Life cycle training and equilibrium unemployment

Arnaud Chéron*
University of Le Mans (GAINS-TEPP) and EDHEC Business School

Anthony Terriau†
University of Le Mans (GAINS-TEPP)

September 23, 2016

Abstract

This paper examines life cycle vocational training investments in the context of a model with search frictions that features skill obsolescence and heterogenous agents. We shed light on some age-dependent externalities. On the one hand, this implies that firms can increase too far from retirement the selection into training programs with respect to what it would be optimal to do. On the other hand, endogenous job creation leads unemployed job finding probabilities to be too low at equilibrium, and also decreasing at the end of the working life. In turn, the latter implies that training externalities are lower for the older workers. We calibrate the model on the french economy and assess the quantitative impact of externalities on employment.

*acheron@univ-lemans.fr

†anthony.terriau@univ-lemans.fr

1 Introduction

Since Becker [1964] it is well known that, in a context of competitive markets, human capital investments are in general efficient. On the opposite, search frictions on the labor market give rise to some inefficiency issues, so that there is a room for an optimal policy to promote vocational training investments. More particularly, Acemoglu [1997] and Acemoglu and Pischke [1999] highlighted the impact of a poaching externality: as general human capital investments can benefit, with some probability, to some future employers, the current firm's private return of training investment is lower than its social return. Recently, Belan and Chéron [2014] also argued that due to a higher job finding rate of workers with a higher general human capital, the vocational training of workers accounts for an additional unemployment externality: the social return of training indeed embodies the fact that unemployed worker with higher human capital will switch faster from home production to market production.

This paper develops a life cycle approach to vocational training investments in the context of search frictions. Our main goal is actually to examine to what extent the impact of skill obsolescence and continuous vocational training externalities is age-dependent. This seems all the more important that life cycle issues for a labor market with search frictions have been pointed out by recent works. Chéron et al. [2011, 2013] for instance emphasized distance-to-retirement effects on workers' flows and showed that there exists an age-specific externality related to the job creation process. Menzio et al. [2015] provided a similar life cycle approach with human capital accumulation to predict US labor market flows and wage growth, but with an exogenous accumulation process. Lastly, Messe and Rouland [2014] built on Chéron et al. [2011] to propose a model with search frictions and endogenous human capital accumulation, in order to examine how training investments are related to age-dependent employment protection. But, as the latter paper deals with specific human capital accumulation, it does not raise any age-dependent externality. In turn, in this paper we examine how search frictions and externalities related to training investments in general human capital can interact each other over the life cycle.

On the one hand, this allows to emphasize that, with respect to what it would be optimal to do, firms typically reduce too far from retirement the entry of workers into training programs. Otherwise stated, some older workers that are not trained would have to be trained when poaching and unemployment externalities are internalized. But on the other hand, it comes that both private and social return of training investments converge to zero when retirement gets closer. The latter is furthermore reinforced in the context of endogenous matching. Indeed, as job matching probabilities decrease at the end of the working life (due to shorter distance-to-retirement), training externalities also collapse. Overall, there exists some opposite forces on externalities as worker is aging. However, our model simulation suggests that to offset the gap between equilibrium and optimal allocations we would primarily need to raise the job creation far from retirement. By doing so, this reduces workers' risk of skill obsolescence, hence lowers future social costs.

The remaining of the paper is organized as follows. The next section presents a partial equilibrium version of the model with exogenous job finding probabilities. We lay out the economic environment and characterize both equilibrium and optimal age-dynamics of training policy to identify related externalities over the life cycle. Then, we extend these results to the endogenous matching case. Additional theoretical results that exhibit the interaction between job creation and training are first provided, and we lastly implement an empirical investigation with a calibration of the model on the french economy. A final section concludes.

2 Search frictions and life cycle training

We aim to develop a life cycle model that features search frictions and general human capital depreciation during unemployment spell, which depends on a "turbulence" parameter in line with what have been pointed out by Ljungqvist and Sargent [1998]. Human capital accumulation relates to firms' endogenous decision to train workers at the time of hiring, as in the paper by Belan and Chéron [2014] with infinite-lived agents. We here consider a finite horizon (retirement), and this leads to an age-dependent selection of workers by firms into training programs, that is due to a distance-to-retirement effect. Yet, in this section we first consider exogenous job finding probabilities in order to characterize the equilibrium and efficient properties of training, and point out the age-dynamics of poaching and unemployment externalities.

2.1 Economic environment

Workers are characterized by their ability level, denoted by a, distributed over the interval $[\underline{a}, \bar{a}]$ according to p.d.f. $f(a)$, and by their age, denoted by t. The model is in discrete time and at each period the older worker generation retiring from the labor market is replaced by a younger generation of the same size (normalized to unity) so that the population on the labor market is constant. We assume that each worker of the new generation enters the labor market at age $t = 0$ and retires at a deterministic age T.

We assume that workers enter the labor market with up-to-date knowledge, so they get the highest level of general human capital, hence related productivity $(1 + \Delta)a$ with $\Delta > 0$. But then they can face skill obsolescence (human capital depreciation) over the life cycle: this occurs during unemployment spells with a per period probability π; if so, productivity is falling to a. Then, at the time of hiring, firms can choose to train workers whose human capital has been depreciated, in order to restore productivity $(1+\Delta)a$ instead of a. This leads firms to bear an instantaneous training cost γ_f. Obviously, this intertemporal decision shall depend on workers' ability, so that the training policy consists of determining an ability threshold that is age-dependent, denoted $\tilde{a}_t$.

Therefore, workers are heterogenous according to three dimensions: (i) ability a, (ii) age t, and (iii) status wrt. skill obsolescence. This implies in particular that we need to distinguish three types of agents at the time of hiring[1]:

- Type-0, with obsolete knowledge and unable for training ($a < \tilde{a}_t$), with productivity a;
- Type-1, able for training ($a \geq \tilde{a}_t$) but with obsolete knowledge, with productivity $(1 + \Delta)a$ once the cost γ_f has been paid;
- Type-2, with up-to-date skills and instantaneous productivity $(1+\Delta)a$ without any additional cost.

Furthermore, at this stage we first consider a partial equilibrium framework where the frictional labor market is featured by exogenous job finding probabilities, constant across ages. We assume that worker's status with respect to skill obsolescence and ability a is perfectly observable by the employer. The probability for an unemployed worker of age t to be employed at age $t + 1$ is assumed to be given by:

- p_0 for individuals with obsolete skills, unable for training if they are hired at the next period ($a < \tilde{a}_{t+1}$)

[1]The firm decide at time t to train a worker with obsolete skills recruited at time $t-1$ if his ability level a is superior or equal to the ability threshold $\tilde{a}_t$.

- p_1 for individuals with obsolete skills, able for training if they are hired at the next period $(a \geq \tilde{a}_{t+1})$
- p_2 for individuals with up-to-date skills.

with $p_2 \geq p_1 \geq p_0$.

2.2 Value functions and Nash bargaining of wages

Let $w_{j,t}(a)$ be the wage, β the discount factor, δ the job destruction probability and b the home production. The expected values of income streams, denoted by $E_{j,t}(a)$ for a worker and $U_{j,t}(a)$ for an unemployed, are defined, $\forall\ t \leq T-1$, by:

Type 0:

$$E_{0,t}(a) = w_{0,t}(a) + \beta\left[(1-\delta)\, E_{0,t+1}(a) + \delta U_{0,t+1}(a)\right] \quad ,\forall\ a < \tilde{a}_t$$

$$U_{0,t}(a) = b + \beta\left[p_0 E_{0,t+1}(a) + (1-p_0)\, U_{0,t+1}(a)\right] \quad ,\forall\ a < \tilde{a}_t$$

Type 1:

$$E_{1,t}(a) = w_{1,t}(a) + \beta\left[(1-\delta)\, E_{1,t+1}(a) + \delta U_{2,t+1}(a)\right] \quad ,\forall\ a \geq \tilde{a}_t$$

$$U_{1,t}(a) = b + \beta \begin{cases} p_1 E_{1,t+1}(a) + (1-p_1)\, U_{1,t+1}(a) & ,\ \forall\ a \geq \tilde{a}_{t+1} \\ p_0 E_{0,t+1}(a) + (1-p_0)\, U_{0,t+1}(a) & ,\ \forall\ a \in [\tilde{a}_t ; \tilde{a}_{t+1}[\end{cases}$$

Type 2:

$$E_{2,t}(a) = w_{2,t}(a) + \beta\left[(1-\delta)\, E_{2,t+1}(a) + \delta U_{2,t+1}(a)\right] \quad ,\forall\ a$$

$$U_{2,t}(a) = b + \beta\left[p_2 E_{2,t+1}(a) + (1-p_2)\,(1-\pi) U_{2,t+1}(a) + (1-p_2)\,\pi \begin{cases} U_{1,t+1}(a) & ,\ \forall\ a \geq \tilde{a}_{t+1} \\ U_{0,t+1}(a) & ,\ \forall\ a < \tilde{a}_{t+1} \end{cases} \right]$$

where we let $E_{j,T}(a) = U_{j,T}(a) \equiv R \ \ \forall a, j$. Value functions for unemployed of type 1 and 2 deserve further discussion. Indeed, it can be the case that the ability of workers is high enough at age t to be trained, but no longer at age $t+1$, so that they switch from type-1 to type-0 from t to $t+1$ (see the expression $U_{1,t}(a)$). Similarly, type-2 workers that remain unemployed and face human capital depreciation can directly switch to the 0-type, if $a < \tilde{a}_{t+1}$. We should also notice that type-1 employed workers switch to type-2, only once they experience an unemployment spell. Obviously, this raises a conventional hold-up issue: once a type-1 worker has been trained, he gets some incentives to ask for wage $w_{2,t}(a)$ instead of $w_{1,t}(a)$ which is lower in equilibrium. This adds another source of inefficiency that have been for instance discussed by Belan and Chéron [2014] in an infinite lived agents context. We choose here to focus on age-dependent externalities due to transferability of general human capital.[2]

Turning to the expected values of filled jobs by a worker of age t and ability a, we have the following value functions, $\forall\ t \leq T-1$:

$$\begin{aligned} J_{0,t}(a) &= a - w_{0,t}(a) + \beta\,(1-\delta)\,J_{0,t+1}(a) && (1)\\ J_{1,t}(a) &= (1+\Delta)a - w_{1,t}(a) + \beta\,(1-\delta)\,J_{1,t+1}(a) && (2)\\ J_{2,t}(a) &= (1+\Delta)a - w_{2,t}(a) + \beta\,(1-\delta)\,J_{2,t+1}(a) && (3) \end{aligned}$$

with $J_{j,T}(a) = 0\ \ \forall a, j$.

