Firm- and Self-sponsored General Human Capital, Wage Floors and Credit Constraints.

Felipe Balmaceda, Assoc Prof.

Available at: https://works.bepress.com/felipe_balmaceda/15/
Firm- and Self-sponsored General Human Capital, Wage Floors and Credit Constraints.

Felipe Balmaceda *
Economics Department
Diego Portales University †
First-Version: January 2014
October 19, 2016

*I would like to express my gratitude to participants to the regular seminar at CEA and Ilades-Georgetown and the Chilean Economic Association Conference, 2014. I would also thank the financial support of FONDECYT project # 1140140, the support of the Instituto Milenio de Sistemas Complejos en Ingeniería and Instituto Milenio para la investigación de imperfecciones de mercado y políticas públicas grant ICM IS130002.

†Avda. Santa Clara 797, Santiago, 8580000, Chile. Email: felipe.balmaceda@udp.cl
Abstract

This paper provides a model to organize our thinking and to explore the consequences of credit constraints and the existence of wage floors on firm- and self-sponsored general human capital. Workers are subject to credit constraints due to incomplete creditor’s protection, wage are determined by Rubinstein’s alternating offers with outside options bargaining game and training is general in Becker’s sense; i.e., it is equally productive across all firms. In the absence of wage floors, Becker’s result is fully recovered; for workers with a low personal wealth level, firms invest and pay for general training together with the workers themselves, while for workers with a higher personal wealth level, training is fully financed by the workers themselves. General Training increases with personal wealth only for those workers whose wealth is neither large nor small. Contractual provisions that can modelled as a wage floor such as up-or-out rules or efficiency wages moves away from Becker’s result; that is, firms pay for training of workers who has enough personal wealth to pay for it. These results suggests an explanation for the observed heterogeneity in firm- and self-sponsored training across countries and a rationale for the evidence documenting higher wage returns to firm-sponsored human capital than to self-sponsored human capital.

**JEL:** J3, D2, J30

**Key Words:** Firm- and self-sponsored General Human Capital, Liquidity constraints, Wage Floors, Promotions, Up-or-out Rules.
1 Introduction

Human capital is a fundamental determinant of economic performance. Post-school training is key to augment and adapt the existing human capital to technical and structural changes, to decrease the risk of unemployment, to increase wages and to improve career prospects. For a long time economists have argue that contractual provisions and credit constraints as play a crucial role in human capital acquisition. The former is due to they compress the wage structure and the latter is due to the fact that young workers cannot pledge their future skills or labor as collateral. This paper develops a framework for analysis of firm- and self-sponsored general training investments in the presence of imperfect credit markets and under contractual provisions that compress the wage structure, and uses it to examine the relationship between general training investments, training financing sources, skills, personal wealth and returns.

The empirical fact related to training are as follows:1 (i) firms finance a large share of training costs despite most training being general in nature; (ii) employees also contribute to their human capital; (iii) wage compression increases firm-provided training; (iv) wage returns due to firm-sponsored training are greater than those due to self-sponsored training; (v) there is strong correlation between skills and training incidence and more educated appear to receive less training; and (vi) there is a great cross-country variation in firm- and self-sponsored training. In addition, there is no evidence regarding credit constraints, financial frictions and general training investments.2 This may in part due to the fact that training theories do not provide very specific predictions regarding the relationship between credit constraints and general training, beyond that firms could pay for training when workers do not have money to pay for it. Thus the evidence suggests that in order to understand the demand and financing of general training one has to look at employers and employees together, we have to take a closer look at the relationship between firm- and self-sponsored training, workers’ skills must be considered, and somehow we have to take into account the different contractual provisions governing the employment relationship into different countries. In spite of the lack of evidence regarding the relationship between workers’ credit constraints and training, we also think is important to take them into account. If firms’ constraints matter, the it is highly plausible that workers’ constraints are more important as they are for schooling and college

1These are discussed in more detail in section XX.
2There is evidence regarding credit constraints and schooling and another regarding firms’ credit constraints and training. Mainly, ? finds that firms’ lack of access to finance in general, mainly to bank credit, results in a significantly lower investment in on-the-job training. This effect is stronger in education-intensive industries and in industries facing good global growth opportunities. Lochner and Monge-Naranjo (2011) propose a model that helps to explain the persistent strong positive correlation between ability and schooling in the United States, as well as the rising importance of family income for college attendance. It also explains the increasing share of undergraduates borrowing the GSL maximum and the rise in student borrowing from private lenders.
This paper provides a framework that brings together all these dimensions. The model does a better job at fitting the data than the available models and provides new and interesting predictions that can be taken to data.

Since Becker (1964), economists have argued that the analysis of firms’ investment in general training in a competitive labor market is straightforward; due to perfect competition, workers capture the full return to their general human capital and thus firms should not pay for this type of human capital. In other words, if firms invest in general human capital, the non-pledgeability of skills will allow workers to fully hold-up firms. More recent theories of firm-provided training can explain why firms provide general training, yet they mainly focus on employers’ incentives to training financing. In particular, Balmaceda (2005) considers a Becker’s type of model in which bargaining and employment on the spot market are mutually exclusive and there are match-specific productivity shocks. He shows that firms invest in both general and specific training, although at a sub-optimal level. Acemoglu and Pischke (1998, 1999a) show that when workers face credit constraints and labor market imperfections distort the wage structure within the firm, pushing it away from the competitive benchmark and favoring skilled workers, it will be profitable for firms to provide workers with general human capital. This is because labor market imperfections make general abilities de-facto specific in the sense that trained workers do not get their full marginal product of training when they switch to another job. Chang and Wang (1996) and Katz and Ziderman (1990) make the same point in a slightly different context. They assume that trained workers’ marginal productivity is the incumbent firm’s private information. This implies that trained workers’ outside offers do not fully reflect their productivity. Thus, the incumbent firm can appropriate a share of the return to general training and therefore the investment in general training is positive and positively related to the probability of a worker staying with the incumbent employer. In addition, the existence of a hold-up problem as well as liquidity constraints have motivated economists to study all kinds of contractual provisions that can deal with these issues. For instance, Waldman (1990), Waldman (2003), Prendergast (1993), MacLeod and Malcolmson (1993) study. None of these papers is capable of explaining the empirical fact summarized above and

This paper proposes a two-period competitive labor market model between risk neutral firms and workers. The crucial assumptions of the model are:3 (i) there are match-specific productivity shocks and aggregated shocks that do not affect the marginal product of human capital; i.e., skills and training; (ii) the human capital technology follows Ben-Porath and therefore human capital rises at decreasing rate with training and this and innate skills are complements; (iii) wages are determined by Rubinstein’s alternating-offer bargaining game with outside options. In other words,

---

3This paper builds on Balmaceda’s (2005).
assuming that bargaining and employment on the spot market are mutually exclusive. Thus, in contrast to most models in the literature, here the no-trade payoffs for the firm and worker enter the bargaining process as outside options instead of as inside options; (iv) workers are initially endowed with a skill level (schooling) and a personal wealth level that determines, together with financial frictions, their ability to pay for training; and (v) it considers any contractual provision that results in a wage floor such as minimum wages, up or out rules, unemployment benefits, internal labor market policies that attach wages to jobs, etc...

Wage floors have three counterweighing forces of which only the wage compression effect has been noted in the literature. First, a wage floor decreases the set of states under which the bargaining outcome is sharing the surplus according to each party’s bargaining power. This may either increase or decrease the firm and worker’s incentives to train; second, when the wage floor binds and therefore the wage structure within the firm get compressed, the firm’s marginal return to train rises, while the worker’s marginal return to train falls; and third, the set of states under which a separation occurs rises and there are inefficient separations. This harms the firm’s incentives to train and may either increase or decrease the worker’s incentive to train. Incentives increase when the outside payment that a worker receives after leaving the current firm increases with training at a higher rate than the wage floor. This may happen for instance when the alternative payment is working in the informal sector or self-employment. Hence, a-priori the impact of wage floors on each party’s incentives to train and the overall effect on total training is also ambiguous.

Most of the contractual provisions studied so far to deal with the hold-up problem have an impact on training because they result in a different kinds of wage floors. Furthermore, the articles considering these contractual provisions ignore how the existence of a labor market constraints the ability of wage floors to increase (decrease) firm and workers’ incentives to invest on general training and to hire workers. For instance, how the possibility to work as a self-employed or in an informal sector limits the ability of an up-or-out contractual provision to induce investment in non-verifiable general training. In addition, most papers either study firm-sponsored general training or self-sponsored general training, but not both and when they study the latter they ignore liquidity constraints. In this paper, we wish to clarify the interaction between credit constraints and contractual provisions that give rise to wage floors on firm and workers’ incentives to invest in general training in a labor market where workers’ outside option is endogenously determined by their skills and training and the structure of the labor market; for instance, whether the labor market is dual or not. This is done for two cases: when training is decided cooperatively between firms and workers and there is a-la-Bertrand competition for workers; and when it is decided in a non-cooperative way and there is also a-la-Bertrand competition. Since, contractual provisions
that result in wage floors do not result in the first-best efficient investment level, we also study public policies that could improve firm and workers’ incentives to undertake training such as public financing of training and tax breaks to firms in order to lower firm’s training costs.

The main results are as follows. Formality rises general human capital, yet it may increase or decrease firm-provided human capital. Mainly, for skilled workers; i.e., those for which in equilibrium the wage floor never binds, formality decreases firm-sponsored human capital and increases self-sponsored human capital in such a way that overall general human capital rises. For unskilled workers, formality may either increase or decrease firm-provided human capital; when formality decreases firm-provided human capital, it increases self-sponsored human capital in such a way that overall general human capital rises, while when formality increases firm-provided human capital it might also increase firm-sponsored human capital. For unskilled workers in the formal sector, low severance payments, large firing costs and a large wage floor are more likely to result in an increase in firm-sponsored training in relation to that in the informal sector.

These results provide support for the country-difference in training incidence as well as in training incidence by funding source. Furthermore, the model suggests that focusing on firm-sponsored training paints a rather biased picture of the effects of dualization on training. The theoretical results derived here also rationalize the empirical evidence showing that firm-sponsored training results in larger wage returns than self-sponsored training. Finally, it seems that a natural policy implication of the model is to eliminate the informal sector in order to increase general human capital. However, this is not necessarily correct since duality contributes to increase the investment in general human capital in the formal sector since a worker who loses his job in that sector and cannot find one in the formal sector, could find a job in the informal sector, which will allow him to capitalize on his human capital while unemployment does not.

The rest of the paper is structured as follows. The next section presents the model and derives the first-best efficient training. In Section 4, the optimal training investment in a dual labor market is derived and the relationship with the literature is discussed. In the following section, we compare firms and workers’ incentives to invest in training in the formal sector relative to those in the informal sector. Section 8 discusses the model’s theoretical predictions regarding wage returns. In Section 9, we offer some concluding remarks.
Table 1. Training and LMI: Means by Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Training</th>
<th>EPL</th>
<th>MW</th>
<th>UD</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>9.99</td>
<td>2.67</td>
<td>n.a.</td>
<td>0.36</td>
<td>1.00</td>
</tr>
<tr>
<td>Belgium</td>
<td>6.50</td>
<td>1.72</td>
<td>0.52</td>
<td>0.52</td>
<td>1.15</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>5.57</td>
<td>3.29</td>
<td>0.33</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>Denmark</td>
<td>22.72</td>
<td>1.63</td>
<td>n.a.</td>
<td>0.73</td>
<td>1.59</td>
</tr>
<tr>
<td>Finland</td>
<td>19.61</td>
<td>2.22</td>
<td>n.a.</td>
<td>0.74</td>
<td>1.42</td>
</tr>
<tr>
<td>France</td>
<td>4.81</td>
<td>2.42</td>
<td>0.60</td>
<td>0.08</td>
<td>1.38</td>
</tr>
<tr>
<td>Germany</td>
<td>6.25</td>
<td>2.80</td>
<td>n.a.</td>
<td>0.23</td>
<td>1.98</td>
</tr>
<tr>
<td>Greece</td>
<td>1.53</td>
<td>2.29</td>
<td>0.48</td>
<td>0.26</td>
<td>0.33</td>
</tr>
<tr>
<td>Hungary</td>
<td>3.40</td>
<td>1.92</td>
<td>0.45</td>
<td>0.22</td>
<td>0.36</td>
</tr>
<tr>
<td>Iceland</td>
<td>23.75</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0.87</td>
<td>0.80</td>
</tr>
<tr>
<td>Ireland</td>
<td>6.43</td>
<td>1.60</td>
<td>0.54</td>
<td>0.37</td>
<td>0.51</td>
</tr>
<tr>
<td>Italy</td>
<td>5.23</td>
<td>1.77</td>
<td>n.a.</td>
<td>0.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>6.45</td>
<td>n.a.</td>
<td>0.41</td>
<td>0.43</td>
<td>1.87</td>
</tr>
<tr>
<td>Netherlands</td>
<td>15.20</td>
<td>3.04</td>
<td>0.46</td>
<td>0.22</td>
<td>0.54</td>
</tr>
<tr>
<td>Norway</td>
<td>16.24</td>
<td>2.25</td>
<td>n.a.</td>
<td>0.55</td>
<td>0.25</td>
</tr>
<tr>
<td>Poland</td>
<td>4.66</td>
<td>2.06</td>
<td>0.42</td>
<td>0.21</td>
<td>0.87</td>
</tr>
<tr>
<td>Portugal</td>
<td>3.62</td>
<td>4.27</td>
<td>0.47</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>4.85</td>
<td>2.40</td>
<td>0.42</td>
<td>0.33</td>
<td>1.38</td>
</tr>
<tr>
<td>Spain</td>
<td>5.58</td>
<td>2.54</td>
<td>0.43</td>
<td>0.16</td>
<td>0.79</td>
</tr>
<tr>
<td>Sweden</td>
<td>27.00</td>
<td>2.86</td>
<td>0.30</td>
<td>0.78</td>
<td>0.81</td>
</tr>
<tr>
<td>Switzerland</td>
<td>30.15</td>
<td>1.16</td>
<td>n.a.</td>
<td>0.20</td>
<td>n.a.</td>
</tr>
<tr>
<td>Turkey</td>
<td>1.31</td>
<td>2.57</td>
<td>n.a.</td>
<td>0.10</td>
<td>n.a.</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>23.60</td>
<td>1.07</td>
<td>0.43</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>Total</td>
<td>11.42</td>
<td>2.31</td>
<td>0.45</td>
<td>0.37</td>
<td>0.94</td>
</tr>
</tbody>
</table>
2 Cross-country Variation in Training, Returns, Skills and Financing Constraints

Estimates suggest that general training makes up somewhere between 60 and 90 percent of all the training provided by firms. Estimates from Europe typically document that 80 to 90 percent of the training is general in nature, whereas U. S. based studies provide estimates of general training in the vicinity of 60 to 70 percent of all training spells (see, for instance, Barron, Berger, and Black, 1999, Loewenstein and Spletzer, 1998, 1997, Booth and Bryan, 2002).

![Fig. 1. Training Incidence (%) workers aged 25-64 (source: OECD 2004)](image)

In spite of the importance of training, Bassanini, Booth, Brunello, De Paola, and Leuven (2005) find, after controlling for a relatively large set of time varying individual, job and firm characteristics, that cross-country variation in training incidence is large across European countries.\(^4\) Indeed, after controlling for individual characteristics, country effects account for almost one-half of the explained variation in training across European countries.\(^5\) Furthermore, the OECD International Adult Literacy Survey (IALS) shows that in almost all countries there is a large variation between firm- and self-sponsored training across countries. Figure 2 shows self-reported training incidence by financing source.\(^6\)

\(^4\)For example, a Danish employee has still a 20 percentage point greater probability of taking training than a Portuguese. The estimated range of variation among country effects is far greater than that estimated for educational levels (7.6 percentage points), age classes (6.2), firm size classes (7.7), occupations (13) and industries (12.4).

\(^5\)Part of this variation is probably due to measurement error and cross-country differences in definitions and perceptions of training.

