{\beta}
\]

Proof: See Appendix C
The case of overinvestment is depicted in Figure 8 below.

![Figure 8. Overinvestment in General Training](image)

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 above, the employer takes this effect into account and is thus willing to provide general skills to the worker with higher intensity. If the positive effect on worker’s incentives is sufficiently strong, then the employer’s incentives to invest in training can be excessively strong.

Note that overinvestment in training on the part of the firm 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 substitutes in the production function assumed here (see subsection 5.1). Of course, the opposite (overinvestment for high values of expected ability) might be the case with a production technology where training 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 both for too low values of power and for too high values of $\delta$: If the worker’s power is very small ($\delta < \delta_5$), then the employer anticipates that his investment in training will not have a very strong positive effect on worker’s first-period incentives; as a result, the employer does not
invest too much. On the other hand, if the worker is too strong \( (δ > δ₀) \) then the firm’s investment incentives are again pushed below the optimum (note that, in this case, even a small level of training investment implies that the worker will enjoy the full second-period expected output). Therefore, the model predicts that over-investment may occur only for intermediate values of worker’s power. In this respect, the high levels of general training observed in Japanese firms may indicate that workers employed in these firms are neither too weak (for example, there are high levels of employment security) nor too strong vis-à-vis the employer. It would be useful to test the validity of this theoretical conclusion by use of relevant empirical data.

### 5.4 Welfare Analysis

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

\[
TS = W^* = (1 + \beta)m₀ + a'_i - \frac{P}{2}(a'_i)^2 + \beta I' - \frac{θ}{2}(I')^2
\]

We focus here on the effect of worker’s power \( δ \) on social welfare:

\[
\frac{∂W^*}{∂δ} = \frac{∂a'_i}{∂δ}(1 - \rho a'_i) + \frac{∂I^*}{∂δ}(β - θ I^*) \tag{15}
\]

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 incentives to invest in general training. 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 total surplus maximization. It will be made clear that a zero level of power is socially optimal only for a limited range of parameters, while in 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 case (parameter range) separately:
Case 1a: \[ m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2^* \]

\( \beta < \frac{\theta(1-k)}{\rho} \left( \frac{h_z}{h_0 + h_z} + \theta k \right) \) and \[ m_0 < \frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - \frac{\beta \theta(1-k)(2-k)}{\beta + \theta^2(1-k)^2} \]
or

\( \beta > \frac{\theta(1-k)}{\rho} \left( \frac{h_z}{h_0 + h_z} + \theta k \right) \) and \[ m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2I^{FB} \]

Then,

- If \( \delta_5 < \delta < \delta_2 \):

\[ \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_1}{\partial \delta} (1 - \rho \alpha^*_1) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) > 0 \]

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

- If \( \delta_2 < \delta < \delta_6 \):

\[ \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_1}{\partial \delta} (1 - \rho \alpha^*_1) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) > 0 \]

A further increase in worker’s power has again an unambiguously positive effect on welfare, because it implies now a lower (thus closer to the first-best) investment in training, while the worker’s incentives remain unaffected.

- If \( \delta > \delta_6 \):

\[ \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_1}{\partial \delta} (1 - \rho \alpha^*_1) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) < 0 \]

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 again remain unaffected).

Note that \( \delta = \delta_6 \) is a local (and perhaps also global) welfare maximizer in this case.
We proceed in the same way to examine the other parameter intervals (in which underinvestment is always the case).

(ii) \( \beta < \frac{\theta(1-k)}{2} \left( \frac{1}{\rho} \frac{h_z}{h_0 + h_z} + \theta k \right) \)

and \( \frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \cdot \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - \beta \theta (1-k)(2-k) < m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2I^{FB} \),

or \( \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2I^{FB} < m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2I^{*} \).

Then:

- If \( \delta \delta_2 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) > 0 \)

- If \( \delta > \delta_2 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) < 0 \)

Note that \( \delta = \delta_2 \) is the socially optimal level of worker’s power in this case.

**Case 1b:** \( \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} - 2I^{*} < m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} \). Then:

- If \( \delta < \delta_2 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) \) which can be either positive or negative.