Then, we consider standard Nash bargaining of wages, and let α be the bargaining power of workers. Therefore, wages are derived from the following sharing rules:

$$\begin{aligned} (1-\alpha)[E_{0,t}(a) - U_{0,t}(a)] &= \alpha J_{0,t}(a)\\ (1-\alpha)[E_{1,t}(a) - U_{1,t}(a)] &= \alpha[J_{1,t}(a) - \gamma_f]\\ (1-\alpha)[E_{2,t}(a) - U_{2,t}(a)] &= \alpha J_{2,t}(a) \end{aligned}$$

This implies the following wage equations, $\forall t < T-1$:[3]

$$\begin{aligned} w_{0,t}(a) &= \alpha\,[a + \beta p_0 J_{0,t+1}(a)] + (1-\alpha)b && (4)\\ w_{1,t}(a) &= \alpha\,[(1+\Delta)a + \beta p_0 J_{0,t+1}(a) - \gamma_f(1-\beta(1-\delta))] \\ &\quad +(1-\alpha)\,\{b - \beta\delta\,[U_{2,t+1}(a) - U_{0,t+1}(a)]\} && (5)\\ w_{2,t}(a) &= \alpha\,[(1+\Delta)a + \beta p_2 J_{2,t+1}(a)] \\ &\quad +(1-\alpha)\,\{b - \beta\pi\,(1-p_2)\,[U_{2,t+1}(a) - U_{0,t+1}(a)]\} && (6) \end{aligned}$$

Wage equation for type-0 workers is standard: it features the weighted average impact of productivity plus expected profits flows and home production. Wage equation for type-1 shows in addition that workers share the cost

[2]See Appendix A for a variant of the model that deals with the hold-up issue.
[3]See Appendix D for more details

for training through lower wages. It comes also that workers' threat point depends on the expected unemployed value of training $[U_{2,t+1} - U_{0,t+1}]$. The higher the latter, the lower is type-1's threat point hence the wage. This issue is also a determinant of type-2 wages, but is here related to the probability π for type-2 workers to face skill obsolescence.

2.3 Equilibrium training policy

The firm's training policy consists in determining the age-specific ability threshold $\tilde{a}_t$. Above this threshold, at the time of hiring the employer trains any worker that faced human capital depreciation during the unemployment spell. Hence, $\tilde{a}_t$ satisfies the following condition:

$$J_{1,t}(\tilde{a}_t) - \gamma_f = J_{0,t}(\tilde{a}_t) \tag{7}$$

This problem can be solved recursively starting from terminal condition at $t = T - 1$, and it comes that:

$$\begin{aligned} \Delta\tilde{a}_{T-1} &= \gamma_f \\ \Delta\tilde{a}_{T-2} &= \frac{\gamma_f}{\sum_{i=0}^{1}[\beta(1-\delta)]^i} \end{aligned}$$

$$\Delta\tilde{a}_t = \frac{\gamma_f - \beta\delta\sum_{i=0}^{T-3-t}[\beta(1-\delta)]^i\left[U_{2,t+1+i}(\tilde{a}_t) - U_{0,t+1+i}(\tilde{a}_t)\right]}{\sum_{i=0}^{T-1-t}[\beta(1-\delta)]^i}, \forall\ t \leq T-3$$

- For $T-1$, the condition is static and shows that the instantaneous productivity gain must be at least equal to training expenditures.
- For $T-2$, there exists a capitalization effect that depends both on the discount factor and the probability of job destruction, since productivity gain can last two periods with probability $1-\delta$.
- Then, $\forall\ t \leq T-3$, the larger the unemployment gap $[U_{2,t+1} - U_{0,t+1}]$, the lower the wage $w_{1,t}(a)$ and the ability threshold $\tilde{a}_t$. Indeed, while an unemployed worker with up-to-date skills (type-2) can benefit from productivity $(1+\Delta)a$ without incurring any training cost, type-1 workers that faced human capital depreciation agree with wage cuts to become of type-2 in the future. Such a wage cut is as much important as $[U_{2,t+1} - U_{0,t+1}]$ is high, because the latter gives the relative value of training for unemployed workers who undergo skill obsolescence.

2.4 Efficient training policy

The efficient training policy is now derived by considering per worker social values, according to age, ability and types (0, 1 or 2). The social value of an employed worker with up-to-date skills (type 2) is:

$$\tilde{Y}_t(a) = (1+\Delta)a + \beta\left[(1-\delta)\tilde{Y}_{t+1}(a) + \delta Y^{u2}_{t+1}(a)\right] \quad ,\forall\ a$$

where the social value of an unemployed with up-to-date skills (type 2) satisfies:

$$\begin{aligned} Y^{u2}_t(a) \quad &= \quad b + \beta\Big[p_2\tilde{Y}_{t+1}(a) + (1-p_2)(1-\pi)Y^{u2}_{t+1}(a) \\ &\qquad +(1-p_2)\pi\left\{\begin{array}{ll} Y^{u1}_{t+1}(a) & ,\ \forall\ a \geq a^\star_{t+1} \\ Y^{u0}_{t+1}(a) & ,\ \forall\ a < a^\star_{t+1} \end{array}\right.\Big] \end{aligned}$$

Yet, attention must be paid on the fact that if the worker does not find a job and faces human capital depreciation he switches at the next period, either to the type-1 or the type-0 status if its ability is too low. Similarly, we can define $Y^{u1}_t(a)$ for $a \geq a^\star_t$:

$$Y^{u1}_t(a) \quad = \quad b + \beta\left\{\begin{array}{ll} p_1[\tilde{Y}_{t+1}(a) - \gamma_F] + (1-p_1)Y^{u1}_{t+1}(a) & ,\ \forall\ a \geq a^\star_{t+1} \\ \\ p_0\hat{Y}_{t+1}(a) + (1-p_0)Y^{u0}_{t+1}(a) & ,\ \forall\ a < a^\star_{t+1} \end{array}\right.$$

and also $Y^{u0}_t(a)$ for $a < a^\star_t$:

$$Y^{u0}_t(a) \quad = \quad b + \beta\left[p_0\hat{Y}_{t+1}(a) + [1-p_0]Y^{u0}_{t+1}(a)\right] \quad ,\forall\ a < a^\star_t$$

where the social value of an employed worker with obsolete knowledge is given by:

$$\hat{Y}_t(a) = a + \beta\left[(1-\delta)\hat{Y}_{t+1}(a) + \delta Y^{u0}_{t+1}(a)\right] \quad ,\forall\ a < a^\star_t$$

Therefore, from the planner's point of view it is optimal to train an unemployed worker that faced skill depreciation only if $\tilde{Y}_t(a) - \gamma_f > \hat{Y}_t(a)$. This means that there exists an efficient ability threshold. The latter is denoted $a^\star_t$ and solves $\forall\ t : \tilde{Y}_t(a^\star_t) - \gamma_F = \hat{Y}_t(a^\star_t)$. This implies that the efficient training policy is then characterized by:

$$\begin{aligned} \Delta\tilde{a}_{T-1} \quad &= \quad \gamma_f \\ \Delta\tilde{a}_{T-2} \quad &= \quad \frac{\gamma_f}{\sum_{i=0}^{1}[\beta(1-\delta)]^i} \end{aligned}$$

$$\Delta a^\star_t \quad = \quad \frac{\gamma_f - \beta\delta \sum_{i=0}^{T-3-t}[\beta(1-\delta)]^i \left[Y^{u2}_{t+1+i}(a^\star_t) - Y^{u0}_{t++1+i}(a^\star_t)\right]}{\sum_{i=0}^{T-1-t}[\beta(1-\delta)]^i} \ , \ \forall \ t \leq T-3$$

This shows (wrt. to the equilibrium policy) that what matters now is the relative social value of training for the unemployed, as defined by $Y_t^{u2}(a) - Y_t^{u0}(a)$, *ie.* the social value of having up-to-date knowledge when unemployed with respect to the value of having obsolete knowledge without any perspective to be trained. In particular, this takes into account of the fact that type-2 unemployed workers get probability p_2 to find a job (instead of p_0), and once employed they get a social value $\tilde{Y}_t(a)$ (instead of $\hat{Y}_t(a)$).