\(^6\)Note that training incidence does not add up to 1 since training could be financed by different sources (firms,
Most research also points to substantially larger returns to training financed by the employer. In fact few studies have been able to document any returns to self-sponsored training. Usually this is explained by the arguable superior knowledge of what training is needed to perform better and improve productivity. Booth and Bryan (2002) using British data find no effects on wages from individual financed training. Similarly, Loewenstein and Spletzer (1998), using data from NLSY, show that non-employer financed training yields no wage return. Blundell, Dearden, Meghir, and Sianesi (1999) also note that employer-provided training has a positive impact on wages whereas training not provided by the employer has an insignificant effect on wages. There are some indications that vocational institute and business school training yield higher returns for the individual when changing employer (Loewenstein and Spletzer, 1998). There is also some evidence that some forms of initial training are inversely related to the starting wage and thus workers somehow pay at least a small share of training costs (see, for instance, Veum, 1999). In addition, the evidence points to a positive effect of general training on job rotation (see, for instance, Loewenstein and Spletzer (1998), Lynch (1992), Brunello and De Paola (2004)). This results are more pronounced for self-sponsored training. Training also affect the career development and internal employability. Mainly, in bad times training provides job stability, and in general is associated with promotions.

The evidence also points to a strong correlation between on-the-job training and skills and less clear correlation between on-the-job training and schooling. Yet, studies controlling for different measures of skills, shows positive return to training (see, Leuven and Oosterbeek (2008)) and the evidence shows larger return to training to those with secondary edcation than to those with college degree (see, for instance, Lynch (1992) and Brunello (2001)). However, Arulampalam, Booth, and workers, government, universities and others).

Fig. 2. Training Incidence by Financing Source (source: IALS, ALL, and AEPS 1993-1999)
Bryan (2004) using male workers in ECHP finds no return difference between low- and high-wage earners. There is no evidence that we are aware of regarding training investment, workers’ credit constraints and financial frictions.

3 The Model

3.1 Set-Up

We consider a two-period model between homogeneous firms \((f)\) and heterogeneous workers \((l)\), both of whom are risk neutral. All firms have access to the same constant-return to scale technology; i.e. the total productivity of a firm is equal to the sum of each worker’s productivity.\(^7\) Each worker has a publicly known multidimensional ability level \(a = (a_1, \ldots, a_n)\), with \(a_i \in A \equiv [a, \bar{a}]\), with \(a > 0\), and an initial personal cash flow or wealth \(A \in [0, \bar{A}]\). For instance, \(a_1\) could be cognitive skills, \(a_2\) mathematical skills and \(a_3\) social skills.

At the beginning of period 1, which is viewed as the early stage of a worker’s career, firms and workers negotiate one period contracts for the supply of one unit of labor and then firms and workers simultaneously decide whether or not to invest in non-contractible general training, where \(\tau_{if} \in \mathbb{R}^+\) is the firm’s investment decision in skill \(i\) and \(\tau_{il} \in \mathbb{R}^+\) is the worker’s investment decision in skill \(i\). The cost of training, which is independent of skills and identical for either type of training, is incurred in the first period and is equal to \(\sum_i \tau_{ij}\) for \(j = f, l\).

At the beginning of period 2, workers’ non-specific productivity, denoted by \(y(t) = \lambda h(t)\), \(t = (a_1 \tau_1, \ldots, a_n \tau_n)\), and a worker-firm specific productivity shock, denoted by \(\eta\), is publicly realized.\(^8\) Hence, if a worker stays with the first-period employer in period 2, he produces \(\lambda h + \eta\), while he produces \(\lambda h\) with an alternative employer. At the beginning of a match, the firm only knows the skill of the worker, but it does not know how well the worker will match to the firm. This is a standard way to endogenize separations in a matching framework (see, for instance, Mortensen and Pissarides (1994)). After productivity becomes known, the parties either negotiate a one period contract for the supply of one unit of labor, or alternatively, they may either refuse to trade, or agree to trade with a third party instead. Wages within the current firm are determined by Rubinstein’s alternating offers with outside options; see, for example, Sutton (1986) and Balmaceda (2005).

Observe that firm- and self-sponsored training are separable in the production technology and they have the same marginal product inside as well as outside of the current firm; that is \(\tau\) is

\(^7\)This assumption is not as restrictive as it appears at first glance. If the technology is of constant returns to scale and inputs can be freely adjusted, the marginal contribution of a worker will be independent of the other inputs. The reason is that profit maximizing firm will keep the ratio between inputs constant.

\(^8\)For the sake of brevity, in what follows the arguments \((\tau, a, \epsilon)\) will be omitted when there is no risk of confusion.
completely general in the sense of Becker and they are neither substitutes nor complements. The latter is due to that: (i) we are agnostic with respect to the relationship between firm- and self-sponsored training; and (ii) these implies that the marginal return to training inside as well as outside of the incumbent firm is the same and therefore in the absence of contractual provisions there are no wage compression as defined by Acemoglu and Pischke (1998, 1999b).

From here onwards for any function \( b(x) \) we will denote the first partial derivative with respect to \( x \) by \( b'(x) \) and the second one by \( b''(x) \). All functions considered here are twice-continuously differentiable.

**Assumption 1.**

i) For all \( \tau \in \mathbb{R}_+ \), \( l(0) = 0, l'(\cdot) > 0, l''(\cdot) < 0, l'''(\cdot) \leq 0 \) and \( x_l'(x) \) is non-decreasing in \( x \).

ii) For all \( a \in A \), \( \lim_{\tau \to 0} l'(a\tau) \to \infty \) and \( \lim_{\tau \to \infty} l'(a\tau) \to 0 \).

iii) \( \epsilon \) is identically and independently distributed with density with density function \( f(\epsilon) \) and full support \([-\bar{\epsilon}, \bar{\epsilon}]\).

iv) \( \eta \) is identically and independently distributed with density \( g(\cdot) \) and full support \([-\bar{\eta}, \bar{\eta}]\), with \( \bar{\eta} \geq \lambda h(\bar{a}, \tau^{**}) + \bar{\epsilon} \). \( G(\cdot) \) is log-concave and symmetric around zero.

Part (i) establishes that the cost function is strictly increasing and convex. Part (ii) assumes that the total and marginal cost at zero is zero and Inada’s type conditions with respect to training. Part (iii) and (iv) are self-explanatory.

### 3.2 Credit Constraints

A typical worker have to finance his training with his own resources and if his wealth is insufficient to finance his training, he may attempt to borrow the difference between the amount he wishes to invest and his wealth. Creditors are not perfectly protected and therefore when facing a defaulting borrower they must forfeit a fraction \( \gamma \in [0, 1] \) of their labor earnings.\(^9\) Enforcement is difficult due to the fact that payments may take different forms, many employers may not pay their workers’s social security taxes or may sub-declare them and many may not keep accurate accounting records of their payments to workers, and so on and so forth. Lenders are competitive and therefore a worker who borrows \( D \), has to repay exactly \( rD \), where \( r \) is the exogenously given interest rate.

To avoid defaulting, creditor limits the loan that a worker can obtain. We assume to keep matters

\(^9\)This is consistent with wage garnishment and penalty avoidance actions like relocation costs, borrowing from loan sharks, renting instead of buying a home, working in the informal sector, which are costly to those who default.
as simple as possible that workers pay whatever they borrow at the end of period one, after they have been paid the first-period wage $w_1$. They cannot use the second-period to pay for the loan. This can be incorporated at a much higher algebraic complexity which does not add much intuition to the problem.

### 3.3 Wage Determination

Here, we turn to the issue of how the worker’s compensation is determined after $(\epsilon, \eta)$ is realized. The key is that the no-trade payoffs enter the bargaining process as outside options instead of as inside options.

The bargaining between a firm and a worker adopted here is Rubinstein’s alternating-offer game with the addition of outside options for both, the firm and worker. At the beginning of each round, the worker is chosen to be a proposer with probability $\alpha$—the worker’s bargaining power—and the firm with probability $1 - \alpha$—the firm’s bargaining power.

If the proposer is the worker, he proposes a wage $w$. The firm can either accept or reject this offer, if it accepts, then the firm gets $\lambda h + \eta - w$, while if it rejects, then the firm and the worker get zero and bargaining either goes to the next round where the firm makes a proposal or chooses to terminate the bargaining process taking its outside option. If bargaining is terminated because the responder takes his or her outside option, the worker gets his outside option which is equal to $\lambda h$.

Note that only the responder is allowed to choose to terminate bargaining. This ensures a unique solution to the bargaining game. Furthermore, because complete information is assumed, the bargaining process ensures that trade is ex-post efficient conditional on that the worker cannot be paid less than wage floor and there are positive firing costs; that is, the firm-worker relationship continues whenever continuing the relationship generates more than separating; i.e., $\lambda h + \eta \geq \lambda h$ and what is generated within the relationship exceeds the maximum between the wage floor and the alternative wage minus the cost of separation; that is, $\lambda h + \eta \geq w$.

It follows from this and the outside-option and inside-option principle that when neither the outside option nor the wage floor binds, each party gets his insider option plus the surplus from continuing the relationship divided according to each party’s bargaining power (hereafter, the output-sharing outcome);\(^\text{10}\) that is, the worker gets $\alpha(\lambda h + \eta)$ and the firm gets $(1 - \alpha)(\lambda h + \eta)$; when only the worker’s outside option binds and it is optimal to continue the relationship, the worker gets the maximum between his outside option, the wage floor and the alternative wage, and the firm gets the total surplus minus the worker’s wage; that is, $\lambda h + \eta - \lambda h$; and when only the firm’s outside

---

\(^{10}\)See, Sutton (1986).
option binds, the worker gets the total surplus from continuing the relationship and the firm gets its outside option 0.

Finally, when the worker and the firm’s outside options are both binding, they are better-off terminating the relationship and each getting his or her outside option because what is generated by continuing the relationship is less than what can be generated if the firm and worker terminate their relationship.

Observe that the worker’s outside option is $\lambda h$ since an outside firm is willing to hire the worker and $\lambda h \geq 0$. Let’s define the lowest match-specific productivity shock, denoted by $\eta_s$, above which the output inside the relationship exceeds the maximum between the outside option, the wage floor and the the alternative wage minus separation costs; that is $\lambda h + \eta_s = \lambda h$, and therefore $\eta_s(t) = 0$.

Similarly, let’s define the lowest match-specific productivity shock, denoted by $\eta_o$, above which the output-sharing outcome occurs; that is, $(1 - \alpha)(\lambda h + \eta_o) \geq 0$ and $\alpha(\lambda h + \eta_o) \geq \lambda h$. It is easy to check that this implies that $\eta_o(t) = \frac{1 - \alpha}{\alpha} \lambda h$. Hence, if $\eta \leq \eta_s$, the relationship is terminated and the worker takes an outside job, which yields an income $\lambda h$ and the firm’s income is 0. If $\eta_o \geq \eta > \eta_s$, the worker remains with the current employer and he is paid his outside option $\lambda h$ and the firm’s return is $\lambda h + \eta - \lambda h$. Finally, if $\eta > \eta_o$, the worker remains with the current employer and the firm and worker share the worker’s productivity according to each other’s bargaining power, which here is assumed to be the same.

It follows from the discussion above that firm’s period-2’s expected payoff is given by:

$$U_f(\tau|w) \equiv E_a\left( \int_{\eta_o}^{\eta_s} (1 - \alpha)(\lambda h + \eta)dG(\eta) + \int_{\eta_s}^{\eta_o} \eta dG(\eta) \right) - \sum_i \tau_{if} \tag{1}$$

and the worker’s period-2 expected payoff is given by:

$$U_l(\tau|w) \equiv E_a\left( \int_{\eta_o}^{\eta} \alpha(\lambda h + \eta)dG(\eta) + \int_{\eta_s}^{\eta_o} \lambda h dG(\eta) \right) - \sum_i \tau_{il}. \tag{2}$$

It follows from equations (1) and (2) that total second-period expected welfare for a firm-worker pair is given by:

$$W(\tau|w) \equiv E_a\left( \int_{\eta_s}^{\eta} (\lambda h + \eta)dG(\eta) + \int_{\eta}^{\eta_s} \lambda h dG(\eta) \right) - \sum_i (\tau_{if} + \tau_{il}). \tag{3}$$

Before ending this section it is worthwhile to discuss the case in which wages in other firms are treated as inside options instead of outside options by the bargaining parties. At a theoretical level, it has been shown by Binmore, Rubinstein, and Wolinsky (1986b) that in order to treat them as inside options, the worker should be able to take the alternative job and keep negotiating with the incumbent employer, which
is a very uncommon situation. Yet more importantly, if that were to be case, then in the absence of wage floors, the second-period wage will be

\[ \lambda h + \beta(\lambda h + \eta - \lambda h - \epsilon) = \lambda h + \beta \eta \]

and second period profits will be

\[ (1 - \beta)(\lambda h + \eta - \lambda h - \epsilon) = (1 - \beta)\eta. \]

Hence, the firm and worker share the return to the match specific shock and the worker gets the full return to training in each possible state. Hence, Becker’s result obtains; i.e., firms do not pay for general training. Thus, a necessary condition for firms to pay for training is that the impact of general training on the inside option is lower than that on the incumbent firm’s product, which is exactly what the wage compression assumption does (see, for instance, Acemoglu and Pischke (1999a) and Booth and Zoega (2004)). Another necessary assumption in this setting is that the firm cannot lower the first-period wage all the way down so as to recover training costs when the worker is young.

Thus, the wage determination procedure adopted here is consistent with how outside options should be dealt correctly according to Binmore et al. (1986b) and it is more in the spirit of Becker’s general human capital theory since the marginal return to training is the same outside as well as inside the firm. Our result do not rely on the idea of wage compression to generate firm-sponsored training, yet many contractual provisions result in wage compression as defined by Acemoglu and Pischke (1999a) and therefore our model combined some of the insights of Acemoglu and Pischke (1999a) with those in Balmaceda (2005).

4 Non-cooperative Training Investments

Provided that a firm hires a worker with a skill level \( a \in A \) in the first period, the firm’s program is

\[
\max_{\tau_f \in \mathbb{R}_+} \{ U_f(\tau) \}
\]

and the firm’s first-order condition is given by:

\[
\frac{\partial U_f(\tau|w)}{\partial \tau_f} = \lambda a_i h_i f(t) \int_{\epsilon(\tau)}^{\epsilon} \int_{\eta_o}^{\eta} (1 - \alpha)dG(\eta)dF(\epsilon) + \int_{0}^{\epsilon(\tau)} \left( \int_{\eta_0}^{\eta} (1 - \alpha)\lambda a l'(a \tau)dG(\eta) + \int_{\eta_0}^{\eta_o} (\lambda a l'(a \tau) - w_r(a, \tau))dG(\eta) \right) dF(\epsilon) - 1 \leq 0.
\]

The worker’s program is

\[
\max_{\tau_l \in \mathbb{R}_+} \{ U_l(\tau|w) \} \text{ subject to } \tau_l \leq r^{-1} w_1(1 - \gamma) + A \}.
\]
and his first-order condition is given by:

\[
\frac{\partial U_l(\tau|w)}{\partial \tau} = \lambda a_l'(a_\tau) \int_{\tau}^{\epsilon} \left( \int_{\eta_o}^{\bar{\eta}} \alpha dG(\eta) + \int_{-\bar{\eta}}^{\eta_o} dG(\eta) \right) dF(\epsilon) + \\
\int_{0}^{\epsilon(\tau)} \left( \int_{\eta_o}^{\bar{\eta}} \alpha \lambda a_l'(a_\tau) dG(\eta) + \int_{\eta_s}^{\eta_o} w_\tau(a, \tau) dG(\eta) + \int_{-\bar{\eta}}^{\eta_s} b_\tau(a, \tau) dG(\eta) \right) dF(\epsilon) + \\
\left( \lambda a_l'(a_\tau) - w_\tau(a, \tau) \right) \int_{0}^{\epsilon(\tau)} (\eta_a + T) g(\eta_s) f(\epsilon) d\epsilon - 1 \begin{cases} 
= 0 & \text{if } \tau_l < r^{-1} w_1 (1 - \gamma) + A, \\
\geq 0 & \text{if } \tau_l = r^{-1} w_1 (1 - \gamma) + A.
\end{cases}
\]

If \( \eta \geq \eta_o \), the worker’s outside option does not bind and the firm and worker are mutually held-up as both receive half of the return to their corresponding investment and half of the return to the investment made by the other player. Hence, in this case there is wage compression in the sense that the wage increases with general training at a lower rate than productivity does. This is very different from the wage compression as defined by Acemoglu and Pischke (1999b), since that refers to the fact that the wage outside of the incumbent firm increases with general training at a lower rate than the worker’s productivity. When \( \eta_s \leq \eta < \eta_o \), the worker is paid the maximum between his outside option \( \lambda h \) and the wage floor \( w \). When \( \epsilon \geq \epsilon(\tau) \), the worker’s outside option exceeds the wage floor \( w \) and therefore the firm is fully held-up as it receives no return to firm-sponsored training since the worker becomes full residual claimant to the return to general training regardless of whom undertakes the investment. In this case the worker’s productivity and his outside option increase with general training at the same rate; i.e., there is no wage compression. When \( \epsilon < \epsilon(\tau) \); i.e., retention requires to pay the worker the wage floor, the worker is partially or fully held-up now since his wage raises with general training at a rate lower than his productivity with the incumbent firm. In this case a wage compression in the sense defined by Acemoglu and Pischke (1999b) occurs since the contractual provisions considered results in a wage that is less sensitive to general training. In this case, the wage floor transforms general training de-facto specific.