- If \( \delta > \delta_2 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) < 0 \)

**Case 2:** \( m_0 > \frac{1-k}{\rho} \frac{h_z}{h_0 + h_z} \). Then:

- If \( \delta < \delta_4 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) \) which can be either positive or negative.

- If \( \delta_4 < \delta < \delta_0 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) < 0 \)

- If \( \delta > \delta_4 \): \( \frac{\partial W^*}{\partial \delta} = \frac{\partial \alpha^*_i}{\partial \delta} (1 - \rho \alpha^*_i) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) = 0 \)
The above results already show that a zero level of worker’s power is a possible welfare maximizer only in cases 1b and 2. In all other parameter intervals, the welfare maximizing level of power is always positive (and bounded from above).

In order to derive more precise results, we write down the general welfare maximization problem for the parameter intervals where \( I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0) \), i.e.

\[
\frac{\partial W^*}{\partial \delta}
\]

for \( \delta < \delta_2 \) in case 1 or \( \delta < \delta_0 \) in case 2 (note that we have already found the sign of \( \frac{\partial W^*}{\partial \delta} \)

for \( \delta > \delta_2 \) or \( \delta > \delta_0 \)). Define \( \delta^* \) as follows:

\[
\delta^* = \delta_2 \text{ in case 1}
\]

\[
\delta^* = \delta_0 \text{ in case 2}
\]

Then, we must solve the following maximization problem:

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

s.t. \( 0 \leq \delta \leq \delta^* \)

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

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

\[
m_0 \geq \frac{1 - k}{\rho} \frac{h_x}{h_y + h_x} \left[ \frac{\beta(1 - \beta k \frac{h_x}{h_y + h_x})}{1 + \frac{h_x}{h_y + h_x} \frac{\beta - \theta I}{\beta - \theta I}} \right] - 2 I
\]

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

Proof: See Appendix D

Note that this is only a necessary (not sufficient) condition for \( \delta = 0 \) to be optimal. 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 \( (\frac{\partial I^*}{\partial \delta} < 0) \) – thus pushing \( I^* \) further below the optimum. This negative impact on welfare dominates the positive impact associated with the
worker’s stronger incentives to provide effort in the first period \( \left( \frac{\partial a_i^*}{\partial \delta} > 0 \right) \). As a result, any initial increase in worker’s power above \( \delta=0 \) is detrimental to welfare.

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

**Proposition 6.** If (16) does not hold, then there is a positive and bounded level of power \( \delta^* \) which maximizes social welfare. In particular, there is an inverse-U relationship between power and welfare: Initial increases of power up to the critical value \( \delta^* \) enhance welfare, whereas further increases of power above that critical value are detrimental to welfare.

We close the subsection related to the welfare analysis with some numerical examples, each of which corresponds to one of the different cases generally depicted above.

5.4.1 Numerical Examples

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

These parameter values correspond to case 1b (implying underinvestment for all values of \( \delta \)). The equilibrium outcome involves:

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

\[
I^* = 1/2 \delta, \text{ if } \delta \geq \sqrt{3} + 1
\]

\[
\alpha_i^* = \frac{1}{2} (\frac{1}{2} + \delta I^*) \text{, if } \delta \leq \delta_2 = \sqrt{3} + 1
\]

\[
\alpha_i^* = \frac{1}{2}, \text{ if } \delta \geq \sqrt{3} + 1
\]

and: \( I_{FB} = a_{FB} = 1 \).

The welfare maximization problem \((P')\) becomes:

\[
\max_{(\delta)} W^*(\delta) = a_i^* - \frac{1}{2} \left(a_i^*\right)^2 + I^* - \frac{1}{2} \left(I^*\right)^2
\]

s.t. \( 0 \leq \delta \leq \sqrt{3} + 1 \)
The solution of \((P')\) is \(\delta = 0\). Since \(\frac{\partial W^*}{\partial \delta} < 0\) for all \(\delta > \delta_2 = \sqrt{3} + 1\), a zero level of worker’s power maximizes equilibrium welfare in this case. Indeed, for these parameter values it can be directly verified that the condition (16) holds.

\[\text{(ii)} \] Let \(\beta = \rho = h_c = h_0 = 1, k = 1/2, m_0 = 0, \theta = 8\).