2.5 Training externalities

We can now compare both equilibrium and efficient outcomes to highlight training externalities. At this stage, we focus on analytical insights by looking at $t = [T-3, T-2, T-1]$. It is then straightforward to see that equilibrium and optimal training policies are characterized by:

$$\Delta\tilde{a}_{T-1} \quad = \quad \Delta a^\star_{T-1} = \gamma_f \tag{8}$$

$$\Delta\tilde{a}_{T-2} \quad = \quad \Delta a^\star_{T-2} = \frac{\gamma_f}{\sum_{i=0}^{1}[\beta(1-\delta)]^i} \tag{9}$$

but

$$\Delta a^\star_{T-3} \quad = \quad \frac{\gamma_f - (a^\star_{T-3} - b)\beta^2\delta(p_2 - p_0)}{\sum_{i=0}^{2}[\beta(1-\delta)]^i + \beta^2\delta p_2} \tag{10}$$

$$\Delta\tilde{a}_{T-3} \quad = \quad \frac{\gamma_f - (\tilde{a}_{T-3} - b)\alpha\beta^2\delta(p_2 - p_0)}{\sum_{i=0}^{2}[\beta(1-\delta)]^i + \alpha\beta^2\delta p_2} \tag{11}$$

It should be first noticed that at the end of the life cycle (from $t \geq T-2$), the equilibrium training policy converges to what it is optimal to do. This is quite intuitive because the potential externalities related to training in general human capital falls short when retirement gets closer. On the opposite, for $t = T-3$, we see that the equilibrium training policy is efficient if and only if $\alpha = 1$, *ie.* if the workers get all the bargaining power. Indeed, two types of externalities are not internalized by firms (whereas they are by workers). By inspecting the two equations for $a^\star_{T-3}$ and $\tilde{a}_{T-3}$, we see a similar capitalization effect that is determined by $\sum_{i=0}^{2}[\beta(1-\delta)]^i$, but for the equilibrium condition α weights two terms:

- $\beta^2 \delta p_2$ which relates to a "poaching externality": general human capital investments can benefit, with some probability, to future employers. Specifically, a worker with up-to-date knowledge fired with probability δ can be re-employed at the next period with probability p_2.

- $p_2 - p_0$ which relates to an "unemployment externality": workers whose ability is high enough for training switch faster from home production to market production (due to a higher job finding probability) and this embodies a social gain that is not valuated by the employers whose bargaining power is $1 - \alpha$.

Therefore, this already emphasizes that externalities related to vocational training are age-dependent. From this first perspective, this suggests that training externalities are lower for the older workers. But another issue is the moment/age at which the horizon effect on training policies starts to be significant and leads selection into training programs to increase. Obviously, this age should depend on whether training externalities are internalized or not. Otherwise stated, it can be the case firms start increasing selection into training programs too far from retirement, with respect to what they should do if evaluating the social gain of training.

2.6 On life cycle effects of training externalities

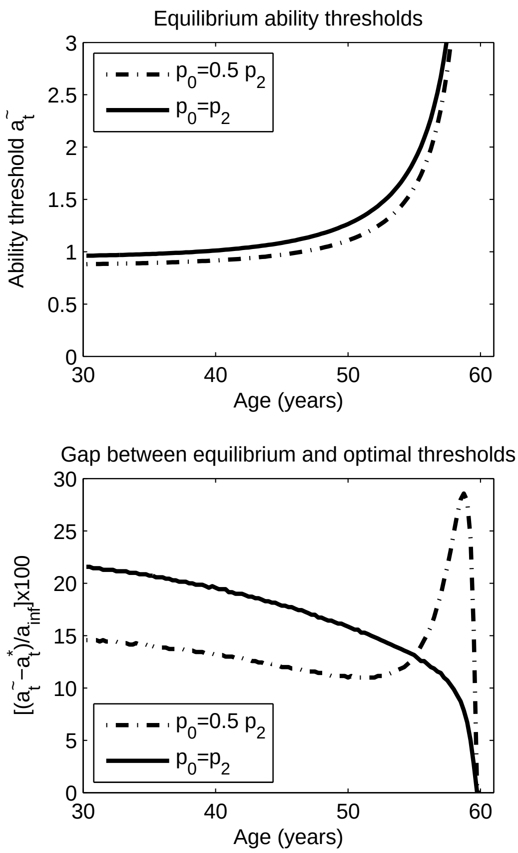
To illustrate this point we run preliminary quantitative simulations of the life cycle dynamics of training selection, on a quarterly frequency basis for the french economy. A first set of parameters is calibrated in a fairly standard way: $\Phi = \{\beta = 0.99, \underline{a} = b = 0.7, \alpha = 0.5, \Delta = 0.1, \pi = 0.5, \delta = 0.0356\}$.[4] Then, it remains three parameters to calibrate, p_2, p_0 and γ_f. We know that in France average unemployment is a about one year. Accordingly, we choose $p_2 = 1/3$ and $p_0 = 1/6$ as a benchmark, but also consider $p_0 = 1/3$ as a variant to illustrate model properties. Lastly, we set $\gamma_f = 2.8$ that is equivalent to one year of the lowest worker's productivity value.[5]

Figure 1 plots the age-dynamics of the productivity thresholds $\tilde{a}_t$, by disentangling cases $p_2 = p_0$ and $p_0 < p_2$. The left panel focuses on the life cycle until 57 years old, while the right panel shows this dynamics until retirement with a sharp increase of the selection in the very end of the working life. This first highlights that equilibrium thresholds are lower with $p_0 < p_2$

[4] Information on that calibration can be found in section 3.3 who implement a detailed quantitative investigation in an extented version of the model.

[5] Again, more attention will be paid in section 3.3 on the calibration of training costs.

Figure 1: Life cycle training policy



because in such circumstances workers valuate the job finding impact of training, hence agree for lower wages ($[U_2 - U_0]$ is relatively higher). The rise of the selection into training programs is then found to be quantitatively significant from 45 years old. This shows a distance-to-retirement (horizon) effect: the private intertemporal value of training investments is decreasing as worker is aging.

Figure 1 also focuses on the efficient training selection (bottom panel), and computes the gap between the equilibrium and optimal training policies, with a gap between ability thresholds expressed in percentage of the lowest threshold $\underline{a}$, *ie.* $\frac{\tilde{a}_t - a_t^\star}{\underline{a}} \times 100$. This shows that in our benchmark case with both poaching and unemployment externalities (that is with $p_0 < p_2$), the age-dependent gap between the two thresholds is not monotonous. As the two externalities combine each other, it comes indeed that the optimal rise of selection into training programs is significantly delayed with respect to what firms do in equilibrium. The latter account for an increase in the gap between equilibrium and efficient thresholds for workers aged between 52 and 58. Then, as both private and social values of training collapse, this gap falls to zero in the very end of the working life. But when we consider $p_0 = p_2$, overall externalities are lowered since only the poaching externality holds, and it turns out that the gap between equilibrium and efficient ability thresholds is unambiguously decreasing with worker's age. To some extent, these results already suggest that job finding probabilities matter a lot in the age-design of training inefficiencies.

3 Life cycle dynamics of unemployment and vocational training

We now propose to revisit these formers results in the context of endogenous matching, that is with explicit job creation decisions. To that end, we first describe the equilibrium conditions, then show related optimal conditions with a discussion of some (in)efficiency issues. In a last step we calibrate the model and simulate the joint age-dynamics of training selection, job finding probabilities and employment, both at equilibrium and at the optimum.

3.1 Hiring decisions and endogenous matching probabilities

We consider a frictional labor market with directed search over the life cycle where firms choose how many vacancies to open. The type of vacancy is 3-dimensional, since it is defined by worker's age t, ability a and status of knowledge (up-to-date or obsolete), both assumed to be perfectly observable. Let $v_{j,t}(a)$, $u_{j,t}(a)$ be respectively the number of vacant jobs and the number of unemployed workers of type j, age t and ability a,[6] the number of hires per unit of time is assumed to be equal to $M(u_{j,t}(a), v_{j,t}(a)) = [u_{j,t}(a)]^{\eta}[v_{j,t}(a)]^{1-\eta}$ with $0 < \eta < 1$. The labor market tightness is the ratio of the number of vacancies to the number of unemployed, that is $\theta_{j,t}(a) \equiv \frac{v_{j,t}(a)}{u_{j,t}(a)}$, so that the rate at which vacancies are filled is given by $q(\theta_{j,t}(a)) \equiv \theta_{j,t}(a)^{-\eta}$, and the probability for an unemployed worker of type j, age t and ability a to be employed at the next period is $p(\theta_{j,t}(a)) \equiv \theta_{j,t}(a)^{1-\eta}$. We now restate the probabilities for an unemployed worker of age t to be employed at age $t+1$ as follows:

- $p(\theta_{0,t}(a)) = \theta_{0,t}(a)^{1-\eta}$ for individuals with obsolete skills, unable for training if they are hired at the next period ($a < \tilde{a}_{t+1}$)
- $p(\theta_{1,t}(a)) = \theta_{1,t}(a)^{1-\eta}$ for individuals with obsolete skills, able for training if they are hired at the next period ($a \geq \tilde{a}_{t+1}$)
- $p(\theta_{2,t}(a)) = \theta_{2,t}(a)^{1-\eta}$ for individuals with up-to-date skills.