When \( \eta < \eta_s \), the worker leaves the firm and this is fully held-up since the worker takes the general human capital with him regardless of whom undertook the investment in it. Whether the worker benefits from general training will depend upon the situation in which he is at. If \( \epsilon \geq \epsilon(\tau) \), the worker’s outside option is a regular employment in a different firm where a-la-Bertrand competition ensures that the worker gets paid his marginal product and therefore he is full residual claimant on the return to general training regardless of who paid for it. When \( \epsilon < \epsilon(\tau) \), his outside option is the alternative wage which is a job where training is less productive than within a regular firm. For instance, this could be a job in the informal sector or self-employment where the general training he received is less productive.

Furthermore, notice that the the contractual provisions considered here results in either a first-order improvement or worsening of the worker and firm’s incentives to invest in training. For the firm, this is different from a marginal increase in any parameter since they affect the thresholds \( (\eta_s, \eta_o) \), which are optimally chosen from the firm’s point of view. For the worker a marginal increase can also have a first-
order effect on his incentives. This is captured by the third term in the worker’s first-order condition in equation (5). This arises because at the separation threshold \( \eta_s \), the worker is better-off staying with the incumbent firm and being paid \( \lambda h \) than leaving and being paid \( w(a, \tau) \). Because an increase in training increases the probability of being retained by the incumbent firm, this increases the worker’s return to self-sponsored training. This term does not arise in the firm’s first-order condition since the separation threshold \( \eta_s \) is chosen optimally by the firm; that is, at \( \eta = \eta_s \), the firm is indifferent between paying the wage floor \( \max\{\lambda h, w\} \) to the worker and letting him go.

### 4.1 Absence of Wage Floors and Alternative Wages

We study first the case in which there are no contractual provision in placed; that is, \( w(a, \tau) = 0 \) for all \((a, \tau)\) and \( T = 0 \). In order to do so, lets define player \( i \)’s best-response function in the absence of a wage floor and an alternative wage as \( BR_i(\tau_j) \equiv \arg\max_{\tau_i \in \mathbb{R}_+} \{U_i(\tau)\} \).

**Proposition 1.** Suppose a non-cooperative regime and that \( w(a, \tau) = 0 \) for all \((a, \tau)\) \( \in A \times \mathbb{R}_+ \) and \( T = 0 \). Then

1. If \( BR_f(r^{-1}w_1(1 - \gamma) + A) > 0 \) and \( BR_i(0) > r^{-1}w_1(1 - \gamma) + A \), then the equilibrium investment is \( \tau^* = BR_f(r^{-1}w_1(1 - \gamma) + A) + r^{-1}w_1(1 - \gamma) + A \), where \( \tau^*_f = BR_f(r^{-1}w_1(1 - \gamma) + A) \) and \( \tau^*_i = r^{-1}w_1(1 - \gamma) + A \).

2. If \( BR_f(r^{-1}w_1(1 - \gamma) + A) \leq 0 \) and \( BR_i(0) > r^{-1}w_1(1 - \gamma) + A \), then the equilibrium investment is \( \tau^* = r^{-1}w_1(1 - \gamma) + A \), where \( \tau^*_f = 0 \) and \( \tau^*_i = r^{-1}w_1(1 - \gamma) + A \).

3. If \( BR_i(0) \leq r^{-1}w_1(1 - \gamma) + A \), then the equilibrium investment is \( \tau^* = BR_i(0) \), where \( \tau^*_f = 0 \) and \( \tau^*_i = BR_i(0) \).

**Proof.** Observe that in this case \( \epsilon(\tau) = 0 \) for all \( \tau \in \mathbb{R}_+ \), and therefore

\[
\frac{\partial U^2_f(\tau|a)}{\partial \tau_f \partial \tau_i} = (1 - \alpha) \left( \lambda a^2 g' \left( a \tau \right) \int_0^\frac{\eta}{\lambda a} dG(\eta) dF(\epsilon) - \frac{1 - \alpha}{\alpha} (\lambda a l'(a \tau))^2 \int_0^{\epsilon} g(\eta_o) dF(\epsilon) \right) < 0.
\]

This implies that \( U_f(\tau) \) is strictly concave and that \( BR_f(\tau_i) \) is a function that decreases with \( \tau_i \). In fact, \( BR_f'(\tau_i) = -1 \).

Next, observe that

\[
\frac{\partial U^2_f(\tau|a)}{\partial \tau_f \partial \tau_i} = \lambda a^2 g'' \left( a \tau \right) \int_0^\frac{\eta}{\lambda a} \left( \int_0^{\lambda a \tau} \alpha dG(\eta) + \int_{-\frac{\eta}{\lambda a}}^0 dG(\eta) \right) dF(\epsilon) + \frac{(1 - \alpha)^2}{\alpha} (\lambda a l'(a \tau))^2 \int_0^{\epsilon} g(\eta_0) dF(\epsilon).
\]

Observe that \( \frac{\partial U^2_f(\tau|a)}{\partial \tau_i \partial \tau_i} < 0 \), since \( l''(\cdot) \leq 0 \) and \( g'(\eta_o) < 0 \) due to the fact that \( G(\cdot) \) is single-peaked at zero. In addition,

\[
\lim_{\tau_i \to 0} \frac{\partial U_i(\tau)}{\partial \tau_i \partial \tau_i} < 0, \quad \lim_{\tau_i \to 0} \frac{\partial U_i(\tau)}{\partial \tau_i} > 0 \quad \text{and} \quad \lim_{\tau \to \infty} \frac{\partial U_i(\tau)}{\partial \tau_i} < 0.
\]
Proposition 2. 

BR such that zero and hiring an untrained worker or closing down. Thus, the firm cannot recoup its investment by paying This readily follows from the fact that a firm can always ensure itself a payoff of at least zero by investing small so that the firm’s marginal return to training evaluated at

\[ \tau \]

limits the worker’s investment, it becomes optimal for the firm to invest since

\[ \alpha \]

otherwise, only the worker invest in training and fully bears the cost of it. The reason stands for the fact that the firm gets a surplus sharing outcome occurs and the full return when his outside option binds. Hence, ceteris-paribus, the worker gets a larger expected share of the return to human capital than the firm does, yet the resource constraint may limit his ability to invest as much as he would like to. When the resource constraint strongly limits the worker’s investment, it becomes optimal for the firm to invest since \( r^{-1}w_1(1-\gamma)+A \) is sufficiently small so that the firm’s marginal return to training evaluated at \( \tau = r^{-1}w_1(1-\gamma)+A \) is greater than 1. Otherwise, only the worker invest in training and fully bears the cost of it.

In the first period, firms compete for workers in a Bertrand-like fashion and therefore total firm’s expected profits should be zero. Because an untrained worker’s productivity in period 1 is given by \( \lambda h(a,0) \), Bertrand-like competition implies that the first-period wage is equal to \( w_1^* = \lambda h(a,0) + E(\epsilon) + U_f(\tau^*) \geq 0 \). This readily follows from the fact that a firm can always ensure itself a payoff of at least zero by investing zero and hiring an untrained worker or closing down. Thus, the firm cannot recoup its investment by paying workers less than the their marginal product as untrained workers.

Lets define \( A_f(\gamma) \) as the largest \( A \) such that \( BR_f(r^{-1}w_1(1-\gamma)+A) \geq 0 \) and \( A_l(\gamma) \) as the smallest \( A \) such that \( BR_l(0) \geq r^{-1}w_1(1-\gamma)+A \). Hence, we have the following result.

**Proposition 2.**

i) If \( A < A_f(\gamma) \), then \( \tau_f^* > 0 \) and \( \tau_l^* > 0 \) and firms pay for firm-sponsored general training. \( \tau_f^* \) falls with \( A \) and raises with \( \gamma \), \( \tau_f^* \) rises with \( A \) and falls with \( \gamma \) and \( \tau^* \) is independent of \( A \) and \( \gamma \).
ii) If $A_f(\gamma) > A \geq A_f(\gamma)$, then $\tau_f^* = 0$ and $\tau_i^* > 0$ and $\tau^* = \tau_i^*$ rises with $A$ and falls with $\gamma$.

iii) If $A \geq A_f(\gamma)$, then $\tau_f^* = 0$ and $\tau_i^* > 0$ and $\tau^* = \tau_i^*$ is independent of $A$ and $\gamma$.

iv) If $A_f(\gamma) > A \geq A_f(\gamma)$, then $w_i^*$ falls with $A$ and rises with $\gamma$, otherwise is independent of $A$ and $\gamma$, and $r^{-1}w_i^*(1 - \gamma) + A$ increases with $A$ and falls with $\gamma$.

v) $A_i(\gamma)$ rises with $\gamma$ and $A_f(\gamma)$ also rises with it.

**Proof.** Observe that

$$\frac{\partial w_1}{\partial A} = U'_f(\tau^*) \frac{\partial \tau^*}{\partial A} \geq 0 \quad \text{and} \quad \frac{\partial(r^{-1}w_1(1 - \gamma) + A)}{\partial A} = r^{-1}U'_f(\tau^*) \frac{\partial \tau^*}{\partial A} (1 - \gamma) + 1$$

and

$$\frac{\partial w_1}{\partial \gamma} = U'_f(\tau^*) \frac{\partial \tau^*}{\partial \gamma} \leq 0 \quad \text{and} \quad \frac{\partial(r^{-1}w_1(1 - \gamma) + A)}{\partial \gamma} = r^{-1}U'_f(\tau^*) \frac{\partial \tau^*}{\partial \gamma} (1 - \gamma) - r^{-1}w_1$$

Observe that $\tau^*$ increases with $A$ and falls with $\gamma$ if and only if $BR_f(r^{-1}w_1(1 - \gamma) + A) \leq 0$ and $BR_i(0) > r^{-1}w_1(1 - \gamma) + A$; otherwise $\tau^*$ is independent of $A$. This follows from noticing that

$$\frac{\partial \tau^*}{\partial A} = r^{-1}U'_f(\tau^*) \frac{\partial \tau^*}{\partial A} (1 - \gamma) + 1 \Rightarrow \frac{\partial \tau^*}{\partial A} = \frac{1}{1 - r^{-1}(1 - \gamma)U'_f(\tau^*)} > 0$$

and

$$\frac{\partial \tau^*}{\partial \gamma} = r^{-1}U'_f(\tau^*) \frac{\partial \tau^*}{\partial \gamma} (1 - \gamma) - r^{-1}w_1 \Rightarrow \frac{\partial \tau^*}{\partial A} = \frac{-r^{-1}w_1}{1 - r^{-1}(1 - \gamma)U'_f(\tau^*)} < 0.$$  

Furthermore, whenever $BR_f(r^{-1}w_1(1 - \gamma) + A) \leq 0$ and $BR_i(0) > r^{-1}w_1(1 - \gamma) + A$, $U'_f(\tau^*) \leq 0$ and, therefore, $w_1$ falls with $A$ and rises with $\gamma$, yet it follows from the equations above that $r^{-1}w_1(1 - \gamma) + A$ increases with $A$ and decreases with $\gamma$. Let’s define $A_f(\gamma)$ as the largest $A$ such that $BR_f(r^{-1}w_1(1 - \gamma) + A) \leq 0$ and $A_i(\gamma)$ as the smallest $A$ such that $BR_i(0) \geq r^{-1}w_1(1 - \gamma) + A$. The existence of $A_i(\gamma)$ follows from the fact that $BR_i(\cdot)$ decreases continuously with $\tau_3$; at a rate equal to -1 and $r^{-1}w_1(1 - \gamma) + A$ rises with $A$.

Because $BR_f(x) > BR_f(x)$, $\forall x$, $A_i(\gamma) > A_f(\gamma)$, $\forall \gamma \in [0, 1]$. Also observe that $A_i(\gamma)$ rises with $\gamma$.

This proposition shows firms invest only on the general training of workers whose personal wealth is sufficiently small, while worker always invest in their general training. The reason is that firm- and self-sponsored training are perfect substitutes and therefore when workers ability to undertake training on their own is sufficiently constraint, firms’ return to training becomes positive due to the fact that when the surplus sharing outcome occurs, firms appropriate a share $1 - \alpha$ of the return to training.

In addition, this result shows that general training is independent of a worker’s personal wealth level is either small or large, while it increases with it when this is between $A_f(\gamma)$ and $A_i(\gamma)$. The reason stands for
the fact that when the wealth level is greater than \( A_l(\gamma) \), the worker’s ability to invest is no constraint by the resources at his disposal. In contrast, when personal wealth is below \( A_f(\gamma) \), the firm also invest in training and since firm- and self-sponsored training are perfect substitutes, for each extra dollar that a worker invests on training the firm reduces its investment in exactly that amount. When personal wealth is between \( A_f(\gamma) \) and \( A_l(\gamma) \), the firm does not invest in training and the workers invest all his first-period wealth on training and therefore as the personal wealth rises, general training rises with it.

The more severe are financial frictions (the higher is \( \gamma \)), the lower the personal wealth threshold below which the firm invest in the general training of the worker. Hence, as financial development improves, ceteris-paribus, we should observe more self-sponsored training.

The proposition also shows that firms pay for the training of poor workers in conjunction with them and do not pay for the training of rich workers. Furthermore, richer individuals invest more on training than poorer individuals. For wealth levels higher than the lowest wealth level satisfying \( B_l(0) = r^{-1}w_1(1 - \gamma) + A \), the total amount invested in self-sponsored training is independent of the wealth level as well as the level of financial development. While for wealth levels below \( A(\gamma) \) and above the lowest wealth level such that \( BR_f(r^{-1}w_1(1 - \gamma) + A) \geq 0 \), the total investment increases with \( A \) and decreases with \( \gamma \), since the firm does not invest in general human capital and the worker invests as much as possible. For wealth levels below the lowest wealth level such that \( BR_f(r^{-1}w_1(1 - \gamma) + A) \geq 0 \), self-sponsored training rises with \( A \) and falls with \( \gamma \) and firm-sponsored training falls with \( A \) and rises with \( \gamma \) in such a way to compensate exactly for the drop in self-sponsored training. The reason is that investments are perfect substitutes in the production technology and therefore the firm’s best response slope is equal to \(-1\).

### 4.2 Presence of Wage Floors and Alternative Wage

In order to ensure that the worker’s first-order condition is necessary and sufficient, the following is assumed from here onwards.