These parameter values correspond to case 1a (but again with underinvestment in equilibrium for all values of \(\delta\)). The equilibrium outcome involves:

\[
I^* = \frac{\delta + 2}{8(\delta + 4)}, \quad \text{if } \delta \leq \delta_2 = \sqrt{17} + 1
\]

\[
I^* = 1/2\delta, \quad \text{if } \delta \geq \sqrt{17} + 1
\]

\[
\alpha_i^* = \frac{1}{2} \left(\frac{1}{2} + \delta I^*\right), \quad \text{if } \delta \leq \sqrt{17} + 1
\]

\[
\alpha_i^* = \frac{1}{2}, \quad \text{if } \delta \geq \sqrt{17} + 1
\]

and: \(I^F = 1/8, \ a^F = 1\).

The welfare maximization problem \((P')\) becomes:

\[
\max_{\delta} W^*(\delta) = a_i^* - \frac{1}{2}(a_i^*)^2 + I^* - 4(I^*)^2
\]

\[\text{s.t. } 0 \leq \delta \leq \sqrt{17} + 1\]

The solution now is \(\delta = \delta_2 = \sqrt{17} + 1\). Since \(\frac{\partial W^*}{\partial \delta} < 0\) 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 (see 1a-ii).

\[\text{(iii)} \] Let \(\beta = \rho = h_c = h_0 = 1, k = 1/2, m_0 = 0, \theta = 16\).

These parameter values correspond to case 1a-i, with overinvestment for \(\delta_5 < \delta < \delta_6\).

The equilibrium values correspond to case 1a-i, with overinvestment for \(\delta_5 < \delta < \delta_6\).
In this case, there is overinvestment in equilibrium for \( \delta < \delta_5 < \delta_6 \), i.e. \( 4 < \delta < 8 \).

The welfare maximization problem \( (P') \) becomes:

\[
\max_{\{\delta\}} W^*(\delta) = a_i^* - \frac{1}{2} (a_i^*)^2 + I^* - 8(I^*)^2
\]

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

The solution now is \( \delta = \delta_2 = \sqrt{33} + 1 \). Since \( \frac{\partial W^*}{\partial \delta} > 0 \) for \( \delta_2 < \delta < \delta_6 = 8 \) and \( \frac{\partial W^*}{\partial \delta} < 0 \) 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 \( \delta = 0 \) cannot be a welfare maximizer in either of these cases.

6. Policy Implications and Possible Extensions

A standard question often addressed in the literature concerns the appropriate policy instrument and the optimal degree of public intervention in order to alleviate inefficiencies associated with the employer’s incentives to invest in general training. In this respect, the prediction of underinvestment which is shared by all contributions on the subject implies that a government subsidy to the firm is an appropriate policy measure to strengthen the firm’s investment incentives and thus enhance equilibrium welfare\(^4\). However, the richer framework adopted here shows that this might not necessarily be an appropriate policy recommendation. The possibility of over-investment 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 will 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 her

---

\(^4\) A notable exception here is Moen and Rosen (2002).
7. Concluding Remarks

In this paper we have studied the provision of firm-sponsored general training when worker’s career concerns are also taken into account. A crucial assumption was that a higher level of general training received in the first period shifts the balance of class forces in favour of workers and thus increases the employee’s bargaining power vis-à-vis the employer in the next period (in this sense, power might be interpreted as the wage premium associated with an additional unit of general human capital). In this framework, it has been shown that a 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, overinvestment in training may occur in equilibrium. Furthermore, 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 the predictions of both the radical economic literature and orthodox models of general training (underinvestment in general human capital along with a negative impact of worker’s power on investment incentives and on social welfare) might not hold in an enriched framework that takes workers’ incentives into account. Alternative policy measures to alleviate the associated inefficiencies as well as possible extensions of the model have also been discussed.
Appendix A

We have: \( I^* = I_{H} \iff m_0 \leq R - \frac{2\beta \delta + \theta}{\beta \delta} I_{H} \iff \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} - m_0 \) \hspace{1cm} (A1)