Related labor market flows are presented in Appendix C.

Any firm is free to open a job vacancy directed toward a worker of type j, age t, and ability a, with a related recruitment cost that we denote c. The job vacancy at t is then filled at time $t+1$ with a probability $q(\theta_{j,t}(a))$. We let $V_{j,t}(a)$ be the inter-temporal value of a vacant position defined, $\forall\ t < T-1$, by: [7]

$$\begin{aligned}
V_{0,t}(a) &= -c + \beta\left[q(\theta_{0,t}(a))J_{0,t+1}(a) + [1 - q(\theta_{0,t}(a))]\max_{j,t}\{V_{j,t}(a)\}\right] \\
V_{1,t}(a) &= -c + \beta\left[q(\theta_{1,t}(a))[J_{1,t+1}(a) - \gamma_f] + [1 - q(\theta_{1,t}(a))]\max_{j,t}\{V_{j,t}(a)\}\right] \\
V_{2,t}(a) &= -c + \beta\left[q(\theta_{2,t}(a))J_{2,t+1}(a) + [1 - q(\theta_{2,t}(a))]\max_{j,t}\{V_{j,t}(a)\}\right]
\end{aligned}$$

[6] The assumption of directed search allows to focus on the role of training externalities. Indeed, it has already been pointed out that non-directed search over the life cycle can give rise to intergenerational externalties in the matching process which add another source of inefficiencies (see Chéron et al. [2011, 2013]).

[7] It is obvious that there is no opening of job vacancies at time $T-1$, because once the job is filled time T the worker retire.

where the firms' job value function are defined by (1)-(3).

New firms enter the labor market and open job vacancies until $V_{j,t}(a) = 0$, $\forall\ j, t, a$. Consequently, the free-entry condition implies that, $\forall\ t < T-1$:

$$\begin{aligned}
\frac{c}{q(\theta_{0,t}(a))} &= \beta J_{0,t+1}(a) \\
\frac{c}{q(\theta_{1,t}(a))} &= \beta\left[J_{1,t+1}(a) - \gamma_f\right] \\
\frac{c}{q(\theta_{2,t}(a))} &= \beta J_{2,t+1}(a)
\end{aligned}$$

whereas $\theta_{j,T-1}(a) = 0\ \ \forall a, j$.

Then, as the equilibrium training condition (equation (7)) is satisfied, on the one hand it is straightforward to see that at the equilibrium we do have $q(\theta_{0,t}(\tilde{a}_t)) = q(\theta_{1,t}(\tilde{a}_t))$, hence $p(\theta_{0,t}(\tilde{a}_t)) = p(\theta_{1,t}(\tilde{a}_t))$.

On the other hand, let substitute out p_0 by $p(\theta_{0,t-1}(a))$, and p_2 by $p(\theta_{2,t-1}(a))$ in equations (4)-(6) we get equilibrium wages. Let substitute again the latter into (1)-(3), this shows in turn that the equilibrium is characterized by $q(\theta_{0,t}(\tilde{a}_t)) < q(\theta_{2,t}(\tilde{a}_t))$, hence $p(\theta_{0,t}(\tilde{a}_t)) < p(\theta_{2,t}(\tilde{a}_t))$. Otherwise stated, endogenous hiring decisions are consistent with the "unemployment externality", that is unemployed workers with up-to-date knowledge get a higher probability to find a job than workers who faced skill obsolescence.

3.2 Revisiting the impact of training externalities

Obviously, due to training externalities, we do expect that equilibrium labor tightness/hiring decisions are not efficient. Indeed, optimal hiring policies are consistent with the fact that at each age the social planner chooses the labour market tightness $\theta^{\star}_{j,t}(a)$ that maximizes the social value of an worker. These values now include the loss due to recruitment costs per unemployed, $c\theta^{\star}_{j,t}(a)$, so that we now restate:

$$
\begin{aligned}
Y_t^{u0}(a) &= b - c\theta_{0,t}(a) + \beta\left[p(\theta_{0,t}(a))\hat{Y}_{t+1}(a) + [1-p(\theta_{0,t}(a))]Y_{t+1}^{u0}(a)\right] \\
Y_t^{u1}(a) &= b - c\theta_{1,t} + \beta\begin{cases} p(\theta_{1,t}(a))[\tilde{Y}_{t+1}(a)-\gamma_F] + [1-p(\theta_{1,t}(a))]Y_{t+1}^{u1}(a) & , \forall\ a \geq a_{t+1}^{\star} \\ p(\theta_{0,t}(a))\hat{Y}_{t+1}(a) + [1-p(\theta_{0,t}(a))]Y_{t+1}^{u0}(a) & , \forall\ a < a_{t+1}^{\star} \end{cases} \\
Y_t^{u2}(a) &= b - c\theta_{2,t}(a) + \beta\Bigg[p(\theta_{2,t}(a))\tilde{Y}_{t+1}(a) + [1-p(\theta_{2,t}(a))](1-\pi)Y_{t+1}^{u2}(a) \\
&\quad + [1-p(\theta_{2,t}(a))]\pi\begin{cases} Y_{t+1}^{u1}(a) & , \forall\ a \geq a_{t+1}^{\star} \\ Y_{t+1}^{u0}(a) & , \forall\ a < a_{t+1}^{\star} \end{cases}\Bigg]
\end{aligned}
$$

Optimal hiring decisions solve $\max_{\theta^{\star}_{j,t}(a)} Y_t^{u,j}(a) \quad \forall j, a, t$, and the optimal labour market tightness therefore satisfies:

$$
\begin{aligned}
\frac{c}{q(\theta^{\star}_{0,t}(a))} &= (1-\eta)\beta\left[\hat{Y}_{t+1}(a) - Y_{t+1}^{u0}(a)\right] \\
\frac{c}{q(\theta^{\star}_{1,t}(a))} &= (1-\eta)\beta\left[\left(\tilde{Y}_{t+1}(a) - \gamma_f\right) - Y_{t+1}^{u1}(a)\right], \forall\ a \geq \tilde{a}_{t+1} \\
\frac{c}{q(\theta^{\star}_{2,t}(a))} &= (1-\eta)\beta\left[\left(\tilde{Y}_{t+1}(a) - Y_{t+1}^{u2}(a)\right) + \pi\left(Y_{t+1}^{u2}(a) - \begin{cases} Y_{t+1}^{u1}(a) & , \forall\ a \geq \tilde{a}_{t+1} \\ Y_{t+1}^{u0}(a) & , \forall\ a \in [\tilde{a}_t; \tilde{a}_{t+1}[\end{cases}\right)\right]
\end{aligned}
$$

Then, we can provide some additional analytical insights by solving recursively these free entry conditions. First recall that it is obvious that, both at the optimum and the equilibrium, the labor tightness is zero at $t = T-1$, whatever worker's type and ability ($theta_{j,T-2}(a) \;\; \forall a, j$). Then, at $t = T-2$ we find that:

$$
\begin{aligned}
&\tfrac{c}{q(\theta_{0,T-2}(a))} = \beta(1-\alpha)(a-b) &;\quad &\tfrac{c}{q(\theta^{\star}_{0,T-2})} = \beta(1-\eta)(a-b) \\
&\tfrac{c}{q(\theta_{1,T-2}(a))} = \beta(1-\alpha)((1+\Delta)a-b-\gamma_f) &;\quad &\tfrac{c}{q(\theta^{\star}_{1,T-2}(a))} = \beta(1-\eta)((1+\Delta)a-b-\gamma_f) \\
&\tfrac{c}{q(\theta_{2,T-2}(a))} = \beta(1-\alpha)((1+\Delta)a-b) &;\quad &\tfrac{c}{q(\theta^{\star}_{2,T-2}(a))} = \beta(1-\eta)((1+\Delta)a-b)
\end{aligned}
$$

This shows that that for $t = T-2$, the Hosios condition $\eta = \alpha$ is sufficient for the equilibrium labor market tightness to be efficient whatever worker's type and ability. For $t = T-3$, we can show that this result holds also for type-0 and type-1 workers but no longer for type-2, *i.e.* training externalities distort job creation for type-2 workers. We have indeed:[8]

[8] To state this we here consider $\tilde{a}_{T-3} \leq \tilde{a}_{T-2}$.