**Assumption 2.** \( U_i(\tau|w) \) is quasi-concave in \( \tau_i \).

Let’s define player \( i \)’s best-response function in the presence of a wage floor and alternative wage as \( BR_i(\tau_i|w) \equiv \arg\max_{\tau_i \in \mathbb{R}_+} \{ U_i(\tau|a) \} \). It is worthwhile to understand the next result to notice that for any \( \tau_i \), we can have \( BR_f(\tau|w) \gtrless BR_f(\tau_i) \). To see this, notice that

\[
\frac{\partial U_f(\tau|w)}{\partial \tau_f} - \frac{\partial U_f(\tau)}{\partial \tau_f} = a'l'(a\tau) \int_{\epsilon}^{\epsilon(\tau)} \left( \int_{\max\{\eta_o, \eta_s(w)\}}^{\eta_o(w)} (\alpha \lambda - \lambda) dG(\eta) \right) \left( \int_{\min\{\eta_o, \eta_s(w)\}}^{\eta_s(w)} (1 - \alpha) \lambda dG(\eta) \right) dF(\epsilon) \geq 0
\]
A contractual provision that establishes a wage floor that is not highly sensitive to general human capital, will increase, ceteris-paribus, the firm’s incentives to provide firm-sponsored training. There two counter-weighting forces. On the one hand, the wage floor decreases the probability that the surplus-sharing occurs, which, ceteris-paribus, lowers the firm’s incentives to invest in training. On the other hand, when the wage floor binds the firm gets a positive return to training whenever \( \lambda < \lambda^* \), which, ceteris-paribus, rises the firm’s incentives to invest in training. This is known as the wage-compression effect since Acemoglu and Pischke (1999a). When \( \lambda \) is small, the latter effects dominates the former and therefore, ceteris-paribus, the firm’s incentives to train rise. Hence, when the wage compression effect is sufficiently strong, a contractual provision that establishes a wage floor increases the firm’s incentives to provide training.

Next, observe that for any \( \tau_f \), \( BR_l(\tau_f|\omega) \gtrless BR_l(\tau_f) \). This follows from the following

\[
\frac{\partial U_l(\tau|\omega)}{\partial \tau_l} - \frac{\partial U_l(\tau)}{\partial \tau_l} =
\begin{align*}
&\alpha^* \int_{\tau}^{\infty} \left( \int_{\max\{\eta_o(\omega),\eta_s(\omega)\}}^{\eta_o(\omega)} (\Delta - \alpha \lambda) dG(\eta) + \int_{\eta_s(\omega)}^{\max\{\eta_o(\omega),\eta_s(\omega)\}} (\Delta - \lambda) dG(\eta) + \\
&\int_{\min\{\eta_o(\omega),\eta_s(\omega)\}}^{\eta_s(\omega)} (\lambda_b - \alpha \lambda) dG(\eta) + \int_{\eta_s(\omega)}^{\min\{\eta_o(\omega),\eta_s(\omega)\}} (\lambda_b - \lambda) dG(\eta) + \\
&\lambda - \Delta \right) (\eta_s(\omega) + T) g(\eta_s(\omega)) \right) dF(\epsilon) \gtrless 0.
\end{align*}
\]

The effect of contractual provisions that result in a wage floor and in the presence of an alternative wage has an ambiguos effect on workers’ incentives to invest on general training. On the one hand, when the wage floor binds, the worker’s marginal return to training rises when \( \lambda < \lambda^* \) is sufficiently large and falls otherwise. The reason is that in the absence of wage floors, the worker gets \( \alpha \) of the return to his investment when the surplus-sharing outcome occurs and the full return to it otherwise, while in the presence of a wage floor, the worker’s marginal return to training whenever he stays with the incumbent firm is \( \lambda_b \). On the other hand, in the absence of wage floors, when the worker leaves the current firm, he gets the full return to training, while in the presence of a wage floor, his marginal return to training is \( \lambda_b \). There is a third effect which is captured by the last term. This arises because an increase in training in the presence of a wage floor, decreases the probability that the wage floor binds. This raises the worker’s return to training since in the states where the wage floor does not bind the worker gets the full return to training while when the wage floor binds he gets a marginal return equal to \( \lambda < \lambda^* \).

Overall, contractual provisions that result in wage floors and alternative payments have counterweighing forces. On the one hand, they compress the wage structure and therefore, as shown by Acemoglu and Pischke (1999a), this increases firms’ incentives to provide general training, while it decreases workers’ incentives to invest in general training. On the other hand, wage floors change the probability that a worker leaves the firm, which lowers the firm’s marginal return to training and increases the worker’s marginal return to it. Third, they decrease the probability that the surplus sharing outcome occurs which has an ambiguous effect.
on the worker and firm’s incentives to train. In the case in which \( \alpha \) is large, the firm’s return to training increases and the worker’s return to it decreases and vice-versa.

**Proposition 3.** Suppose that \( \overline{w}(a, \tau) > 0 \) for some \((a, \tau) \in A \times \mathbb{R}_+ \) and \( T \geq 0 \). If \( U_i(\tau | w) \) is quasi-concave, then

i) If \( BR_f(r^{-1}w_1(1-\gamma)+A|w) > 0 \) and \( BR_i(BR_f(r^{-1}w_1(1-\gamma)+A|w)) > r^{-1}w_1(1-\gamma)+A \), then the equilibrium investment is \( \tau^* = \tau^*_f + r^{-1}w_1(1-\gamma)+A \), where \( \tau^*_f = BR_f(r^{-1}w_1(1-\gamma)+A|w) \) and \( \tau^*_i = r^{-1}w_1(1-\gamma)+A \).

ii) If \( BR_f(r^{-1}w_1(1-\gamma)+A|w) \leq 0 \) and \( BR_i(0|w) > r^{-1}w_1(1-\gamma)+A \), then the equilibrium investment is \( \tau^* = r^{-1}w_1(1-\gamma)+A \), where \( \tau^*_f = 0 \) and \( \tau^*_i = r^{-1}w_1(1-\gamma)+A \).

iii) If \( BR_i(0|w) \leq r^{-1}w_1(1-\gamma)+A \), then the equilibrium investment is \( \tau^* = BR_i(0) \), where \( \tau^*_f = 0 \) and \( \tau^*_i = BR_f(BR_i(0)) \) if \( BR_f(BR_i(0)) < 0 \) and \( \tau^* = BR_f(\tau^*_f) + \tau^*_i \), where \( \tau^*_f = BR_i(BR_f(\tau^*_i)) \) and \( \tau^*_i = BR_f(\tau^*_i) \).

Otherwise, if \( BR_f(r^{-1}w_1(1-\gamma)+A|w) \leq 0 \) the equilibrium investment is \( \tau_f + \tau_i = 0 + r^{-1}w_1(1-\gamma)+A \), while if \( BR_f(r^{-1}w_1(1-\gamma)+A|w) > 0 \), \( \tau_f + \tau_i = \tau_f^* + r^{-1}w_1(1-\gamma)+A \), where \( \tau_f^* = BR_f(r^{-1}w_1(1-\gamma)+A|w) \).

**Proof.** Observe that the firm’s first-order condition re-writes as follows:

\[
\frac{\partial U_f(\tau | w)}{\partial \tau_f} = al'(a\tau) \int_{\epsilon(\tau)}^{\epsilon_0} \left( (1-\alpha)\lambda dG(\eta)dF(\epsilon)+ \right. \\
\left. al'(a\tau) \int_{\epsilon(\tau)}^{\epsilon_0} \left( \int_{\eta_0}^{\eta} (1-\alpha)\lambda dG(\eta)+ \int_{\eta_s}^{\eta_0} (\lambda-\lambda) dG(\eta) \right) dF(\epsilon) - 1 \leq 0, \right)
\]

where \( \lambda = \lambda_m \) if \( m(a, \tau) \geq b(\tau, a) \) and \( \lambda = \lambda_b \) if \( m(a, \tau) < b(\tau, a) \).

Observe that in this case \( \epsilon(\tau) \geq 0 \) and therefore

\[
\frac{\partial U^2_f(\tau | a)}{\partial \tau_f \partial \tau_f} = a^2 l''(a\tau) \int_{\epsilon(\tau)}^{\epsilon_0} \left( (1-\alpha)\lambda dG(\eta)dF(\epsilon)- \left( \lambda al'(a\tau) \right)^2 \int_{\epsilon(\tau)}^{\epsilon_0} \frac{(1-\alpha)^2}{\alpha} g(\eta_0)dF(\epsilon)+ \\
\left( al'(a\tau) \right)^2 \left( \int_{\eta_0}^{\eta} (1-\alpha)\lambda dG(\eta)+ \int_{\eta_0}^{\eta_0} (\lambda-\lambda) dG(\eta) \right) dF(\epsilon)- \\
\left( al'(a\tau) \right)^2 \left( \int_{\eta_0}^{\eta} \frac{1}{\alpha} (\alpha\lambda-\lambda)^2 g(\eta_0)- \left( \lambda-\lambda)^2 g(\eta_0) \right) dF(\epsilon)+ f(\epsilon(\tau)) \int_{\omega}^{\omega} (\lambda-\lambda)^2 dG(\eta) \right) \leq 0.
\]
Next, observe that
\[
\int_{0}^{\epsilon(\tau)} (\frac{d}{d\epsilon} (g(\eta_s + T)g(\eta_s) + f(\epsilon) \int_{\eta_s}^{\eta} \alpha \lambda dG(\eta))) d\eta = 0
\]
Using the mean-value theorem for integrals, there exists \( z \in (0, \epsilon(\tau)) \) such that
\[
\int_{0}^{\epsilon(\tau)} (g(\eta_s) + (\eta_s + T)g'(\eta_s)) f(\epsilon) d\epsilon
\]
Hence, the third term in equation (11) is positive.
Next observe that

\[
\frac{\partial U_l(\tau|w)}{\partial \tau_l} - \frac{\partial U_f(\tau|w)}{\partial \tau_f} = \lambda a l'(a^\tau) \int_{\epsilon(\tau)}^{\epsilon(\tau)} \left( (2\alpha - 1) \int_{\eta_0}^{\eta} dG(\eta) + \int_{\epsilon(\tau)}^{\eta_0} \lambda dG(\eta) \right) dF(\epsilon) + \\
al l'(a^\tau) \int_{0}^{\epsilon(\tau)} \left( (2\alpha - 1) \int_{\eta_0}^{\eta} \lambda dG(\eta) + \int_{\eta_0}^{\eta} (2\lambda - \lambda dG(\eta) + \int_{\eta_0}^{\eta} \lambda dG(\eta) + \\
(\lambda - \lambda) (\eta_0 + T) g(\eta_0) \right) dF(\epsilon) \leq 0
\]

\[
= a l'(a^\tau) \int_{0}^{\epsilon(\tau)} \int^{\eta_0}_{\eta} \lambda dG(\eta) dF(\epsilon) - a l'(a^\tau) \int_{0}^{\epsilon(\tau)} \left( \int_{\eta_0}^{\eta} 2(\lambda - \lambda) dG(\eta) + \int_{\eta_0}^{\eta} (\lambda - \lambda) dG(\eta) - \\
(\lambda - \lambda) (\eta_0 + T) g(\eta_0) \right) dF(\epsilon) \leq 0.
\]

Partialy differentiating this with respect to \(w\), holding \(\lambda\) constant, one can show that this falls with \(w\) if and only if

\[
\int_{0}^{\epsilon(\tau)} \left( 2(2\lambda - \lambda) g(\eta_0) + (\lambda_b + \lambda - 2\lambda) g(\eta_0) - (\lambda + \lambda) (g(\eta_0) + (\eta_0 + T) g'(\eta_0)) \right) dF(\epsilon) - \\
f(\epsilon(\tau)) \left( \int_{-T}^{0} t \int^{\eta_0}_{\eta} (\lambda - \lambda) dG(\eta) + \int_{\eta_0}^{\eta} (\lambda - \lambda) dG(\eta) \right) \leq 0.
\]

Next, observe that

\[
\lim_{\epsilon(\tau) \to 0} \left( \frac{\partial U_l(\tau|w)}{\partial \tau_l} - \frac{\partial U_f(\tau|w)}{\partial \tau_f} \right) > 0
\]

and

\[
\lim_{\epsilon(\tau) \to 0} \left( \frac{\partial U_l(\tau|w)}{\partial \tau_l} - \frac{\partial U_f(\tau|w)}{\partial \tau_f} \right) = \int_{0}^{\epsilon(\tau)} \left( (2\lambda - \lambda) G(\eta_0) + (\lambda_b + \lambda - 2\lambda) G(\eta_0) + (\lambda - \lambda) (\eta_0 + T) g(\eta_0) \right) dF(\epsilon).
\]

Notice that the last term is positive as shown in equation (12) and the sum of the first and second is also positive. To see this notice that for any \(\lambda \in \frac{1}{2} [\lambda, \lambda_b + \lambda]\), the term multiplying \(g(\eta_0)\) and \(g(\eta_s)\) are both positive. Next, if \(\lambda < \frac{1}{2} \lambda\), the term multiplying \(g(\eta_0)\) is negative and therefore the sum of the two terms is lower than or equal to \(2(2\lambda - \lambda) \max \{ g(\eta_0), g(\eta_s) \} \) + \((\lambda_b + \lambda - 2\lambda) g(\eta_0)\). If \(g(\eta_0) \leq g(\eta_s)\), then

\[
\]

In the first period, firms compete for workers in a Bertrand-like fashion and therefore total firm’s expected profits should be zero. Because an untrained worker’s productivity in period 1 is given by \(\lambda h(a, 0)\), Bertrand-like competition implies that the first-period wage is equal to \(w_1 = \lambda h(a, 0) + E(\epsilon) + U_f(\tau) \geq 0\).
This readily follows from the fact that a firm can always ensure itself a payoff of at least zero by investing zero and hiring an untrained worker or closing down. Thus, the firm cannot recoup its investment by paying workers less than their marginal product as untrained workers. If contractual provision at hand also restricts the first period wage such as it is the case with a minimum wage law, a worker will be able to find a job in period if and only if \( w_1 \geq m(a, 0) \).

Observe that

\[
\frac{\partial w_1}{\partial A} = (U'_f(\tau^*|w) + U'_l(\tau^*|w)) \frac{\partial \tau^*}{\partial A} \geq 0,
\]

\[
\frac{\partial w_1}{\partial \gamma} = (U'_f(\tau^*|w) + U'_l(\tau^*|w)) \frac{\partial \tau^*}{\partial \gamma} \leq 0
\]

and

\[
\frac{\partial w_1}{\partial w} = (U'_f(\tau^*|w) + U'_l(\tau^*|w)) \frac{\partial \tau^*}{\partial w} + \frac{\partial U_f(\tau^*|w)}{\partial w} + \frac{\partial U_l(\tau^*|w)}{\partial w} \leq 0
\]

Lets define \( A(\gamma|w) \) as the largest \( A \) such that \( BR_f(r^{-1}w_1(1 - \gamma) + A|w) \geq 0 \).\(^{11}\) Hence, we have the following result.

**Proposition 4.**

i) For all \( A \) and \( \gamma \), \( \tau^* < \tau^{**} \).

ii) \( \tau^* \) rises with \( A \) and falls with \( \gamma \) and \( \tau^*_f \) falls with \( A \) and rises with \( \gamma \).

iii) If \( A < A(\gamma|w) \), then \( \tau^*_f > 0 \) and firms pay for firm-sponsored general training. Otherwise, \( \tau^*_f = 0 \).

iv) If \( A < A(\gamma|w) \), then \( w_1 \) rises with \( A \) and falls with \( \gamma \).

v) \( A(\gamma|w) < A(\gamma) \).