The condition (A1) can hold only if \( m_0 \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \). In this case:

(A1) \( \iff \beta \left( \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} - m_0 \right) \delta^2 - \beta(1-k)\delta - [m_0 + \theta(1-k)] \geq 0 \) or, equivalently: \( \delta \geq \delta_2 \)

Also:

\[ I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0) \iff R - \frac{2\beta \delta + \theta}{\beta \delta} I_{H} \leq m_0 \leq R \iff \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} - \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \leq m_0 \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} + \frac{1-k}{\delta} \]

(A2)

(i) \( m_0 \geq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} - \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \iff \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \geq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} - m_0 \) \hspace{1cm} (A3)

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

- If \( m_0 \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) then: (A3) \( \iff \delta \leq \delta_2 \)

(ii) \( m_0 \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} + \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \iff \frac{1-k}{\delta} \frac{\beta \delta + \theta}{\beta \delta} \geq m_0 - \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) \hspace{1cm} (A4)

- If \( m_0 \geq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) then: (A4) \( \iff \delta \leq \delta_0 \)

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

Finally: \( m_0 \geq R \iff \frac{1-k}{\delta} \leq m_0 - \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) \hspace{1cm} (A5)

- If \( m_0 \geq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) then: (A5) \( \iff \delta \geq \delta_0 \)

- If \( m_0 \leq \frac{1-k}{\rho} \frac{h_{z}}{h_{0} + h_{z}} \) then (A5) cannot hold.

The above results are summarized in conditions (11a), (11b).
Appendix B (Proof of Proposition 3)

For \( m_0 \leq \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} \) we have:

\[
a_i^* = \frac{\beta}{\rho} \frac{h_e}{h_0 + h_e} (k + \delta I^*) \quad \text{with} \quad I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0), \quad \text{if} \quad \delta \leq \delta_2
\]

\[
\frac{\partial a_i^*}{\partial \delta} = \frac{\beta}{\rho} \frac{h_e}{h_0 + h_e} (I^* + \delta \frac{\partial I^*}{\partial \delta}) > 0 \Leftrightarrow \frac{1-k}{\delta} \frac{\theta}{2(\beta \delta + \theta)} > m_0 - \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} \quad \text{which holds for all} \quad \delta \leq \delta_2 \quad (B1)
\]

Also:

\[
\frac{\partial^2 a_i^*}{\partial \delta^2} = \frac{\beta}{\rho} \frac{h_e}{h_0 + h_e} (2 \frac{\partial I^*}{\partial \delta} + \delta \frac{\partial^2 I^*}{\partial \delta^2}) > 0 \Leftrightarrow m_0 < \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} - 2I \quad (B2)
\]

The conditions (B1) and (B2) yield cases 1a, 1b in Proposition 3.

Similarly, for \( m_0 \geq \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} \) we have:

\[
a_i^* = \frac{\beta}{\rho} \frac{h_e}{h_0 + h_e} (k + \delta I^*) \quad \text{with} \quad I^* = \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0), \quad \text{if} \quad \delta < \delta_0
\]

\[
\frac{\partial a_i^*}{\partial \delta} = \frac{\beta}{\rho} \frac{h_e}{h_0 + h_e} (I^* + \delta \frac{\partial I^*}{\partial \delta}) > 0 \Leftrightarrow \frac{1-k}{\delta} \frac{\theta}{2(\beta \delta + \theta)} > m_0 - \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} \Leftrightarrow \delta \leq \delta_2 \quad (B3)
\]

\[
\frac{\partial^2 a_i^*}{\partial \delta^2} < 0 \Leftrightarrow m_0 < \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} - 2I \quad \text{which holds for all} \quad \delta < \delta_0 \quad (B4)
\]

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

Appendix C (Proof of Proposition 4)

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

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

If \( \delta < \delta_2 \) then: \( I^* > I^{FB} \Leftrightarrow \frac{\beta \delta}{2 \beta \delta + \theta} (R - m_0) > \frac{\beta}{\theta} \Leftrightarrow \frac{k}{\delta} - \frac{1-k}{\rho} \frac{h_e}{h_0 + h_e} - m_0 - 2I^{FB} \quad (C1) \)
(i) For \( m_0 > \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I_{FB} \),
\[ \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I_{FB} < m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I \]