$$
\begin{aligned}
\frac{c}{q(\theta_{0,T-3}(a))} &= \beta(1-\alpha)(a-b)[1+\beta(1-\delta)-\alpha\beta p(\theta_{0,T-2}(a))] \\
\frac{c}{q(\theta^\star_{0,T-3}(a))} &= \beta(1-\eta)(a-b)[1+\beta(1-\delta)-\eta\beta p(\theta^\star_{0,T-2}(a))] \\
\frac{c}{q(\theta_{1,T-3}(a))} &= \beta(1-\alpha)\left\{(a-b)[1+\beta(1-\delta)-\alpha\beta p(\theta_{0,T-2}(a))]\right. \\
&\quad \left.+\Delta a-\gamma_f+\beta(1-\delta)\Delta a\right\} \\
\frac{c}{q(\theta^\star_{1,T-3}(a))} &= \beta(1-\alpha)\left\{(a-b)[1+\beta(1-\delta)-\eta\beta p(\theta^\star_{0,T-2}(a))]\right. \\
&\quad \left.+\Delta a-\gamma_f+\beta(1-\delta)\Delta a\right\} \\
\frac{c}{q(\theta_{2,T-3}(a))} &= \beta(1-\alpha)[(1+\Delta)a-b][1+\beta(1-\delta)-\alpha\beta p(\theta_{2,T-2}(a))] \\
\frac{c}{q(\theta^\star_{2,T-3}(a))} &= \beta(1-\eta)\left\{[(1+\Delta)a-b][1+\beta(1-\delta)-\eta\beta p(\theta^\star_{2,T-2}(a))]\right. \\
&\quad \left.+\pi\beta\eta\left[((1+\Delta)a-b)[p(\theta^\star_{2,T-2}(a))-p(\theta^\star_{0,T-2}(a))]+\Delta a p(\theta^\star_{0,T-2}(a))\right]\right\}
\end{aligned}
$$

The last term of the last equation shows the interaction between training externality and job creation: efficient type-2 tightness is higher than at equilibrium even though the Hosios condition holds ($\alpha=\eta$). The planner actually internalizes that a higher labor market tightness generates an intertemporal externality due to training costs: for type-2 unemployed workers, a higher tightness reduces the risk of skill depreciation (with probability π). The point is indeed that when skill depreciation occurs, this accounts for longer unemployment spell (which depends on the gap $p(\theta^\star_{2,T-2}(a))-p(\theta^\star_{0,T-2}(a))$) and lower productivity. There is a social loss given by $(1+\Delta)a-b$ that is not valuated by firms once worker faces skill obsolescence and stays longer unemployed. Overall, this leads job finding rates of type-2 workers to be too low at equilibrium.

Lastly, we may also wonder to what extent the inefficiency of training selection is affected by endogenous matching. The main point is that the social values related to unemployment now include the recruitment costs. Again this point can be highlighted by solving recursively the model. We first notice that solutions at equilibrium and optimum for $t=T-1$ and $t=T-2$ are the same as at partial equilibrium (equations (8) and (9)). But, solutions for $t=T-3$ show new interesting results:

$$\Delta \tilde{a}_{T-3} = \frac{\gamma_f - (\tilde{a}_{T-3} - b)\alpha\beta^2\delta[p(\theta_{2,T-2}(\tilde{a}_{T-2})) - p(\theta_{0,T-2}(\tilde{a}_{T-2}))]}{\sum_{i=0}^{2}[\beta(1-\delta)]^i + \alpha\beta^2\delta p(\theta_{2,T-2}(\tilde{a}_{T-2}))} \quad (12)$$

$$\Delta a^{\star}_{T-3} = \frac{\gamma_f - (a^{\star}_{T-3} - b)\eta\beta^2\delta[p(\theta^{\star}_{2,T-2}(a^{\star}_{T-2})) - p(\theta^{\star}_{0,T-2}(a^{\star}_{T-2}))]}{\sum_{i=0}^{2}[\beta(1-\delta)]^i + \eta\beta^2\delta p(\theta^{\star}_{2,T-2}(a^{\star}_{T-2}))} \quad (13)$$

We now have that the Hosios condition ($\alpha = \eta$) is sufficient for the training decision at $t = T-3$ to be efficient. In the context of endogenous matching, the planner indeed internalizes that the poaching externality induces recruitment costs (which depend on $c\theta_{j,t}$).[9] With respect to the efficient partial equilibrium case, we have therefore a higher optimal training threshold (by comparison with the partial equilibrium case), and it comes that the equilibrium training policy is now efficient for $\alpha = \eta$.[10] Accordingly, under the Hosios condition, the point is therefore that training externality distorts the allocation only through its impact on hiring.

Analyzing what is going on for $t = T-4$ actually comforts this result. It comes indeed that if the Hosios condition is satisfied, $p(\theta^{\star}_{2,T-3}(a^{\star}_{T-3})) = p(\theta_{2,T-3}(\tilde{a}_{T-3}))$ is a sufficient condition for $\tilde{a}_{T-4} = a^{\star}_{T-4}$ (see Appendix B). Again, in words, once job creation is efficient, we do necessarily have that the training selection is also optimal (for $\alpha = \eta$).

All in all, an interpretation of these results is that training externalities now rather require higher job finding rates (for type-2 workers) to reduce the probability of experiencing skill depreciation (hence future training costs), than accepting to train lower ability unemployed workers with obsolete knowledge. This suggests that enhancing hirings is crucial to internalize training externalities.

3.3 Model simulations

Our last objective is to provide a quantitative evaluation of the potential impact of training externalities over the life cycle. As emphasized earlier, we need to focus on the life cycle dynamics of job finding rates, hence employment. The model is simulated at a quarterly frequency. A first set parameters is calibrated in a fairly standard way $\Phi_1 = \{\beta, b, \underline{a}, \alpha, \eta\}$ (see

[9]Let remark that we do not allow for job-to-job transition, so training externalities relate to employment-unemployment-employment transitions.

[10]This can be seen by comparing equations (10) and (13) since $\eta < 1$.

table 1). In particular, the turbulence parameter π is taken from Ljungqvist and Sargent [2004], and the corresponding value suggests that in turbulent times the expected unemployment duration before undergoing skill loss is two quarters.

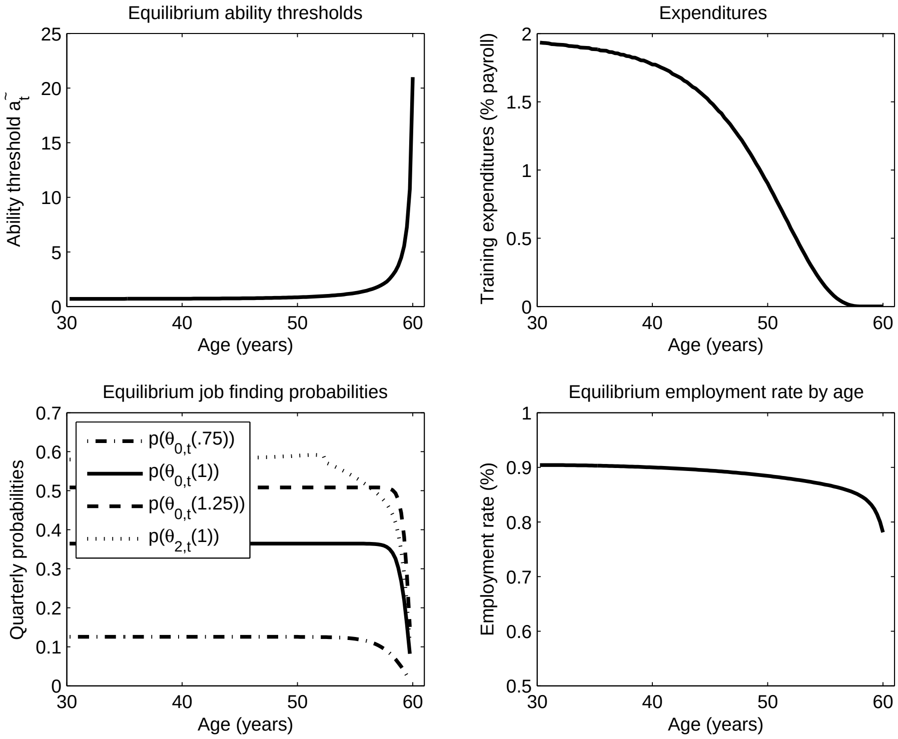
Table 1: Model parameters

Parameter	Description	Value
β	Discount factor	0.99
b	Home production	0.7
$\underline{a}$	lowest ability level	0.7
α	Bargaining power of workers	0.5
η	Matching elasticity	0.5
π	Turbulence parameter	0.5
T	Retirement age	160
δ	Job separation probability	0.0356
Δ	Additional output from up-to-date knowledge	0.1
γ_f	Training cost	2.1
c	Recruitment cost	1.8

Then, a second subset of parameters is calibrated on french data, $\Phi_2 = \{T, \delta, c, \gamma_f, \Delta\}$ (see table 1). This leads to consider $t = [1, 160]$ by referring to workers from 20 to 60 years old (corresponding to the average retirement age over this period). Actually, since in our benchmark model, life cycle only relies to distance-to-retirement, our quantitative investigation will focus on 30-60 years old workers, hence we do not aim at discussing the labor market entry of the youth. We consider a job destruction probability $\delta = 3.56\%$ per quarter which is in accordance with Hairault et al. [2015]. We use statistics for the life dynamics of vocational training expenditures that are taken from Chéron et al. [2015] which refer to the french Labor Force Survey. Total firms' training sponsored expenditures amount approximately to 1.3% of total wage costs for workers over 25 years old. Setting $\gamma_f = 2.1$ allows to account for this statistics, which means that our training cost is equal to the total output of the lowest ability worker over three quarters. Measuring the additional output related to training is a somewhat complex and disputed issue. Chéron et al. [2010] estimated on French Data that training participation increases wages by 6.7%. In accordance with bargaining of wages, implying that workers only get part of the productivity gain, it seems that setting $\Delta = 0.1$ is reasonable. The recruitment cost c is chosen to be consistent with an average unemployment rate for workers aged 25 to 49 of 10%.[11]