Recall that a worker with a skill level \( a \) will be eligible for a job in sector \( S \) if and only if the first-period wage \( h(0, a) + E(\epsilon) + U_f(\tau|w) \) exceeds the wage floor \( w \). Due to the envelope theorem the first-period wage rises with skills if and only if

\[
1 + \frac{\partial U_f(\tau|w)}{\partial \tau_f} \frac{\partial \tau_f}{\partial a} + \frac{\partial U_f(\tau|w)}{\partial w} = 1 + c'(\tau_f)(\frac{\partial \tau_f}{\partial a} + 1) \geq 0. \tag{14}
\]

The first-term is the increase in the first-period productivity due to a marginal increase in innate skills. The second-term is the increase in the firm’s expected payoff due to the impact that an increase in skills has over the optimal self-sponsored training. And the third term is the direct effect of an increase in the innate skill level over the firm’s second-period expected payoff. It is easy to check that the first-order condition with respect to \( \tau_f \) implies that in equilibrium this term is equal to \( c'(\tau_f) \). The reason stands for the fact that an increase in innate skills increases the worker’s productivity exactly in the same amount as an equal increase

\(^{11}\) The existence of \( A(\gamma) \) follows from the fact that \( BR_f(\cdot) \) decreases continuously with \( \tau_f \) and \( r^{-1}w_1(1 - \gamma) + A \) rises with \( A \).
in training does, and in equilibrium this is equal to $c'(\tau_f)$. The second term is positive due to part (v) in proposition ??.

It readily follows from equation (14) and proposition ?? that the first-period wage rises with skills. Let denote by $a(w)$ the lowest skill level that is required to be hired in the formal sector; that is, the lowest skill level such that the following holds $h(0, a(w)) + E(\epsilon) + U_f(\tau|a(w), S) = w$. If the wage floor $\bar{w}$ is such that $a(w) \in (0, \bar{a})$, only workers whose initial skills are large enough are eligible for a job in the formal sector, while the least skillful workers are not, yet they are eligible, as well as skillful workers, for a job in the informal sector, since $\delta^I = 0$ and $U_f(\tau^I|a, I) \geq 0$. This readily follows from the fact that a firm in the informal sector can always ensure itself a payoff of at least zero by investing zero and hiring an untrained worker or closing down. Thus, a firm in the informal sector is always willing to hire any worker with initial skills $a \in [0, \bar{a}]$, and when the firm trains the worker, it cannot recoup investment costs by paying the worker less than his marginal product as an untrained worker, $h(0, a) + E(\epsilon)$.

### 4.3 Discussion

In the standard human capital theory as developed by Becker (1964), the analysis of firms’ investment in general training in a competitive labor market is straightforward; due to perfect competition, workers capture the full return to their general human capital and thus firms should not pay for this type of human capital. This is at odds with the empirical evidence. However, more recent theories of firm-provided training can explain why firms provide general training. In particular, Balmaceda (2005), on which this paper builds, considers a Becker’s type of model in which bargaining and employment on the spot market are mutually exclusive and there are match-specific productivity shocks. He shows that general and specific training are strategic complements, despite the fact that they are neither substitutes nor complements in the production technology, and that firms invest in both general and specific training, although at a sub-optimal level.

Observe that in the absence of a match-specific shock, Becker’s result in the informal sector is obtained; that is, there is no firm-sponsored training. The reason stands for the fact that the worker must be paid his marginal productivity in every state of the world and therefore the firm never gets a positive return to training. In contrast, in the presence of a match-specific productivity shock, the output-sharing outcome occurs with positive probability, which implies that the worker is paid more than his outside option, yet only a fraction of his marginal productivity within the firm, and thus the firm is willing to pay for general training.\(^{12}\) Observe also that there could be firm-sponsored training in the formal sector even in the absence of a match-specific productivity shock. The reason stands for the fact that in the states where the worker is paid the wage floor, the firm becomes the full residual claimant on both, the firm and the worker’s return to training. In other words, the wage floor transforms general skills de-facto specific.

\(^{12}\)Becker’s result is consistent bargaining games like Rubinstein’s alternating-offer game. See, Balmaceda (2005) for a more detailed discussion of this.
ture within the firm, pushing it away from the competitive benchmark and favoring skilled workers, it will be profitable for firms to provide workers with general human capital. This is because labor market imperfections make general abilities de-facto specific in the sense that trained workers do not get their full marginal product of training when they switch to another job.\textsuperscript{13} This is different from Balmaceda’s model, since workers are never paid less than their marginal product outside the firm regardless of whether they stay or leave the first-period employer. Thus, Balmaceda’s model predicts firm-sponsored general training in a competitive model, while Acemoglu and Pischke’s model do so in a non-competitive labor market model.

Acemoglu and Pischke (1999b) also study firm- and self-sponsored training when the marginal return to self-sponsored training is independent of the amount of self-sponsored training and vice-versa in a model with labor market frictions, defined as a situation in which a worker who leaves the incumbent firm is not pay his marginal product in the best alternative job. They show that either the firm or the worker invests in general training, but not both. Our model delivers a different result. The difference is due to fact that they treat outside options as inside options, while we treat them as outside options,\textsuperscript{14} they ignore the impact that general training has on the probability of separation and in their model the strategic relationship between the two types of training is exogenous.

The model here combines the intuition of Balmaceda’s model with the wage compression idea on Acemoglu and Pischke (1999a, 2003) augmented with the strategic interaction between firm- and self-sponsored training and endogenous separations. Mainly, firm and workers’ incentives to invest in the informal sector are fully driven by the existence of match-specific productivity shocks combined with the outside option principle, while those in the formal sector considers both the wage compression effect created by wage floors and the output-sharing outcome resulting from the match-specific shocks and the bargaining procedure adopted here. Furthermore, we model the impact of wage floors and EPL on separations, while Acemoglu and Pischke (1999a) and Acemoglu and Pischke (2003) ignore that effect. The main advantage of combining both ideas in one model is that, as the evidence suggests, informal firms have incentives to provide training, while in a pure wage compression model that never happens. Hence, a dual labor market model based only on the wage compression effect will yield the straightforward prediction that wage floors in the formal sector increase firm-sponsored training with regard to the informal sector.

The ideas in this paper are also somewhat related to those in Etienne (2006). He proposes a job-matching model in which workers in more flexible labor markets (that is, markets with little employment protection and low unemployment benefits) tend to invest in general human capital, while in more rigid markets with generous benefits and higher duration of jobs, workers are more inclined to invest in specific training. He

\textsuperscript{13}Chang and Wang (1996) and Katz and Ziderman (1990) make the same point in a slightly different context. They assume that trained workers’ marginal productivity is the incumbent firm’s private information. This implies that trained workers’ outside offers do not fully reflect their productivity. Thus, the incumbent firm can appropriate a share of the the return to general training and therefore the investment in general training is positive and positively related to the probability of a worker staying with the the incumbent employer.

\textsuperscript{14}Treating outside wages as inside options implicitly assumes that a worker, while bargaining with the current employer, can work for a different employer and get paid a positive wage. This is unlikely to be the case, except in a few jobs such as temporary jobs.
focuses on the trade-off between general and specific training, while we focus on the trade-off between firm- and self-sponsored training across sectors and he ignores the dual structure of labor markets.

Hence, in the absence of wage floors and an alternative wage, firms provide and pay for a share of the total investment in general human capita for poor workers, while rich workers are on their own in the sense that firm-sponsored training is nil. This suggests that in countries where lenders’ protection is weak, we should observe more firm-sponsored training and this should be concentrated on low-wealth workers. The results here are consistent with Booth and Bryan (2005). They show that firm-sponsored training is associated with higher wages both in the current and future firms, with some evidence that the impact in future firms is larger. These results are consistent with the human capital theory proposed here, since it is the combination of credit constraints with the fact that the firm gets a positive share of the return to training under certain states that gives rise to positive firm-sponsored training. None of these for itself is able to explain firm-sponsored training when this is studied together with self-sponsored training. Hence, the relatively recent literature on training (see, Balmaceda (2005) and Acemoglu and Pischke (1998, 1999b)) provides a partial or incomplete rationale for firm-sponsored training by suppressing in their analysis self-sponsored training.

Acemoglu and Pischke (1999b) also study firm- and self-sponsored training under the assumption that workers are credit constraint; i.e., wages cannot be negative, and wages are determined by Nash bargaining with inside options; that is, by mean of treating workers’ outside offers as inside options. They show that when the marginal product of training outside of the incumbent firm is lower than inside of it; i.e., wage compression takes place, wages cannot be negative and the worker’s bargaining power is sufficiently small, only the firm invests on training and pays for it; otherwise only the worker invests on it and pays for it. They argue that this can explain why in economies with a more compressed wage structure such as Germany and Sweden, employers should pay for general training, whereas in the UK or United States where apparently wage compression is less intense, workers should do so such as it is the case of vocational training.

Figure 2 shows that in USA and UK firms contribute more to training than their counterparts in Germany and Sweden. Furthermore, in most countries firm- and self-sponsored training co-exist, which is not the case in Acemoglu and Pischke’s (1999b) model.

5 Cooperative Training Investments

In this sub-section, we derive the training profile for each firm-worker pair when training is chosen cooperatively and therefore liquidity constraints as well as the distinction between firm- and self-sponsored training play no role. This can also be understood as the joint maximization training profile that accrues when the firm and worker can commit to a training contract in advance and firms compete so that first-period expected profits are zero.

In this case for any training level \( \tau \) and productivity \( \lambda h + \eta \), trade must be at the efficient level; that is,
separations take place if and only if what is generated by staying together is lower than what can be created by severing the match. Given efficient trading, the efficient investment further requires that $\tau$ maximizes the total expected welfare from the employment relationship. Thus, the training profile $\tau$ maximizes total second-period expected surplus minus training costs; that is,

$$\max_{\tau \in \mathbb{R}_+} \{W(\tau|w)\}.$$ 

Let $\tau^{**}$ be the efficient training profile. It readily follows from assumption 1 that $\tau^{**}$ satisfies the following first-order condition,

$$\lambda a l''(a\tau) \int_{\epsilon(\tau)}^{\ell} dF(\epsilon) + \int_{0}^{\epsilon(\tau)} \left( \int_{\eta_s}^{\ell} \lambda a l'(a\tau)dG(\eta) + \int_{-\ell}^{\eta_s} b_{\tau}(a,\tau)dG(\eta) \right) dF(\epsilon) +$$

$$\left(\bar{w}_a(a, \tau) - \lambda a l'(a\tau)\right) \left( \int_{0}^{\epsilon(\tau)} (b(a, \tau) - \bar{w}(a, \tau))g(\eta_s)dF(\epsilon) +
$$

$$f(\epsilon(\tau)) \int_{-\ell}^{-T} (b(a, \tau) - \bar{w}(a, \tau))dG(\eta) \right) - 1 = 0.$$ 

The first term is the marginal return to training when regardless of the training level chosen neither the wage floor nor the alternative wage binds; that is, the worker’s outside option is to work in another regular firm and produce $\lambda h(a, \tau)$. The second term arises when the worker must be paid $\bar{w}(a, \tau)$ in order to be retained. Because general training increases the marginal productivity within the firm at least as much as it increases $\bar{w}(a, \tau)$, the incentives to invest in general training are, ceteris-paribus, lower than those when there is no wage floor. The third term arises because $\bar{w}(a, \tau)$ places a gap between the surplus inside the relationship when the worker must be paid its outside option in order to be retained (i.e., $\lambda h$) and that when the worker must be paid $\bar{w}(a, \tau)$ in order to be retained. When the aggregate productivity shock is lower than $\epsilon(\tau)$, the worker must be paid the wage floor in order to be retained and therefore in these state the separation rate is inefficiently high with respect to the case in which there is no wage floor. Because general training decreases the probability that inefficient separations occur due to the fact that $\bar{w} - \lambda h$ falls with $\tau$, the incentives to invest in general training are, ceteris-paribus, as least as large as those in the presence of a wage floor and alternative wage. Hence, a-priori the impact of a wage floor and an alternative wage on cooperative training is ambiguous.

Let's denote the optimal investment in training by $\tau^{**}$, which is given by the solution to the first-order
\( \lambda a l'(a \tau) - al'(a \tau)(\lambda - \lambda_b) \int_0^{\epsilon(\tau)} G(\eta_b(\tau, a, \epsilon))dF(\epsilon) + \)
\[
\frac{a l'(a \tau)(\lambda - \lambda) \left( \int_0^{\epsilon(\tau)} (b(a, \tau) - \underline{w}(a, \tau))g(\eta_a)dF(\epsilon) + f(\epsilon(\tau)) \int_{-\tilde{\eta}}^{-T} (b(a, \tau) - \underline{w}(a, \tau))dG(\eta) \right)}{1} = 0.
\]

Let's denote the solution to this first-order condition by \( \tau_b \) when \( \underline{w}(a, \tau) = b(a, \tau) \), by \( \tau_m \) when \( \underline{w}(a, \tau) = m(a, \tau) \) and by \( \tau_o \) when \( \underline{w}(a, \tau) = 0 \). Let's also define the following skills thresholds \( a_{b,o} \), the lowest ability level such that \( \epsilon(\tau_b) = 0, a_{m,o} \) the lowest ability level such that \( \epsilon(\tau_m) = 0, a_b \) the lowest ability level such that \( m(a, \tau_b) = b(a, \tau_b) \), and \( a_m \) the lowest ability level such that \( m(a, \tau_m) = b(a, \tau_m) \).

Hence, we have the following result.

**Proposition 5.** Suppose assumption 1 holds and training is chosen cooperatively. Then,

i) If \( m(a, 0) > b(a, 0) \), \( m(a, \bar{\tau}) < b(a, \bar{\tau}) \) and \( a \leq a_{b,o} \), there exists a skills level threshold, denoted by \( a_m \), such that the optimal investment, denoted by \( \tau^{**} \), is given by \( \tau^{**} = \tau_m \) for all \( a \leq a_m \) and \( \tau^{**} = \tau_b \) for all \( a > a_m, \tau_m < \tau_b \), \( \tau_m \) and \( \tau_b \) increase with \( a \).

ii) If \( m(a, 0) < b(a, 0) \), \( m(a, \bar{\tau}) > b(a, \bar{\tau}) \) and \( a \leq a_{m,o} \), there exists a skills level thresholds, denoted by \( a_b \), such that the optimal investment, denoted by \( \tau^{**} \), is given by \( \tau^{**} = \tau_b \) for all \( a \leq a_b \) and \( \tau^{**} = \tau_m \) for all \( a > a_b, \tau_m > \tau_b \), \( \tau_m \) and \( \tau_b \) increase with \( a \).

iii) If for all \((a, \tau) \in \mathcal{A} \times \mathbb{R}_+, m(a, \tau) \geq b(a, \tau) \) and \( a \leq a_{m,o} \), then \( \tau^{**} = \tau_m \) and \( \tau_m \) increases with \( a \).

iv) If for all \((a, \tau) \in \mathcal{A} \times \mathbb{R}_+, b(a, \tau) \geq m(a, \tau) \) and \( a \leq a_{b,o} \) then \( \tau^{**} = \tau_b \) and \( \tau_b \) increases with \( a \).

v) If \( a > \max\{a_{b,o}, a_{m,o}\} \), then \( \tau^{**} = \tau_o \), \( \tau_o > \max\{\tau_0, \tau_m\} \) and \( \tau_o \) rises with \( a \).

**Proof.** First notice that if \( m(a, 0) > b(a, 0) \) and \( m(a, \bar{\tau}) < b(a, \bar{\tau}) \), then continuity of \( b(a, \tau) - m(a, \tau) \) together with the Intermediate Value Theorem ensures that there exists a training level, denoted by \( \tau_{m,b} \), such that \( m(a, \tau) > b(a, \tau) \) for all \( \tau < \tau_{m,b} \) and \( \underline{w}(a, \tau) = m(a, \tau) \) for all \( \tau \geq \tau_{m,b} \). While if \( m(a, 0) < b(a, 0) \) and \( m(a, \bar{\tau}) > b(a, \bar{\tau}) \), \( m(a, \tau) < b(a, \tau) \) for all \( \tau > \tau_{m,b} \) and \( \underline{w}(a, \tau) = m(a, \tau) \) for all \( \tau \leq \tau_{m,b} \).