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

(ii) For \( m_0 < \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I_{FB} \):

\[ (C1) \Leftrightarrow \delta > \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} = \delta_3 \Leftrightarrow \delta_3 < \delta < \delta_2 \]

with necessary condition: \( \delta_3 < \delta_2 \)

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

with necessary condition:

\[
\delta_2 < \delta_6 \Leftrightarrow \frac{\beta(1-k)}{\beta + \theta^2(1-k)^2} + 4\beta \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - m_0 - \frac{\theta(1-k)}{\beta} < \frac{\theta(1-k)}{\beta}.
\]

\[ m_0 < \frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - \frac{\theta(1-k)(2-k)}{\beta + \theta^2(1-k)^2}, \text{ i.e.}
\]

\[ m_0 < \min \left\{ \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I_{FB}, \frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - \frac{\theta(1-k)(2-k)}{\beta + \theta^2(1-k)^2} \right\}
\]

with:

\[
\frac{\theta^2(1-k)^2}{\beta + \theta^2(1-k)^2} \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - \frac{\theta(1-k)(2-k)}{\beta + \theta^2(1-k)^2} < \frac{1-k}{\rho} \frac{h_z}{h_0+h_z} - 2I_{FB}
\]

\[ \Leftrightarrow \beta < \frac{\theta(1-k)}{2} \frac{1}{\rho} \frac{h_z}{h_0+h_z} + \theta k \]

Finally, note that \( \delta_3 < \delta_6 \) implies \( \delta_3 < \delta_2 \).

The conditions (C1) to (C5) yield the results summarized in Proposition 4.

**Appendix D** (Proof of Proposition 5)

The welfare maximization problem \((P')\) is not (globally) concave with respect to \( \delta \).
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 corresponding to (P') is:

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

The Kuhn-Tucker first order conditions are:

\[ \frac{\partial L}{\partial \delta} = \frac{\partial W^*}{\partial \delta} - \lambda = \frac{\partial a_i^*}{\partial \delta} (1 - \rho a_i^*) + \frac{\partial I^*}{\partial \delta} (\beta - \theta I^*) - \lambda = 0 \]

\[ \frac{\partial L}{\partial \lambda} = \delta^* - \delta \leq 0, \quad \frac{\partial L}{\partial \lambda} \lambda = 0 \]  
\[ \text{(D1)} \]

\[ \frac{\partial L}{\partial \lambda} = \delta^* - \delta \geq 0, \quad \frac{\partial \lambda}{\partial \lambda} \lambda = 0 \]  
\[ \text{(D2)} \]

We check if \( \delta = 0 \) can be a solution to (P'). For \( \delta = 0 \) the conditions (D2) imply \( \lambda = 0 \). In this case (D1) requires:

\[ \text{(D1)} \iff \frac{\partial L}{\partial \delta} = \frac{\partial W^*}{\partial \delta} \leq 0 \iff \frac{\beta}{\rho} \frac{h_\epsilon}{h_0 + h_\epsilon} I(1 - \beta k \frac{h_\epsilon}{h_0 + h_\epsilon}) + (\beta - \theta I) \frac{\partial I^*}{\partial \delta} |_{\delta = 0} \leq 0 \]  
\[ \text{(D3)} \]

We have:

\[ \frac{\partial I^*}{\partial \delta} |_{\delta = 0} = I^{FB} \left( \frac{1 - k}{\rho} \frac{h_\epsilon}{h_0 + h_\epsilon} - m_0 - 2I \right) \]  
\[ \text{(D4)} \]

Then:

\[ \text{(D3)} \iff m_0 \geq \frac{1 - k}{\rho} \frac{h_\epsilon}{h_0 + h_\epsilon} \left[ \frac{\beta(1 - \beta k \frac{h_\epsilon}{h_0 + h_\epsilon})}{\beta - \theta I} \right] - 2I. \]
References


- 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.