[11] As we do not examine participation, in our model employment rate is 1-unemployment rate. Moreover, it is well known that in France the unemployment rate for workers over 50

Figure 2: Training and employment dynamics



In our model, although firms choose whether to train workers (with productivity gain Δa) or not, it is obvious that productivity dispersion highly depends on the distribution of abilities. To be consistent with a long right tailed distribution of productivities, we consider the following Pareto distribution of abilities:

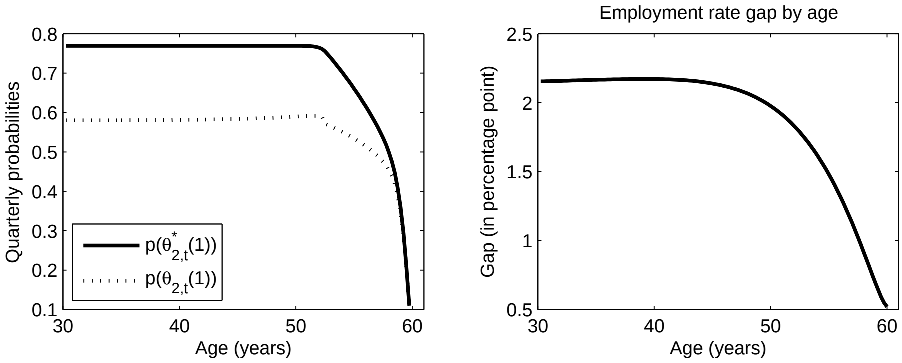
$$F(a) = \frac{1 - (\underline{a}/a)^k}{1 - (\underline{a}/\bar{a})^k}$$

where we set $\bar{a} = \gamma_f/\Delta$, so that an interior solution exists over all the life cycle.[12] Therefore, we choose parameter $k = 3$ such that when computing quartiles related to this distribution, we get Q3/Q1=1.5, which is in accordance with the range of estimates by Bontemps et al. [2000].

years old is not very informative, due to institutions that lead workers out of employment to be referred to as inactive.

[12]We know indeed that $a^{\star}_{T-1} = \tilde{a}_{T-1} = \gamma_f/\Delta$.

Figure 3: Training externalities and inefficient job creation



We can then compute equilibrium life cycle dynamics of the ability threshold, job finding probabilities and employment rates. Figure 2 first shows a strong heterogeneity of job finding rates not only according to ability but also to status wrt training. For instance, at 40 the quarterly probability to find a job for a worker with up-to-date knowledge and ability one ($a = 1$) is close to 60%, hence more than 5 times greater than that of a worker with obsolete knowledge, unable for training, and an ability 25% lower ($a = 0.75$). Simulations also show that due to horizon effect (shorter distance-to-retirement), we do observe not only a sharp rise of selection into training programs, but also a fall of job finding probabilities over 50 years old. This accounts for a decrease of the overall employment rate of about 10 points, that is about a half of the fall we observe in the french data over this range of ages.[13]

Here, our main objective is rather to assess what could expected (in terms of employment) from dealing with training externalities, that interact with job creation. As discussed earlier, in our context of endogenous matching the key variable is the job finding probabililty of type-2 workers (workers with up-to-date knowledge). With respect to the optimal decision, job creation by firms is indeed too low, which means that the risk of skill obsolescence turns out to be to high. Figure 3 plots the gap between equilibrium and the optimum: for type-2 workers with ability $a = 1$, it comes that at 40 the optimal job finding probability is about 80% whereas it is only 60% at the equilibrium. As retirement gets closer, this gap converges to zero. Then, from the aggregation over all abilities, we can compute the employment incidence

[13]But it is obvious that explaining the labor market of older workers is here beyond the scope of that paper, otherwise we would have to deal with the life cycle dynamics of endogenous job destruction (see for instance Chéron et al. [2013]).

of training externalities. We find it significant since it is greater than 2 points under 50 and then falls to 0.5 at the very end of the life cycle.

4 Conclusion

In this paper, we developed a life cycle model to examine the (in)efficiency of vocational training investments in a frictional labor market context. Our main goal was to examine age-dependent impact of training externalities, through distance-to-retirement effects on training but also job creation decisions. Ultimately, we argued that internalizing training externalities, *ie.* both future costs of skill obsolescence and productivity gains related to up-to-date knowledge, require to boost hirings of those workers that are not yet obsolete. At the very end of the life cycle, this issue turns out to be less crucial because optimal job finding probabilities decrease due to shorter horizon, which accounts for a reduction of training externalities. But as the employment is predetermined by past decisions, our quantitative investigation suggests that implementing the optimal allocation would account for significant employment increase, even for the older workers.

The last issue relates to the way the optimal policy should be implemented. Obviously, our model is quite too stylized to propose a pragmatic tool. Indeed, according to our model, the optimal policy would consist in implementing an hiring subsidy targeted toward workers with up-to-date knowledge, and a function of both ability and age. This would require information for the public policy maker that is not straightforward. However, our main message is the following. Training externalities are important, and would require some incentives to reduce selection into training programs and raise hirings, essentially for short term unemployed workers to prevent them from skill obsolescence and related social costs.

References

Daron Acemoglu. Training and innovation in an imperfect labour market. *Review of Economic Studies*, 64(3):445–64, 1997.

Daron Acemoglu and Jorn-Steffen Pischke. The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572, 1999.

Gary S. Becker. Human capital: A theoretical and empirical analysis, with special reference to education. 1964.

Pascal Belan and Arnaud Chéron. Turbulence, training and unemployment. *Labour Economics*, 27(C):16–29, 2014.

Christian Bontemps, Jean-Marc Robin, and Gerard J van den Berg. Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation. *International Economic Review*, 41(2):305–58, May 2000.

Arnaud Chéron, Bénédicte Rouland, and François-Charles Wolff. Returns to firm-provided training in france: Evidence on mobility and wages. *TEPP Working Paper*, (10), 2010.

Arnaud Chéron, Jean-Olivier Hairault, and François Langot. Age-dependent employment protection. *The Economic Journal*, 121(557):1477–1504, 2011.

Arnaud Chéron, Jean-Olivier Hairault, and François Langot. Life-cycle equilibrium unemployment. *Journal of Labor Economics*, 31(4):843 – 882, 2013.

Arnaud Chéron, Pierre Courtioux, and Vincent Lignon. Maintenir la formation continue chez les seniors : Pourquoi, comment, combien ? *EDHEC Position Paper*, 121(2):209–231, 2015.

Jean-Olivier Hairault, Thomas Le Barbanchon, and Thepthida Sopraseuth. The cyclicality of the separation and job finding rates in france. *European Economic Review*, 76:60–84, 2015.

Lars Ljungqvist and Thomas J. Sargent. The european unemployment dilemma. *Journal of Political Economy*, 106(3):514–550, 1998.

Lars Ljungqvist and Thomas J. Sargent. European unemployment and turbulence revisited in a matching model. *Journal of the European Economic Association*, 2(2-3):456–468, 2004.

Guido Menzio, Irina Telyukova, and Ludo Visschers. Directed search over the life cycle. *Review of Economic Dynamics*, 2015.

Pierre-Jean Messe and Bénédicte Rouland. Stricter employment protection and firms' incentives to sponsor training: The case of french older workers. *Labour Economics*, 31(C):14–26, 2014.

A The impact of hold-up

In the context of hold-up, we consider that type-1 workers are able to renegotiate type-2 wages once training occurred (that is before production takes place). Accordingly, $w^h_{1,t} = w_{2,t}$, so we have $J^h_{1,t} = J_{2,t}$. Yet we consider $p_1 < p_2$ (which is an endogenous matching result). The unemployed value of type-1 workers then turns out to be defined by:

$$U_{1,t}(a) = b + \beta \begin{cases} p_1 E_{2,t+1}(a) + (1-p_1)\, U_{1,t+1}(a) & , \forall\ a \geq \tilde{a}_{t+1} \\ p_0 E_{0,t+1}(a) + (1-p_0)\, U_{0,t+1}(a) & , \forall\ a \in [\tilde{a}_t; \tilde{a}_{t+1}[\end{cases}$$

This shows that type-1 unemployed now expects to be type-2 employed, once they find a job.

The related ability threshold, that we now denote a^h_t, is therefore derived from the following condition:

$$J^h_{1,t}(a^h_t) - \gamma_f = J_0(a^h_t) \iff J_{2,t}(a^h_t) - \gamma_f = J_0(a^h_t)$$

This implies:

$$\Delta a^h_t = \frac{\gamma_f + (1-\alpha)(U^h_{2,t}(a^h_t) - U^h_{0,t}(a^h_t)) - \sum_{i=0}^{T-3-t} \beta\delta\, [\beta(1-\delta)]^i \left[U^h_{2,t+1+i}(a^h_t) - U_{0,t+1+i}(a^h_t)\right]}{(1-\alpha)\sum_{i=0}^{T-1-t} [\beta(1-\delta)]^i},$$

which in particular leads to:

$$\Delta a^h_{T-1} = \frac{\gamma_f}{1-\alpha} > \Delta\tilde{a}_{T-1}$$

$$\Delta a^h_{T-2} = \frac{\gamma_f}{1-\alpha}\left(\frac{1+\alpha\beta p_2}{\sum_{i=0}^{1}[\beta(1-\delta)]^i}\right) + \frac{\alpha\beta(p_2-p_0)(a^h_{T-1}-b)}{\sum_{i=0}^{1}[\beta(1-\delta)]^i} > \Delta\tilde{a}_{T-2}$$

It is therefore obvious that hold-up introduces additional inefficiencies. In particular, it comes that at $t = T-1$ we do have $a^h_{T-1} > \tilde{a}_{T-1} = a^\star_{T-1}$. Otherwise stated, whereas unemployment and poaching externalities vanish at the end of the life cycle (due to horizon effect), hold-up still increase wages hence depress training in comparison with what would be required at the optimum.