Observe that when \( \underline{w}(a, \tau) = b(a, \tau) \), the first-order condition is

\[
\lambda a l'(a \tau) - al'(a \tau)(\lambda - \lambda_b) \int_0^{\epsilon(\tau)} G(\eta_b(\tau, a, \epsilon))dF(\epsilon) + 1 = 0.
\]

Let's denote the solution to this first-order condition by \( \tau_b \). Next observe that when \( \underline{w}(a, \tau) = m(a, \tau) >
\[ b(a, \tau), \text{ the first-order condition is} \]

\[
\lambda a l'(a \tau) - a l'(a \tau)(\lambda - \lambda_b) \int_0^{\epsilon(\tau)} G(\eta_s(\tau, a, \epsilon))dF(\epsilon) - \\
al'(a \tau)(\lambda - \lambda_m)(b(a, \tau) - m(a, \tau)) \left( \int_0^{\epsilon(\tau)} g(\eta_s(\tau, a, \epsilon))dF(\epsilon) + f(\epsilon(\tau))G(-T) \right) - 1 = 0. \tag{18}
\]

Let's denote the solution to this first-order condition by \( \tau_m \). It is easy to check that \( \tau_m \geq \tau_b \). In addition, \( W(\tau|a) \) falls with \( w \). Hence, if \( m(a, 0) > b(a, 0) \) and \( m(a, \bar{\tau}) < b(a, \bar{\tau}) \), \( W(\tau|a)|_{w=b} > W(\tau|a)|_{w=m} \) for all \( \tau \leq \tau_{m,b} \) and the opposite holds for all \( \tau > \tau_{m,b} \). This implies that if \( \tau_m \leq \tau_{m,b} \), the optimal investment, denoted by \( \tau^{**} \), is given by \( \tau^{**} = \tau_m \); otherwise, \( \tau^{**} = \tau_b \). In contrast, if \( m(a, 0) < b(a, 0) \) and \( m(a, \bar{\tau}) > b(a, \bar{\tau}) \), \( W(\tau|a)|_{w=b} < W(\tau|a)|_{w=m} \) for all \( \tau \leq \tau_{m,b} \) and the opposite holds for all \( \tau > \tau_{m,b} \). This implies that if \( \tau_b \leq \tau_{m,b} \), the optimal investment, denoted by \( \tau^{**} \), is given by \( \tau^{**} = \tau_b \); otherwise, \( \tau^{**} = \tau_m \).

Notice that for all \( \tau > \tau_{m,b} \)

\[
\frac{\partial W(\tau|w)}{\partial \tau \partial a} = (l'(a \tau) + a \tau l''(a \tau))(\lambda - (\lambda - \lambda_b) \int_0^{\epsilon(\tau)} G(\eta_s(\tau, a, \epsilon))f(\epsilon)d\epsilon) + \\
\lambda a(\tau l'(a \tau))^2 \left( \int_0^{\epsilon(\tau)} g(\eta_s(\tau, a, \epsilon))dF(\epsilon) + G(-T)f(\epsilon(\tau)) \right) > 0.
\]

Hence, \( \tau_b \) increases with \( a \). This implies that \( \epsilon(\tau_b) \) falls with \( a \), since

\[
\frac{\partial \epsilon(\tau_b)}{\partial a} = (\lambda_b - \lambda)(1 + l'(a \tau)) \left( \tau_b + a \frac{\partial \tau_b}{\partial a} \right) < 0.
\]

This together with the fact that \( \lim_{a \to 0} \epsilon(\tau_b) < 0 \) implies that there exists a skill level \( a_{b,o} \) such that for all \( a \geq a_{b,o} \), \( \epsilon(\tau_b) \leq 0 \).

Next observe that for all \( \tau \leq \tau_{m,b} \)

\[
\frac{\partial W(\tau|w)}{\partial \tau \partial a} = (l'(a \tau) + a \tau l''(a \tau))(\lambda - (\lambda - \lambda_b) \int_0^{\epsilon(\tau)} G(\eta_s(\tau, a, \epsilon))f(\epsilon)d\epsilon) + \\
\lambda a(\tau l'(a \tau))^2 \left( \int_0^{\epsilon(\tau)} g(\eta_s(\tau, a, \epsilon))dF(\epsilon) + G(-T)f(\epsilon(\tau)) \right) + \\
\left( (l'(a \tau) + a \tau l''(a \tau))(\lambda_m - \lambda)(b(a, \tau) - m(a, \tau)) + a \tau (l'(a \tau))^2(\lambda_m - \lambda)(\lambda_b - \lambda_m) \right) + \\
\left( \int_0^{\epsilon(\tau)} g(\eta_s(\tau, a, \epsilon))dF(\epsilon) + f(\epsilon(\tau))G(-T) \right) + \\
a \tau (l'(a \tau))^2(\lambda_m - \lambda)^2(b(a, \tau) - m(a, \tau))(g(-T)f(\epsilon(\tau)) + f'(\epsilon(\tau))G(-T)) \geq 0.
\]

30
Hence, τ_m increases with a. This implies that \( \epsilon(\tau_m) \) falls with a, since

\[
\frac{\partial \epsilon(\tau_m)}{\partial a} = (\lambda_m - \lambda)(1 + l'(a\tau_m))(\tau_m + a\frac{\partial \tau_m}{\partial a}) < 0.
\]

This together with the fact that \( \lim_{a \to \bar{a}} \epsilon(\tau_b) < 0 \) implies that there exists a skill level \( a_{m,o} \) such that for all \( a \geq a_{m,o} \), \( \epsilon(\tau_m) \leq 0 \).

It is easy to check that \( \tau_{m,b} \) falls with a and since \( \tau_m \) and \( \tau_b \) increase with a, there ability thresholds, denoted by \( a_m \) and \( a_b \), such that \( \tau_m \geq \tau_{m,b} \) for all \( a \geq a_m \) and \( \tau_b \geq \tau_{m,b} \) for all \( a \geq a_m \). Hence, if \( m(a,0) > b(a,0) \) and \( m(a,\bar{\tau}) < b(a,\bar{\tau}), \tau^{**} = \tau_m \) for all \( a \leq a_m \) and \( \tau^{**} = \tau_b \) for all \( a > a_b \), while if \( m(a,0) < b(a,0) \) and \( m(a,\bar{\tau}) > b(a,\bar{\tau}), \tau^{**} = \tau_b \) for all \( a \leq a_b \) and \( \tau^{**} = \tau_m \) for all \( a > a_b \).

First, this result shows that for high-ability workers neither the wage floor nor the alternative wage are relevant since \( w(a,\tau^{**}) \leq \lambda h(a,\tau^{**}) \). For these workers the cooperative investment is independent of \( m \) and \( b \) and increasing in skills. For all other workers either the wage floor or the alternative wage or both determine the cooperative investment. Mainly, the cooperative investment is lower than that in the absence of a wage floor and alternative wage. This says that when human capital is chosen cooperatively there is no need to create any contractual provision resulting in wage floor or an alternative wage since this harms the relationship. The reason stands for the fact that any contractual provision that results on a wage floor or on an alternative wage increases separations. Because for certain states the marginal return to training upon a separation is lower than that within the current firm, these excessive separations result in a welfare loss.

Perhaps the best example of an alternative wage is the possibility to be self-employed or to work in the informal sector. In either case, the possibility to take self-employment harms the relationship since the worker with positive probability will take self-employment despite the fact that the marginal return to skills there is lower than under the current firm. This happens when self-employment provides a positive benefit \( b > 0 \). This could be a subsistence income or private benefit measured in dollars from working on his own. If the \( b(a,\tau) \) is interpreted as unemployment benefit, this result suggests that the existence of unemployment benefits harms cooperative human capital investments so long as \( \lambda_b < \lambda \).

### 6 Applications

#### 6.1 Promotion Rules

Prendergast (1993) considers a standard-promotion-practices setting in which firms commit to a wage for promoted workers and workers invest in firm-specific human capital in order to increase their chances for promotion. In his setting, if jobs are similar, then the probability of promotion falls and young workers invest less. Prendergast’s informal argument is that, when jobs are similar, firms employ up-or-out in order to increase the incentive for young workers to invest. His logic is that up-or-out increases the probability an old worker will be assigned to the high-level job by eliminating the firm’s ability to assign old workers to the
low-level job (this part of Prendergast’s argument builds on Kahn and Huberman (1988) which is discussed in detail in the next section). Thus, up-or-out will be used to increase the incentive for young workers to invest.

In order to accommodate a promotion rule, we need to consider the existence of two jobs. We will assume that in one job, called the m-job, the worker’s productivity is given by $m + \lambda_m h(a, \tau)$, with $\lambda_m < \lambda$ and the other job, called the d-job, his productivity is $\lambda h(a, \tau) + \eta$. In the m-job, the marginal return to human capital is lower than in the d-job and the productivity in the d-job subject to an aggregated shocks and to a match specific job.

### 6.2 Revisiting Up-or-out Rules

Kahn and Huberman’s (1988) analysis up-or-out rules as contractual provision to induce specific-human capital investments. In that analysis workers choose whether or not to invest in firm-specific human capital, but directly contracting on the investment is not feasible. They first identify a double-moral-hazard problem that results in underinvestment. That is, ex post firms do not reward old workers who invested when young, and anticipating this workers underinvest. They then show that firms can avoid this underinvestment by committing to up-or-out and a high retention wage. Committing in this way results in firms only retaining old workers who invested when young, and knowing this workers invest.

Kahn and Huberman (1988) propose that up-or-out rules are used as a solution to a moral hazard problem. They suppose that the employer wants a worker to invest in human capital, thereby increasing his productivity. But after he made the investment, the firm may claim that he did not, that his productivity was low, and therefore that it will only pay him a low wage. A commitment to either pay a worker the higher wage or else to fire him can overcome the moral hazard problem if the worker is indeed productive, the firm would prefer to retain him at the higher wage rather than to fire him. Realizing this, the worker would be willing to make the investment. Waldman (1990) extends this idea to consider general human capital rather than only firm-specific capital, and to consider the signaling aspects of firing and retention: a firm which retains a worker signals to other firms that the worker likely has high productivity, thereby inducing other firms to offer this worker a higher wage, and inducing the worker to invest in human capital.

A different line of explanation focuses on the information a firm gains about the worker over time. O’Flaherty and Siow (1992) consider a younger and an older worker, where the productivity of the older worker increases with the quality of the younger worker. Though a young worker’s productivity may exceed the market wage, the firm may prefer to fire him because of the option value of finding a younger worker who is even more productive. More generally, Berglas (1976), Brueckner (1991), and McGuire (1991) consider how the peer-group effect affects the characteristics of a competitive equilibrium when firms hire workers with different skills. They do not, however, consider firing decisions.

In contrast to Kahn and Huberman (1988), we assume commitment to a wage floor rather than a specific wage. We believe this is more realistic since mutually beneficial renegotiation would take place provided
that they do not violate the wage floor restriction. Mainly, when the worker’s share of total surplus exceeds the wage floor, the firm and worker will share the surplus according their bargaining power. Notice also that we focus on general human capital and not specific human capital.

In this case, the alternative wage is to work in a different firm since the wage floor does not constraint the first-period wage and therefore the worker can find a job somewhere else. Hence, given that firms compete for workers in period 1, an up-or-out rule will be optimal if and only if the solution to the following problem is positive.

$$\max_{m(a,\tau)\in\mathbb{R}} \{ W(\tau^+|w) \}$$

6.3 Long-Term Contracts

6.4 Unions

6.5 Dual Labor Markets and Minimum Wages

Now, let’s consider the case in which there is formal sector in which a minimum wage and firing costs are enforced and an informal sector where neither the wage floor nor the firing costs are enforced. In this case, the worker’s productivity in the informal sector is \( b + \lambda_b h(a, \tau) \), with \( b = 0 \) and \( \lambda_b < \lambda \) and the worker’s productivity in the formal sector is \( \lambda h(a, \tau) \). Hence, as is usually assumed the worker’s marginal productivity of his human capital is lower in the informal sector than that in the formal sector.

6.6 Training in the Formal vs Informal Sector

Traditional dual labor-market theories, starting with Lewis (1954), assert that labor markets have a formal and an informal sector. The informal sector is the disadvantaged sector into which workers enter to escape unemployment once they are rationed out of the formal sector where wages are set above market-clearing prices (see, for instance, Harris and Todaro, 1970; Fields, 1990; Stiglitz, 1976). However, Maloney (2004) presents evidence for several Latin American countries that challenges this view and instead interprets the informal sector as an unregulated sector. In addition, whether a worker is employed in the formal or the informal sector is to some extent a matter of choice and not a matter of rationing as traditionally assumed. We take this view of the informal sector rather than the traditional view and model this as an unregulated sector and the formal sector as a regulated one. The literature on dual labor markets has focused mainly on the consequences of this structure on unemployment levels, unemployment spells, capital investments, job tenure and earnings, but not on firm- and self-sponsored training. Although some research has found that firm-sponsored training is lower in the informal sector, we are not aware of any study that discusses the theoretical underpinnings of the relationship between dual labor markets as understood here and training.
levels as well as training composition in each sector.

The existence of a wage floor and an EPL in the formal sector has an ambiguous effect on firms and workers’ best responses. On the one hand, a wage floor transforms technologically general skills de-facto specific in the sense that the firm gets a positive (in this case a full) return to general training in those states where the wage floor binds (that is, when $\epsilon \leq \epsilon(\tau)$ and $\eta^F_G \geq \eta > \eta^F_F$). Ceteris-paribus, this increases firms’ incentives and decreases workers’ incentives to invest in training. On the other hand, the imposition of a wage floor: (i) increases the probability of separation since a higher output is needed in order to be profitable to retain the worker; and (ii) decreases the probability that the output-sharing outcome occurs, since the output with the incumbent firm must be larger so that the worker’s share of it is higher than the maximum between the wage floor and the worker’s outside option. This decreases firms’ incentives and increases workers’ incentives to invest in training.

The literature on the impact of employment protection is now rich and has brought some robust results: severance pay tend to be neutral to the extent that wages are downward flexible and incorporate the expected extra costs for the firm. Pure taxes in case of separation affect the firm differently if these costs affect the outcome of bargaining or if they only affect the job destruction margin. Notably, Mortensen and Pissarides (1999) show that, when new entrants bargain the initial wage, this wage is made lower by the existence of future layoffs taxes, which reduces the adverse effects of employment protection legislation. Overall, they show that both job creations and job destructions are reduced, with ambiguous employment effects. The imposition of EPL, be that severance pay or firing costs, is not neutral here: on the one hand, it decreases the probability that the output-sharing outcome occurs, and, on the other hand, it decreases the probability of separation. The former effect rises firms’ incentives to train and lower workers’ incentives to undertake training, and the latter has the opposite effects. Hence, the impact of an EPL on firm- and self-sponsored training is also ambiguous. Observe that here severance payments are not neutral even when wages are not downward rigid as usually argued, since severance payments affect the probability that the output-sharing outcome occurs due to the fact that $P$ increases a worker’s outside option. The difference stands for the fact that most models assume bargaining with inside options instead of outside options as we do here.

Despite the fact that the impact of a wage floor and EPL over the firm and worker’s best responses are ambiguous, we can show the following.

**Proposition 6.** For any $(a, \omega, P, T)$ we have the following:

i) Either $\tau^F_F \leq \tau^l_F$ and $\tau^F_l \leq \tau^l_F$ or $\tau^F_F \geq \tau^l_F$ and $\tau^F_l \geq \tau^l_F$ or $\tau^F_F \geq \tau^l_F$ and $\tau^F_l \geq \tau^l_F$.

ii) $\tau^F_F + \tau^F_l \geq \tau^l_F + \tau^l_l$.