B Ability thresholds with endogenous matching for $t = T - 4$

We can find the solution at $t = T - 4$, by using backward induction. We get at the equilibrium:

$$
\begin{aligned}
&\Delta a_{\tilde{T}-4} \left[1 + \beta(1-\delta) + [\beta(1-\delta)]^2 + [\beta(1-\delta)]^3 + \beta(1-\delta)\beta^2\delta\alpha p_{2,T-3} \left[1 + \beta(1-\delta-\alpha p_{2,T-2})\right] + \beta\delta\beta^2 \left[1 - \pi(1-p_{2,T-3})\right] p_{2,T-2}\right] \\
&= \gamma_f - (a_{\tilde{T}-4} - b)\left[\beta(1-\delta)\alpha\beta^2\delta(p_{2,T-2} - p_{0,T-2}) + \beta\delta\left[\alpha\beta p_{2,T-3} \left[1 + \beta(1-\delta-\alpha p_{2,T-2})\right] - \alpha\beta p_{0,T-3} \left[1 + \beta(1-\delta-\alpha p_{0,T-2})\right]\right.\right. \\
&\left.\left. + \alpha\beta^2 \left[1 - \pi(1-p_{2,T-3}\right] \left[p_{2,T-2} - p_{0,T-2}\right]\right]\right]
\end{aligned}
$$

where $p_{j,t}$ stands for $p(\theta_{j,t}(\tilde{a}_t))$. And at the optimum we also have:

$$
\begin{aligned}
&\Delta a^\star_{T-4} \left[1 + \beta(1-\delta) + [\beta(1-\delta)]^2 + [\beta(1-\delta)]^3 + \beta(1-\delta)\beta^2\delta\eta p^\star_{2,T-3} \left[1 + \beta(1-\delta-\eta p^\star_{2,T-2})\right] + \beta\delta\beta^2 \left[1 - \pi(1-p^\star_{2,T-3})\right] p^\star_{2,T-2}\right] \\
&= \gamma_f - (a^\star_{T-4} - b)\left[\beta(1-\delta)\eta\beta^2\delta(p^\star_{2,T-2} - p_{0,T-2}) + \beta\delta\left[\eta\beta p^\star_{2,T-3} \left[1 + \beta(1-\delta-\eta p^\star_{2,T-2})\right] - \alpha\beta p_{0,T-3} \left[1 + \beta(1-\delta-\eta p_{0,T-2})\right]\right.\right. \\
&\left.\left. + \alpha\beta^2 \left[1 - \pi(1-p^\star_{2,T-3}\right] \left[p^\star_{2,T-2} - p_{0,T-2}\right]\right]\right]
\end{aligned}
$$

where $p^\star_{j,t}$ stands for $p(\theta^\star_{j,t}(a^\star_t))$.

From this, we find that with $\eta = \alpha$, if $p^\star_{2,t} = p_{2,t} \quad \forall t = T-3, T-2$, then we have $\tilde{a}_{T-4} = a^\star_{T-4}$.

C Workers flows

Let $n_{j,t}(a)$ and $u_{j,t}(a)$ be respectively the level of employment and the level of unemployment for individuals of type j, age t and ability a. The age-dynamic of workers flows is then given by :

For $t = 0$:

Initially, all individuals are endowed with up-to-date skills and enter on the labor market as type-2 unemployed, so that :

- $\forall\ a$:

  $u_{2,t}(a) = f(a)$

For $t = 1$:

The firms now determine at each age an ability threshold $\tilde{a}_t$ above which train a hired worker becomes profitable. We note that some unemployed of type 2 may have undergone a skill loss with a probabilty π : They become unemployed of type 0 if their ability is below the threshold $\tilde{a}_1$, or unemployed of type 1 if their ability is superior or equal to $\tilde{a}_1$.

- $\forall\ a < \tilde{a}_t$,

  $u_{0,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi$

  $u_{2,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi)$
  $n_{2,t}(a) = u_{2,t-1}(a)p(\theta_{2,t-1}(a))$

- $\forall \geq \tilde{a}_t$,

  $u_{1,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi$

  $u_{2,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi)$
  $n_{2,t}(a) = u_{2,t-1}(a)p(\theta_{2,t-1}(a))$

For $t = 2$:

From this period, we note that some unemployed workers who were able for training in the previous period now have a level of ability too low to be trained : They become unemployed of type 0.

- $\forall\ a < \tilde{a}_{t-1}$:

  $u_{0,t}(a) = u_{0,t-1}(a)[1 - p(\theta_{0,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi$
  $n_{0,t}(a) = u_{0,t-1}(a)p(\theta_{0,t-1}(a))$

  $u_{2,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta n_{2,t-1}(a)$
  $n_{2,t}(a) = (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a))$

- $\forall\ a \in [\tilde{a}_{t-1}; \tilde{a}_t[$:

  $$\begin{aligned} u_{0,t}(a) &= u_{1,t-1}(a)[1 - p(\theta_{0,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi \\ n_{0,t}(a) &= u_{1,t-1}(a)p(\theta_{0,t-1}(a)) \end{aligned}$$

  $$\begin{aligned} u_{2,t}(a) &= u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta n_{2,t-1}(a) \\ n_{2,t}(a) &= (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a)) \end{aligned}$$

- $\forall\ a \geq \tilde{a}_t$:

  $$\begin{aligned} u_{1,t}(a) &= u_{1,t-1}(a)[1 - p(\theta_{1,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi \\ n_{1,t}(a) &= u_{1,t-1}(a)p(\theta_{1,t-1}(a)) \end{aligned}$$

  $$\begin{aligned} u_{2,t}(a) &= u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta n_{2,t-1}(a) \\ n_{2,t}(a) &= (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a)) \end{aligned}$$

$\forall\ t \in [3; T - 1]$:

From $t = 3$, we get a dynamic of workers flows which breeds until retirement age.

- $\forall\ a < \tilde{a}_{t-1}$,

  $$\begin{aligned} u_{0,t}(a) &= u_{0,t-1}(a)[1 - p(\theta_{0,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi + \delta n_{0,t-1}(a) \\ n_{0,t}(a) &= (1 - \delta)n_{0,t-1}(a) + u_{0,t-1}(a)p(\theta_{0,t-1}(a)) \end{aligned}$$

  $$n_{1,t}(a) = (1 - \delta)n_{1,t-1}(a)$$

  $$\begin{aligned} u_{2,t}(a) &= u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta[n_{1,t-1}(a) + n_{2,t-1}(a)] \\ n_{2,t}(a) &= (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a)) \end{aligned}$$

- $\forall\ a \in [\tilde{a}_{t-1}; \tilde{a}_t[$,

$$u_{0,t}(a) = u_{1,t-1}(a)[1 - p(\theta_{0,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi$$
$$n_{0,t}(a) = u_{1,t-1}(a)p(\theta_{0,t-1}(a))$$

$$n_{1,t}(a) = (1 - \delta)n_{1,t-1}(a)$$

$$u_{2,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta[n_{1,t-1}(a) + n_{2,t-1}(a)]$$
$$n_{2,t}(a) = (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a))$$

- $\forall\ a \geq \tilde{a}_t$,

$$u_{1,t}(a) = u_{1,t-1}(a)[1 - p(\theta_{1,t-1}(a))] + u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))]\pi$$
$$n_{1,t}(a) = (1 - \delta)n_{1,t-1}(a) + u_{1,t-1}(a)p(\theta_{1,t-1}(a))$$

$$u_{2,t}(a) = u_{2,t-1}(a)[1 - p(\theta_{2,t-1}(a))](1 - \pi) + \delta[n_{1,t-1}(a) + n_{2,t-1}(a)]$$
$$n_{2,t}(a) = (1 - \delta)n_{2,t-1}(a) + u_{2,t-1}(a)p(\theta_{2,t-1}(a))$$

D Wage bargaining

Let α be the bargaining power of workers, considered as constant across ages. Wages are renegociated at each period according to age and are solutions of the following Nash-sharing rules :

$$\begin{aligned}
(1 - \alpha)[E_{0,t}(a) - U_{0,t}(a)] &= \alpha J_{0,t}(a) \\
(1 - \alpha)[E_{1,t}(a) - U_{1,t}(a)] &= \alpha[J_{1,t}(a) - \gamma_f] \\
(1 - \alpha)[E_{2,t}(a) - U_{2,t}(a)] &= \alpha J_{2,t}(a)
\end{aligned}$$

D.1 Type-0 worker, unable for training ($a < \tilde{a}_t$):

Workers :

$$[E_{0,t}(a) - U_{0,t}(a)] = w_{0,t}(a) - b + \beta\left[1 - \delta - p(\theta_{0,t}(a))\right][E_{0,t+1}(a) - U_{0,t+1}(a)]$$