This proposition shows that for any given innate skills level general training in the formal sector exceeds that in the informal sector. Yet, firm-provided training might be smaller or larger in the formal sector than

---

15 With regard to the unconditional severance pay, this is different from the models that treat the outside option as an inside option by modeling bargaining as Nash-bargaining game.
in the informal sector. This means that the potential decrease in firm-sponsored training due to formality is more than compensated for by an increase in self-sponsored training. Similarly, any potential decrease in self-sponsored training is more than compensated for by an increase in firm-provided training. As a result of this, aggregated general human capital investments for a worker in the formal sector exceeds that for an equivalent worker in the informal sector. Hence, this, together with the fact that only workers for whom their innate skills are greater than \( a(\omega) \) are eligible for a job in the formal sector, implies that \( h(\tau^F, a) > h(\tau^I, a) \).

This shows that despite the fact that there are no exogenous productivity advantages in the formal sector, post-training productivity in the formal sector is greater than that in the informal sector. This is due to the selection as well as the fact that overall incentives for investments in general training are more powerful in the formal sector.

Pinning down the effect of formality in firm- and self-sponsored training is difficult due to the fact that formality does not only shift best responses either upward or downward, but also may affect the strategic relationship between the two types of general training. In what follows we will discuss how formality shifts the best responses and when possible we will provide a full characterization of the equilibrium.

It readily follows from equation (4) that formality shifts the firm’s best response upwards if

\[
\int_{\epsilon(\tau)}^{\bar{\epsilon}} \frac{1}{2} (G(\eta^F_0) - G(\eta^I_0)) f(\epsilon) d\epsilon + \int_{0}^{\epsilon(\tau)} \frac{1}{2} (G(\eta^I_0) - G(\eta^F_0)) + G(\eta^F_0) - G(\eta^F_s) f(\epsilon) d\epsilon \geq 0. \tag{19}
\]

The output sharing effect regards how formality (that is, the existence of a wage floor and an EPL) changes the probability that the firm and worker share the return to training and, the wage compression effect refers to the fact that the wage floor, when binding, makes the firm full residual claimant on the return to training. The output sharing effect is negative since the EPL and the wage floor increase the wage that the incumbent firm must paid the worker in order to retain him; that is, the maximum between the outside option and the wage floor. In contrast the wage compression effect is positive since formality makes the firm full residual claimant. Hence, formality increases, ceteris-paribus, firm-sponsored training when the wage compression effect outweighs the output sharing effect.

It readily follows from equation (5) that formality shifts a worker’s best response upwards if

\[
- \int_{\epsilon(\tau)}^{\bar{\epsilon}} \frac{1}{2} (G(\eta^F_0) - G(\eta^I_0)) f(\epsilon) d\epsilon - \int_{0}^{\epsilon(\tau)} \frac{1}{2} (G(\eta^I_0) - G(\eta^F_0)) + G(\eta^F_0) - G(\eta^F_s) f(\epsilon) d\epsilon + \int_{0}^{\epsilon(\tau)} (\eta^F_s + \delta^F T) g(\eta^F_s) f(\epsilon) d\epsilon \geq 0. \tag{20}
\]

Because the output-sharing effect has a negative impact on the firm’s incentives to invest in training and the wage-compression effect has a positive incentive effect, they have the opposite effect on the worker’s incentives to invest in training; when the firm’s share of the return to training falls (rises), the worker’s share

35
of the return to it increases (decreases). The EPL and the wage floor decrease the probability that output sharing take place, which increases the worker’s return to training since he is more likely to get his outside option. This implies that the worker is more likely to be full residual claimant on the return to training. The wage compression effect plays against the worker’s incentive to undertake training since in those state where he is paid the wage floor, in the absence of it, he would have been paid his outside option. This implies that if he were to be in the informal sector he would have been the full residual claimant on the return to training on those states, while in the formal sector he gets no return to training in those states. The separation effect is the result of the fact that the wage floor imposes a gap between his productivity in the formal sector and that in the informal sector exactly at the match-specific productivity threshold at which separation is optimal. Mainly, when the wage floor is binding there are states where a worker leaves the formal sector despite the fact that he is more productive in this sector than in the informal sector. Because in those states the worker is the full residual claimant to the return to training in the informal sector and he gets no return to training in the formal sector, his return to training rises when he departs from the formal sector to the informal sector. Hence, when the return from the separation effect combined with the return from the output-sharing effect dominates the wage compression effect, formality increases the return to self-sponsored training.

In order to have a taxonomy of the impact that formality has on general human capital investments, we will define a worker as a skillful worker when his skill level is sufficiently large so that $\epsilon(\tau) = 0$ and a worker as an unskillful when his skill level is such that $\epsilon(\tau) > 0$. Thus, skillful workers are those for whom the wage floor never binds and unskillful workers are those for whom the wage floor binds with positive probability. Let’s define the threshold $a(w, P)$ as the lowest skill level such that $w - P \leq h(\tau_f, a)$. If $w \leq P$, then $a(w, P)$ is assumed to be zero. The existence of a skill threshold follows from the fact that general human capital $h(\tau_f, a)$ rises with innate skills $a$ (see, proposition ??).

**Proposition 7.**

i) For skilled workers formality decreases firm-sponsored training, while for unskilled workers formality may either increase or decrease firm-sponsored training.

ii) For skilled workers formality increases self-sponsored training, while for unskilled workers formality may either increase or decrease self-sponsored training.

For skilled workers, equations (19) and (20) show that formality results in a downward shift in the firm’s best response function and in an outward shift in the worker’s best response function. This, together with the fact that in either sector the firm’s payoff function is such that firm- and self-sponsored training are strategic substitutes and the worker’s best-response function is such that firm- and self-sponsored training are strategic complements (see, proposition ??) explains the result above. For unskilled workers, the shifts in the best-responses depend on the LMP in place and therefore it is difficult to characterized the equilibrium under the assumptions made. However, it is worthwhile to discuss how the different policies affect the output-sharing, the wage-compression and the separation effects.
On the one hand, an increase in $P$, holding $\epsilon(\tau)$ constant, decrease the probability that the output-sharing outcome occurs, which in turn decreases the firm’s incentives and increases the worker’s incentives to invest in general training. On the other hand, an increase in $P$, holding $\epsilon(\tau)$ constant, lowers the probability of separation when the wage floor binds. Ceteris-paribus, this increases the probability that the wage-compression effect takes place. In addition, when the wage floor is such that retaining the worker will require to pay him the wage floor (the wage floor is greater than the worker’s outside option), an increase in $P$ does not affect the probability that the output-sharing outcome occurs. This results in an increase in the firm’s incentives and a decrease in the worker’s incentives to invest in training. In addition, an increase in $P$ lowers the probability that retention requires to pay the worker the wage floor (i.e., lowers $\epsilon(\tau)$). This increases the worker’s return to training and decreases the firm’s return to training since the worker is more likely to be full residual claimant, while the firm is less likely to be so.

Provided that the wage floor is such that $\epsilon(\tau) > 0$, an increase in the wage floor has qualitatively the same effect that a decrease in severance pay $P$. Hence, there are several counterweighing forces that makes hard to sign the impact that an increase in the wage floor and a decrease in $P$ has on firm- and self-sponsored training.

An increase in firing costs $T$ increases the firm’s return to firm-sponsored training and decreases the worker’s incentives to undertake training. The reasons stands for the fact that an increase in $T$ has no impact on the output-sharing effect, increases the wage-compression effect since it decreases the probability of separations and may either increase or decrease the separation effect. The latter depends on the slope of the density function at the separation threshold. The impact of $T$ on the wage compression effect is a first-order effect while the one on the separation threshold is second-order effect and therefore the former dominates the latter in the case they have the opposite sign.

The discussion above leads to the following conclusion: a firm in the formal sector is more likely to increase the firm’s incentives to provide general training to unskillful workers than an equivalent firm in the informal sector when: (i) severance payments are small relative to the wage floor; (ii) firing costs are large; and (iii) the wage floor is high.

7 An Application to Dual Labor Markets

Now that an important link between duality and firm- and self-sponsored training has been established, we need to expose the general equilibrium logic of workers’ choice of a sector in which to begin their working careers. However, the model here was not designed to study the optimal choice of a labor market sector because it ignores other important aspects such access to capital markets, restricciones to mobility across sectors, technological differences across sectors, etc... We have ignored these issues in order to focus on the main trade-off which is firms’ incentives vis-a-vis workers’ incentives to invest in general human capital across different sectors when the only difference between them is the enforcement of LMP. Yet, there are
some preliminary insights that are worthwhile to highlight here despite the fact that the model as is does not provide an answer to the choice of a sector question.

Because only workers whose ability exceeds \( a(w) \) will be able to find a job in the formal sector, workers for whom \( a < a(w) \) will work in the informal sector and their training will be \( \tau^I \). A worker whose ability satisfies \( a \geq a(w) \) will choose the formal sector if and only if total career wages minus the cost of self-sponsored training (i.e., his utility), when he begins his working career in the formal sector exceeds those when he begins his working career in the informal sector; that is, when \( W(\tau^F | a, F) \geq W(\tau^I | a, I) \).

Observe that total career wages minus the cost of self-sponsored training are given by:

\[
W(\tau | a) = h(0, a) + E(\epsilon) + h(\tau, a) + E(\epsilon) + \int_{-T}^{\tau} G(\eta) d\eta - \int_{0}^{T} (\eta_s + T) g(\eta_s) F(\epsilon) d\epsilon - c(\tau_f) - c(\tau_l)
\]

and they increase with innate skills since

\[
\frac{dW(\tau | a)}{da} = \frac{\partial U_l(\tau | w)}{\partial \tau_f} \frac{\partial \tau_f}{\partial a} + \frac{\partial U_f(\tau | w)}{\partial \tau_l} \frac{\partial \tau_l}{\partial a} + \frac{\partial W(\tau | a)}{\partial a} \\
= c'(\tau_l) \left( \frac{\partial \tau_f}{\partial a} + 1 \right) + c'(\tau_f) \left( \frac{\partial \tau_l}{\partial a} + 1 \right) + 1 \\
\geq 0
\]

where the inequality follows from proposition ??.

This leads to the following result

**Proposition 8.** Total career wages as well as total utility increase with a worker’s innate skills regardless of the sector in which he works.

Without further former analysis we can conclude that the results in propositions 6 and 8 imply that there are two inefficiencies in the labor market: first, there is a sub-optimal investment in general human capital in both the formal and informal sector due to the hold-up problem and the non-cooperative nature of the choice of investments; and second, workers’ choice of a sector is inefficient; that is, either too many or too few workers choose the formal sector. Given than training incentives differ in the formal and informal sector, this distorts even further the investment in general human capital vis-a-vis the efficient investment level. Hence, the hold-up problem together with a dual labor market affects general human capital through an intensive as well as extensive margin. Furthermore, the fact that human capital is higher in the formal sector suggests that one could argue that a dual labor market might be welfare enhancing.
8 Wage Returns To Firm- and Self-Sponsored Training

Some research points towards substantially larger returns to training financed by the employer. In fact, few studies have been able to document positive returns to individual financed (self-sponsored) training. Booth and Bryan (2002) using British data find no effects on wages from individual financed training. Similarly, Loewenstein and Spletzer (1998), using data from NLSY, show that non-employer financed training yields no wage return. Blundell et al. (1999) also note that employer-provided training has a positive impact on wages whereas training not provided by the employer has an insignificant effect on wages. There are some indications that vocational institute and business school training yield higher returns for the individual when changing employer (Loewenstein and Spletzer, 1998). There is also some evidence that some forms of initial training are inversely related to the starting wage and thus workers somehow pay at least a small share of training costs (see, for instance, Veum, 1999). However, Leuven and Oosterbeek (2008) have argued that the fundamental problem concerning the recovery of the causal effect of training on earnings lies in the correction for selectivity into training. They show that most studies that depend on fixed-effect methods where non-participants are used as a comparison group predict high return estimates. They show that when selectivity is dealt with by exploiting arguably exogenous variation in training participation there are smaller wage effects of firm-sponsored training, but still positive. More specifically, they find a zero wage effect of employer-provided training when focusing on the control group that wanted to attend training but did not do so because of a random event (group III). Not only are the coefficients very close to zero, but the p-values are also very high. Importantly, most of the training courses in their analysis are employer-financed and most lead to a certificate for the individuals. Thus, empirical evidence is mixed with regard to average returns to firm-sponsored training when controlling for selection into training; that is, some studies point to zero average returns while others to positive average returns. Yet, most studies find that the average returns to firm-sponsored training are greater than those to self-sponsored training.

The model predicts that compensation is front-loaded since firms anticipate the rent they will receive in the second period, and thus they are willing to bid to attract a worker in the first period higher than the worker’s first-period productivity. The fact that there are smaller wage effects of firm-sponsored training when selectivity is dealt with by exploiting arguably exogenous variation in training participation is consistent with the model’s prediction that wages are front-loaded.

To explain the evidence documenting a higher wage return to firm-sponsored training than to self-sponsored training, we need to define the wage return first. The wage return to training is given by:

\[ U_t(\tau|w) + c(\tau) - (h(0, a) + E(\varepsilon) + U_f(\tau|w)) \]

where the first two terms are the second period expected wage and the term in parenthesis is the first-period wage. Let's denote the second period wage by \( w_2(\tau) \) and the first-period wage by \( w_1(\tau) \). Hence, explaining the evidence requires to show that the wage return due to an increase in self-sponsored training holding firm-sponsored training constant is lower than the wage return.
due to an equal increase in firm-sponsored training holding self-sponsored training constant; that is,

\[ w_2(\tau^1_l, \tau_f) - w_1(\tau^0_l, \tau_f) < w_2(\tau_l, \tau^1_f) - w_1(\tau_l, \tau^0_f) \]

where \( \tau^1_l - \tau^0_l = \tau^1_f - \tau^0_f > 0 \). Hence, dividing both sides by \( \tau^1_l - \tau^0_l \), taking the limit as \( \tau^1_l \) goes to \( \tau^0_l \) and evaluating at the equilibrium training profile, we need to show that

\[ \left. \frac{\partial (w_2(\tau) - w_1(\tau))}{\partial \tau_l} \right|_{\tau=\tau_l} < \left. \frac{\partial (w_2(\tau) - w_1(\tau))}{\partial \tau_f} \right|_{\tau=\tau_f} . \]

This entails the following:

\[ U^l_{\tau_l}(\tau|a) + c^l(\tau_l) - U^f_{\tau_l}(\tau|a) < U^l_{\tau_f}(\tau|a) - U^f_{\tau_f}(\tau|a) . \]

Due to the optimality conditions that determine \( \tau \) (i.e., \( U^l_{\tau_l}(\tau|a) = 0 \)), this re-writes as follows

\[ c^l(\tau_l) - U^l_{\tau_f}(\tau|a) < U^l_{\tau_l}(\tau|a) . \]

Because \( U^l_{\tau_f}(\tau|a) = U^l_{\tau_l}(\tau|a) + c^l(\tau_l) \) and \( U^f_{\tau_l}(\tau|a) > 0 \), the inequality holds always. This leads to the following result.

**Proposition 9.** Regardless of the worker’s skill level and in which sector he begins his career, the wage return to firm-sponsored training exceeds the wage return to self-sponsored training.

This theoretical result provides support for the empirical evidence showing that firm-sponsored training results in larger wage returns than self-sponsored training. The intuition has to do with the fact this is derived in a model where both firm- and self-sponsored training are optimally chosen in a non-cooperative way and firms compete for young workers in a Bertrand-like fashion. Competition ensures that young workers appropriate the whole surplus and optimality guarantees that both types of training are chosen in such a way that the marginal impact of firm-sponsored training on the first-period wage is zero and the marginal impact of self-sponsored training in the second-period wage is also zero.

9 Conclusions

This paper shows that dual labor markets and firm- and self-sponsored training interact in a non-straightforward manner such that dual labor markets result in general human capital patterns that depend upon the LMP implemented in each country. However, regardless of the LMP implemented, formality rises general human capital, decreases firm-provided training for skilled workers and may either increase or decrease that for unskilled workers. The results here also suggest that the final effect of formality on workers’ ability to adapt to technological progress and labor market changes is positive if we accept the hypothesis that more general hu-
man capital facilitates this process. In addition, the model explains why wage-returns from firm-sponsored training are greater than those from self-provided training, yet this does not mean that wage returns from self-sponsored training are nil.

There are also other interesting conclusions that arise from this analysis. First, having a dual-labor market structure, increases, ceteris-paribus, workers’ incentives to invest in self-sponsored training. The reason stands for the fact that upon separation, workers can find a job in the informal sector where productivity depends on their general human capital investments. Second, the strategic nature of firm- and self-sponsored training in the informal sector is such that for firms self-and firm-sponsored training are strategic substitutes, while for workers, they are strategic complements. This is due to the fact that an increase in firm-sponsored training increases the probability that the worker’s outside option binds and therefore the probability that the worker becomes full residual claimant on the return to general training. Third, the strategic nature of firm- and self-sponsored training in the formal sector is determined by the LMP adopted. Mainly, it might be the case that for workers self-and firm-sponsored training are no longer strategic complements and for firms they are no longer strategic substitutes. The reason stands for the fact that the wage floor transforms general human capital de-facto specific; that is, there are productivity states in which in the absence of a wage floor, the worker is the full residual claimant on the return to general training, while in the presence of a wage floor, the firm is the full residual claimant. Fourth, severance payments are not neutral even when wages are not downward rigid as usually argued. The reason stands for the fact that most models assume bargaining with inside options instead of outside options as we do here. This implies that severance payments affect the probability that the output-sharing outcome occurs due to the fact that an increase in them results in a higher outside option and as a consequence of this they also affect the probability of separation through the optimal training profile. Fifth, in general a formal sector with low severance payments, large firing costs and a large wage floor is more likely to result in an increase in firm-sponsored training vis-a-vis that in the informal sector.

Finally, we do not have training data to correctly test the real impact of dualization on training since we lack cross-country panel data on training by financing source (self- and firm-sponsored training), and it is not trivial to decompose total training between firm- and self-sponsored training when we have data for instance on firm-provided and total training. The reason stands for the fact that in order to obtain a good proxy for self-sponsored training from firm-sponsored and total training, we need wage data to know whether or not firm-provided training costs are transferred to workers via lower wages during training periods. Hence, a test of the results provided here must await for the right data set to be collected. The paper also suggests that the empirical studies that focus on the effect of dualization on firm-sponsored training and conclude that dualization is bad for training are missing an important part of the story since they ignore the impact of dualization on self-sponsored training. Furthermore, the results here points to the need to control for the determinants of turnover in order to estimate the impact of dualization on training. In fact, in the model turnover is related to both training and the degree to which the wage exceeds the mandated minimum wage. Thus, failing to control for turnover biases the estimated impact of formality on training. Indeed, there is
empirical evidence to suggest that the extent of training is both dependent upon and an important determinant of the rate of labor turnover. For instance, Royalty (1996) examines the effect of the predicted probability of job turnover on the probability of receiving training and finds that predicted turnover is significantly related to receiving training.
References


Proof of Proposition 5. Observe that assumption 1 implies the following

i) For all \((\tau, a) \in T \times A\), \(W_{ij} \tau_f (\tau | a) W_{\tau_j \tau_i} (\tau | a) - W_{\tau_j \tau_i} (\tau | a) W_{\tau_i \tau_j} (\tau | a) > 0\).

ii) For all \((\tau, a) \in T \times A\), \(|W_{\tau_i \tau_j} (\tau | a)| > |W_{\tau_j \tau_i} (\tau | a)|\).

iii) For all \(q \in A\), \(\lim_{\tau_i \to q} W_{\tau_i} (\tau | a) < 0\) and \(\lim_{\tau_i \to 0} W_{\tau_i} (\tau | a) > 0\).

Hence, the result in (i) follows from this and the first-order conditions in equation (15). Part (ii) follows from the first-order condition in equation (15), the symmetry and the concavity of the welfare function.

Let's define \(x \in \{w, P, T, a\}\). Using Cramer's rule one can easily show that for any parameter \(x\) the following holds

\[
\frac{\partial \tau^*_i}{\partial x} = -\frac{W_{\tau_i \tau_2} (\tau^* | a) W_{\tau_2 \tau_j} (\tau^* | a) - W_{\tau_2 \tau_j} (\tau^* | a) W_{\tau_j \tau_2} (\tau^* | a)}{W_{\tau_2 \tau_i} (\tau^* | a) W_{\tau_j \tau_i} (\tau^* | a) - W_{\tau_j \tau_i} (\tau^* | a) W_{\tau_i \tau_j} (\tau^* | a)}
\]

Part (iii) follows from this, the fact that \(\frac{\partial h (\tau^* | a)}{\partial a} = 1 + \frac{\partial \tau^*_i}{\partial x} + \frac{\partial \tau^*_j}{\partial x}\) and \(W_{\tau_j} = W_{\tau_i}a\).

Parts (iv) and (v) follow from using the envelope theorem and noticing that

\[
\frac{W(\tau | a)}{\partial x} = G(-T) \frac{\partial (-T)}{\partial x} - \int_0^{\epsilon(\tau)} \left( g(\eta_s + T) \frac{\partial (w - P)}{\partial x} + (\eta_s + T) g'(\eta_s) \frac{\partial (w - P - T)}{\partial x} \right) f(\epsilon) d\epsilon
\]

(A1)

Using the mean-value theorem for integrals, there exists \(z \in (0, w - P)\) such that

\[
\int_0^{\epsilon(\tau)} (g(\eta_s + T) + (\eta_s + T) g'(\eta_s)) f(\epsilon) d\epsilon
\]

\[
= f(z) \int_0^{\epsilon(\tau)} (g(\eta_s) + (\eta_s + T) g'(\eta_s)) d\epsilon
\]

\[
= -f(z) \int_0^{\epsilon(\tau)} \frac{d((\eta_s + T) g(\eta_s))}{d\epsilon} d\epsilon
\]

\[
= -f(z)(\eta_s + T) g(\eta_s) \int_0^{\epsilon(\tau)} > 0.
\]

It readily follows from this and equation (A1) that total welfare rises with \(P\), falls with \(w\) and may either rise or fall with \(T\).

Observe also that

\[
\frac{\partial W(\tau^* | a, F)}{\partial a} > \frac{\partial W(\tau^* | a, I)}{\partial a},
\]

which follows from the fact that

\[
\frac{\partial W(\tau^* | a)}{\partial a} = 1 + \int_0^{\epsilon(\tau)} (\eta_s + T) g(\eta_s) f(\epsilon) d\epsilon > 0.
\]
where the second term is strictly positive for $S = F$ and $\epsilon(\tau) > 0$.

Proof of proposition\textsuperscript{??}. By partially differentiating the first-order conditions in equations (4) and (5) with respect to the corresponding variable, we get that

$$\frac{\partial^2 U_f(\tau|w)}{\partial \tau_f \partial \tau_l} = -B(\tau) - c''(\tau_f)$$  \hfill (A2)

and

$$\frac{\partial^2 U_l(\tau|w)}{\partial \tau_l \partial \tau_f} = B(\tau) - C(\tau) - c''(\tau_l),$$  \hfill (A3)

where

$$B(\tau) \equiv \frac{1}{2} \left( \int_{\epsilon(\tau)}^{\bar{\epsilon}} g(\eta_o) dF(\epsilon) + \int_{0}^{\epsilon(\tau)} (g(\eta_o) - 2g(\eta_s)) dF(\epsilon) \right) + (G((w + P)) - G(-T)) f(\epsilon(\tau))$$

and

$$C(\tau) \equiv \int_{0}^{\epsilon(\tau)} (g(\eta_s) + (\eta_s + T) g'(\eta_s)) dF(\epsilon).$$

By cross-partially differentiating the first-order conditions in equations (4) and (5), we get that

$$\frac{\partial^2 U_f(\tau|w)}{\partial \tau_f \partial \tau_l} = -B(\tau)$$  \hfill (A4)

and

$$\frac{\partial^2 U_l(\tau|w)}{\partial \tau_l \partial \tau_f} = B(\tau) - C(\tau).$$  \hfill (A5)

Using the mean-value theorem for integrals, there exist an $z \in (0, \epsilon(\tau))$ such that $C(\tau)$ can be written as

$$\int_{0}^{\epsilon(\tau)} (g(w - h - \epsilon - P - T) + (w - h - \epsilon - P) g'(w - h - \epsilon - P - T)) f(\epsilon) d\epsilon$$

$$= f(z) \int_{0}^{\epsilon(\tau)} d((w - h - \epsilon - P) g(w - h - \epsilon - P - T)) d\epsilon$$

$$= -f(z) \int_{0}^{\epsilon(\tau)} (w - h - \epsilon - P) g(w - h - \epsilon - P - T) d\epsilon$$

$$= -f(z)(w - h - \epsilon - P) g(w - h - \epsilon - P - T) \int_{0}^{\epsilon(\tau)} > 0,$$

Using the mean-value theorem for integrals, there exists $z \in (0, \epsilon(\tau))$ such that $B(\tau)$ can be re-written as
Proof of Proposition 3. First observe that assumption 1 implies that for all \( i, j \in \{f, l\} \) with \( i \neq j \), \( U_i(\tau|a) \) is such that for all \((\tau) \in T \times A\) the following holds:

\[
\frac{\partial \tau_i}{\partial x} = \frac{-U_{i, x}^i(\tau|a)U_{i, j, \tau_j}^i(\tau|a) + U_{i, j, \tau_j}^i(\tau|a)U_{i, x}^j(\tau|a)}{U_{i, \tau_i}^i(\tau|a)U_{i, j, \tau_j}^i(\tau|a) - U_{i, \tau_j}^i(\tau|a)U_{i, \tau_i}^j(\tau|a)}
\]

It follows from this that the sign of \( \frac{\partial \tau_l}{\partial a} + \frac{\partial \tau_f}{\partial a} \) is equal to the sign of \((B(\tau) - C(\tau))c''(\tau_f) - B(\tau)c''(\tau_l)\). The sign of \( \frac{\partial \tau_f}{\partial a} \) is given by the sign of \(-B(\tau)c''(\tau_l)\) and the sign of \( \frac{\partial \tau_l}{\partial a} \) is given by the sign of \((B(\tau) - C(\tau))c''(\tau_f)\). Hence, (i) if \( B(\tau|F) < C(\tau|F) \), firm-sponsored training decreases and self-sponsored training increases with the skill level; (ii) if \( C(\tau|F) \geq B(\tau|F) \geq 0 \), firm- and self-sponsored training decrease with the skill level and \( \tau_f + \tau_l \) falls with the skill level \( a \); and (iii) if \( B(\tau|F) < 0 \), firm-sponsored training increases and self-sponsored training decreases with the skill level.

Notice also that convexity of \( c(\cdot) \) together with the fact that \( \tau_l > \tau_f \) implies that \( \tau_f^l + \tau_l^l \) falls with the skill level \( a \). It also follows from the discussion above that

\[
\frac{\partial h(\tau, a)}{\partial a} = \frac{\partial \tau_f}{\partial a} + \frac{\partial \tau_l}{\partial a} + 1 = \frac{c''(\tau_l)c''(\tau_f)}{B(\tau)c''(\tau_l) - (B(\tau) - C(\tau))c''(\tau_f) + c''(\tau_l)c''(\tau_f)}.
\]
\[
\frac{\partial \tau_l}{\partial a} + 1 = \frac{B(\tau) c''(\tau_l) + c''(\tau_l) c''(\tau_f)}{B(\tau) c''(\tau_l) - (B(\tau) - C(\tau)) c''(\tau_f) + c''(\tau_l) c''(\tau_f)}
\]

and

\[
\frac{\partial \tau_f}{\partial a} + 1 = \frac{- (B(\tau) - C(\tau)) c''(\tau_f) + c''(\tau_l) c''(\tau_f)}{B(\tau) c''(\tau_l) - (B(\tau) - C(\tau)) c''(\tau_f) + c''(\tau_l) c''(\tau_f)}
\]

where concavity of the payoff functions ensure that all these are positive.

Also notice that

\[
\frac{\partial \tau_f}{\partial a} + \frac{\partial \tau_l}{\partial a} = \frac{- B(\tau) c''(\tau_l) + (B(\tau) - C(\tau)) c''(\tau_f)}{B(\tau) c''(\tau_l) - (B(\tau) - C(\tau)) c''(\tau_f) + c''(\tau_l) c''(\tau_f)}.
\]

**Proof of Proposition 6.** Adding the the first-order conditions in equations (4) and (5) for the formal sector one gets that

\[
c'(\tau_f^F) + c'(\tau_l^F) = 1 + \int_0^{e(\tau)} (\eta_s^F + \delta^F T) g(\eta_s^F) f(\epsilon) d\epsilon. \tag{A6}
\]

Adding the first-order conditions in equations (4) and (5) for the informal sector one gets that

\[
c'(\tau_f^I) + c'(\tau_l^I) = 1. \tag{A7}
\]

Combining equations (A6) and (A7), we get that

\[
c'(\tau_f^F) - c'(\tau_f^I) = c'(\tau_l^I) - c'(\tau_f^F) + \int_0^{e(\tau)} (\eta_s^F + \delta^F T) g(\eta_s^F) f(\epsilon) d\epsilon. \tag{A8}
\]

Because the marginal cost is increasing and the integral on the right-hand side equation is positive, it is easy to see from equation (A8) that if \(\tau_f^F \leq \tau_f^I\), then \(\tau_f^I \leq \tau_f^F\) and if \(\tau_f^I \geq \tau_f^{F^I}\), then \(\tau_f^I \geq \tau_f^F\). This proves part (i).

For part (ii) first suppose that \(\tau_f^F \geq \tau_f^I\). If \(\tau_f^F \geq \tau_f^I\), we are done. Then suppose that \(\tau_f^F < \tau_f^I\). It follows from the fact that \(c''(\cdot) > 0\) and \(c'''(\cdot) \leq 0\) that

\[
c'(\tau_f^F) - c'(\tau_f^I) \leq c''(\tau_f^I)(\tau_f^F - \tau_f^I) \tag{A9}
\]

and

\[
c'(\tau_f^I) - c'(\tau_f^F) \geq c''(\tau_f^I)(\tau_f^I - \tau_f^F). \tag{A10}
\]

This can be re-written as

\[-c'(\tau_f^I) + c'(\tau_f^F) \leq -c''(\tau_f^I)(\tau_f^I - \tau_f^F) \tag{A11}
\]
Adding the inequality in equation (A9) and that in equation (A11), one gets that

\[ c''(\tau_f^I)(\tau_f^F - \tau_f^I) + c''(\tau_l^I)(\tau_l^F - \tau_l^I) \geq c'(\tau_f^F) + c'(\tau_l^F) - c'(\tau_f^I) - c'(\tau_l^I) \]

Substituting equations (A6) and (A7) into the RHS of the last inequality one gets that

\[ c''(\tau_f^I)(\tau_f^F - \tau_f^I) + c''(\tau_l^I)(\tau_l^F - \tau_l^I) \geq \int_0^{\epsilon(\tau)} (\eta_\epsilon^F + \delta^F T)g(\eta_\epsilon^F) f(\epsilon) d\epsilon. \]

Because \( c''(\cdot) > 0 \) and the term on the RHS is positive, it readily follows that

\[ \max\{c''(\tau_f^I), c''(\tau_l^I)\}(\tau_f^F - \tau_f^I + \tau_l^F - \tau_l^I) \geq c''(\tau_f^I)(\tau_f^F - \tau_f^I) + c''(\tau_l^I)(\tau_l^F - \tau_l^I) \geq 0. \]

Because \( \max\{c''(\tau_f^I), c''(\tau_l^I)\} \geq 0, \tau_f^F - \tau_f^I + \tau_l^F - \tau_l^I \geq 0 \) and the result obtains.

When \( \tau_f^F < \tau_f^I \), it readily follows from equation (A8) that \( \tau_l^F \geq \tau_l^I \). Because the proof it is almost identical to the one above, for the sake of brevity is omitted. \( \square \)