Firms :

$$J_{0,t}(a) = a - w_{0,t}(a) + \beta(1 - \delta)J_{0,t+1}(a)$$

According to the sharing rule :

$$(1-\alpha)\Big[w_{0,t}(a) - b + \beta\left[1-\delta - p(\theta_{0,t}(a))\right]\left[E_{0,t+1}(a) - U_{0,t+1}(a)\right]\Big]$$
$$= \alpha\Big[a - w_{0,t}(a) + \beta(1-\delta)J_{0,t+1}(a)\Big]$$

This implies the following wage for a type-0 worker :

$$w_{0,t}(a) = \alpha a + (1-\alpha)b + \alpha\beta p(\theta_{0,t}(a))J_{0,t+1}(a)$$

The expected value of a filled job by a type-0 worker is :

$$J_{0,t}(a) = (1-\alpha)\left[a - b\right] + \beta(1-\delta)J_{0,t+1}(a) - \alpha\beta p(\theta_{0,t}(a))J_{0,t+1}(a)$$

D.2 Type-1 worker, able for training for the last time ($a \in [\tilde{a}_t; \tilde{a}_{t+1}[$):

Workers :

$$\begin{aligned}\left[E_{1,t}(a) - U_{1,t}(a)\right] = {} & w_{1,t}(a) - b + \beta(1-\delta)\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right] \\ & - \beta p(\theta_{0,t}(a))\left[E_{0,t+1}(a) - U_{0,t+1}(a)\right] + \beta\delta\left[U_{2,t+1}(a) - U_{0,t+1}(a)\right]\end{aligned}$$

Firms :

$$J_{1,t}(a) = (1+\Delta)a - w_{1,t}(a) + \beta(1-\delta)J_{1,t+1}(a)$$

According to the sharing rule :

$$(1-\alpha)\Big[w_{1,t}(a) - b + \beta(1-\delta)\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right]$$
$$-\beta p(\theta_{0,t}(a))\left[E_{0,t+1}(a) - U_{0,t+1}(a)\right] + \beta\delta\left[U_{2,t+1}(a) - U_{0,t+1}(a)\right]\Big]$$
$$= \alpha\Big[(1+\Delta)a - w_{1,t}(a) + \beta(1-\delta)J_{1,t+1}(a) - \gamma_f\Big]$$

This implies the following wage for a type-1 worker :

$$\begin{aligned} w_{1,t}(a) &= \alpha(1+\Delta)a + (1-\alpha)b - \alpha\gamma_f\left[1-\beta(1-\delta)\right] \\ &\quad +\alpha\beta p(\theta_{0,t}(a))J_{0,t+1}(a) - \beta\delta(1-\alpha)\left[U_{2,t+1}(a) - U_{0,t+1}(a)\right] \end{aligned}$$

The expected value of a filled job by a type-1 worker is :

$$\begin{aligned} J_{1,t}(a) &= (1-\alpha)\left[(1+\Delta)a - b\right] + \alpha\gamma_f\left[1-\beta(1-\delta)\right] + \beta(1-\delta)J_{1,t+1}(a) \\ &\quad -\alpha\beta p(\theta_{0,t}(a))J_{0,t+1}(a) + \beta\delta(1-\alpha)\left[U_{2,t+1}(a) - U_{0,t+1}(a)\right] \end{aligned}$$

D.3 Type-1 worker, who will still be able for training at the next period ($a \geq \tilde{a}_{t+1}$):

Workers :

$$\begin{aligned} \left[E_{1,t}(a) - U_{1,t}(a)\right] = {} & w_{1,t}(a) - b + \beta(1-\delta)\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right] \\ & - \beta p(\theta_{1,t}(a))\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right] + \beta\delta\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right] \end{aligned}$$

Firms :

$$J_{1,t}(a) = (1+\Delta)a - w_{1,t}(a) + \beta(1-\delta)J_{1,t+1}(a)$$

According to the sharing rule :

$$\begin{aligned} &(1-\alpha)\Big[w_{1,t}(a) - b + \beta(1-\delta)\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right] \\ &-\beta p(\theta_{1,t}(a))\left[E_{1,t+1}(a) - U_{1,t+1}(a)\right] + \beta\delta\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right]\Big] \\ &= \alpha\Big[(1+\Delta)a - w_{1,t}(a) + \beta(1-\delta)J_{1,t+1}(a) - \gamma_f\Big] \end{aligned}$$

This implies the following wage for a type-1 worker :

$$\begin{aligned} w_{1,t}(a) &= \alpha(1+\Delta)a + (1-\alpha)b - \alpha\gamma_f\left[1-\beta(1-\delta)\right] \\ &\quad +\alpha\beta p(\theta_{1,t}(a))[J_{1,t+1}(a) - \gamma_f] - \beta\delta(1-\alpha)\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right] \end{aligned}$$

The expected value of a filled job by a type-1 worker is :

$$\begin{aligned} J_{1,t}(a) &= (1-\alpha)\left[(1+\Delta)a - b\right] + \alpha\gamma_f\left[1-\beta(1-\delta)\right] + \beta(1-\delta)J_{1,t+1}(a) \\ &\quad -\alpha\beta p(\theta_{1,t}(a))[J_{1,t+1}(a)-\gamma_f] + \beta\delta(1-\alpha)\left[U_{2,t+1}(a)-U_{1,t+1}(a)\right] \end{aligned}$$

D.4 Type-2 worker, with up-to-date knowledge, who will no longer be able for training if he undergoes a depreciation of human capital ($a < \tilde{a}_{t+1}$):

Workers :

$$\begin{aligned} [E_{2,t}(a) - U_{2,t}(a)] = w_{2,t}(a) - b + \beta\left[1-\delta-p(\theta_{2,t}(a))\right]\left[E_{2,t+1}(a)-U_{2,t+1}(a)\right] \\ + \beta\pi\left[1-p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a)-U_{0,t+1}(a)\right] \end{aligned}$$

Firms :

$$J_{2,t}(a) = (1+\Delta)a - w_{2,t}(a) + \beta(1-\delta)J_{2,t+1}(a)$$

According to the sharing rule :

$$\begin{aligned} &(1-\alpha)\Big[w_{2,t}(a) - b + \beta\left[1-\delta-p(\theta_{2,t}(a))\right]\left[E_{2,t+1}(a)-U_{2,t+1}(a)\right] \\ &+\beta\pi\left[1-p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a)-U_{0,t+1}(a)\right]\Big] = \alpha\Big[(1+\Delta)a - w_{2,t}(a) + \beta(1-\delta)J_{2,t+1}(a)\Big] \end{aligned}$$

This implies the following wage for a type-2 worker :

$$\begin{aligned} w_{2,t}(a) &= \alpha(1+\Delta)a + (1-\alpha)b + \alpha\beta p(\theta_{2,t}(a))J_{2,t+1}(a) \\ &\quad -\beta\pi(1-\alpha)\left[1-p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a)-U_{0,t+1}(a)\right] \end{aligned}$$

The expected value of a filled job by a type-2 worker is :

$$\begin{aligned} J_{2,t}(a) &= (1-\alpha)\left[(1+\Delta)a - b\right] + \beta(1-\delta)J_{2,t+1}(a) - \alpha\beta p(\theta_{2,t}(a))J_{2,t+1}(a) \\ &\quad +\beta\pi(1-\alpha)\left[1-p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a)-U_{0,t+1}(a)\right] \end{aligned}$$

D.5 Type-2 worker, with up-to-date knowledge, who will still be able for training at the next period if he undergoes a depreciation of human capital ($a \geq \tilde{a}_{t+1}$):

Workers :

$$\begin{aligned}[E_{2,t}(a) - U_{2,t}(a)] = {} & w_{2,t}(a) - b + \beta\left[1 - \delta - p(\theta_{2,t}(a))\right]\left[E_{2,t+1}(a) - U_{2,t+1}(a)\right] \\ & + \beta\pi\left[1 - p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right]\end{aligned}$$

Firms :

$$J_{2,t}(a) = (1+\Delta)a - w_{2,t}(a) + \beta(1-\delta)J_{2,t+1}(a)$$

According to the sharing rule :

$$\begin{aligned}&(1-\alpha)\Big[w_{2,t}(a) - b + \beta\left[1 - \delta - p(\theta_{2,t}(a))\right]\left[E_{2,t+1}(a) - U_{2,t+1}(a)\right] \\ &+\beta\pi\left[1 - p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right]\Big] = \alpha\Big[(1+\Delta)a - w_{2,t}(a) + \beta(1-\delta)J_{2,t+1}(a)\Big]\end{aligned}$$

This implies the following wage for a type-2 worker :

$$\begin{aligned} w_{2,t}(a) &= \alpha(1+\Delta)a + (1-\alpha)b + \alpha\beta p(\theta_{2,t}(a))J_{2,t+1}(a) \\ & \quad -\beta\pi(1-\alpha)\left[1 - p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right]\end{aligned}$$

The expected value of a filled job by a type-2 worker is :

$$\begin{aligned} J_{2,t}(a) &= (1-\alpha)\left[(1+\Delta)a - b\right] + \beta(1-\delta)J_{2,t+1}(a) - \alpha\beta p(\theta_{2,t}(a))J_{2,t+1}(a) \\ & \quad +\beta\pi(1-\alpha)\left[1 - p(\theta_{2,t}(a))\right]\left[U_{2,t+1}(a) - U_{1,t+1}(a)\right]\end{aligned}$$