UNIVERSITÉ DU MAINE

Faculté de Droit, Sciences Economiques et Gestion

Ecole Doctorale Droit, Economie, Gestion, Environnement, Sociétés et Territoires

Année 2011 Numéro attribué par la bibliothèque

| | | | | | | | | | | | |

Thèse de Doctorat de Sciences Economiques

Présentée et soutenue publiquement par

Bénédicte Rouland

le 6 Décembre 2011

Formation continue des salariés, chômage et efficience : Analyses empiriques et théoriques

---

Directeur de Thèse

Arnaud Chéron Professeur à l'Université du Maine, GAINS-TEPP

Jury

Rapporteurs

Elena Stancanelli Chargée de recherche, CNRS

Olivier Charlot Professeur à l'Université de Cergy-Pontoise, THEMA

Examinateurs

François Langot Professeur à l'Université du Maine, GAINS-TEPP

Michael Rosholm Professeur à l'Université d'Aarhus (Danemark)

Contents

Acknowledgements

This thesis was written in the period from September 2006 to December 2011 while I was a Ph.D. student at GAINS-TEPP, University of Le Mans. I am grateful to the lab for having provided me with a pleasant and stimulating working environment and with funding to attend conferences. I also benefited from the administrative help of many people.

I would like to thank a number of people who have helped me in the writing of this thesis. First and foremost, I would like to thank Arnaud Chéron for having directed this dissertation with patience, attentiveness and for his teaching ability. Arnaud not only helped me in solving problems with this thesis, but also taught me a lot about how economic research should be done. Thanks to him, I was able to develop the necessary scientific skills to undertake and complete the work. François Langot also deserves thanks for the constructiveness of his comments and suggestions, especially on Chapter 2. I have benefited tremendously from working along side Arnaud Chéron, Pierre-Jean Messe and François-Charles Wolff on parts of the thesis. I also had many fruitful discussions with my office-mates, Ph.D. students and Professors. I am very grateful to Yves Zenou who was a source of much sound advices, for his conscientiousness and his willingness to go the extra mile to help a student. Finally, I would like to thank the assessment committee Elena Stancanelli, Olivier Charlot, François Langot and Michael Rosholm for agreeing to review my dissertation.

I also want to thank Gaëlle Heude for her amazing Economics courses in high school that first lured me in the direction of Economics. During my "dissertation years", Anne Bucher gave me friendly support and much joyful laughter. I believe our friendship will continue to stand the test of time and distance. Finally, I thank my parents for their trust in what I am doing. My lasts words go to Richard, husband *extraordinaire*. Thank you.

Résumé français

Cette thèse a pour thème central les liens entre les décisions de formation et de licenciement, ainsi que l'efficience économique de ces décisions. La formation est ici entendue comme formation *spécifique* à l'entreprise dont les salariés bénéficient au cours de leur vie professionnelle *via* leur employeur. Le premier chapitre évalue, à partir de données individuelles, les rendements de la formation en France, à la fois sur le plan de la mobilité (emploi-emploi et emploi-chômage) qu'en termes de gain salarial. Le deuxième chapitre examine comment une protection de l'emploi, différenciée selon l'âge des salariés, affecte la volonté des entrepreneurs de former leurs travailleurs. Le troisième chapitre met en évidence que les décisions de formation et de destruction d'emploi sont fortement complémentaires. Dès lors, des subventions à la formation ciblées par niveau de qualification et combinées à des taxes sur le licenciement (également ciblées par niveau de qualification) doivent être mises en place pour que ces décisions soient socialement optimales. Le quatrième chapitre analyse comment le risque de licenciement, différencié entre les travailleurs d'un même niveau de qualification selon leur niveau d'aptitude, peut être source d'inégalités salariales. Enfin, le dernier chapitre souligne que, face aux disparités salariales, de formation et de risque de licenciement entre les salariés d'un même niveau de qualification, les subventions à la formation et les taxes sur le licenciement, nécessaires à l'efficacité économique, devraient non seulement différer selon la catégorie socioprofessionnelle, mais également au sein de chacune.

Mots-clés : formation professionelle spécifique, incitations à former, rendements de la formation, appariement sur le score de propension, expérience naturelle, holdup, hétérogénéité des travailleurs, destruction d'emploi, volume de formation, dispersion salariale, taxes sur le licenciement, subventions à la formation, efficience économique.

Introduction and summary

This thesis is comprised of five chapters on firm-specific training investments, job destruction and inefficiencies issues. Firm-specific training investments refer to firm-provided training spells that workers benefit from in their adult life while in employment. This thesis is concerned both with the positive aspects of firm-specific training and job destructions, and with their normative implications, but does not going about general training. Throughout, “firm-specific training” is synonymous with “on-the-job training” or “training”.

Becker (1962) is the first reference in the economics of training. In this seminal paper, Becker distinguishes between investments in general-usage and specific human capital, on the basis of the transferability of the acquired skills. As pointed out, this distinction is important if these investments take the form of employer-provided training. Returns to specific training can be realized only in an ongoing relationship with the training firm. Accordingly, the cost and the return should be shared by the worker and the employer to reduce the likelihood of either party unilaterally terminating the employment relationship and imposing on the other party a loss in her return. In this context, if training investments can be preceded by non-renegotiable contracts specifying the sharing decision, there is no need for public policies to encourage training investments since private decisions should be optimal. But this conclusion stands in a competitive labor market. On the contrary, the beginning of the 1990s was a turning point in the theoretical economic literature on training by considering market imperfections and information asymmetries. Strategic interaction between employers and employers changes investment incentives (Leuven (2005)). In particular, in a frictional labor market, wages are determined by an ex-post bargaining and hence contracts are not enforceable. In the renegotiation process, the non-investing party is often able to capture part of the returns. The investor under-invests since she no longer receives the full marginal return on her investment, which finally leads to a "holdup" problem (Malcomson (1997)). Therefore, there is room for training advocacy. Potential for holdups has been studied in case of physical capital (Acemoglu and Shimer (1999)) and general human capital (Sato and Sugiura (2003)) investments. Essays in this thesis are in line with these recent developments in the economics of training but with a focus on firm-specific investments. The step we take relative to this literature is to note the link between firms' job destruction decisions and their incentives to finance firm-specific training in such a context. This brings new insights into the economics of training, both positively and normatively.

It is often hard to clearly distinguish between specific and general training empirically. Some studies rather try to distinguish between formal and informal training, others according to the source of financing. Two main themes are developed empirically: documenting stylized facts about the distribution of training and explaining training access on one hand and evaluating benefits for recipients on the other hand. Among many others, Bassanini et al. (2007) provide accurate information about the distribution of training in Europe, across countries and regions, across firms and across employees. Country of residence, individual and firms characteristics are actually the main determinants of training. In particular, the Nordic countries show the highest participation rates. Besides, in all countries, the low-educated and older workers are under-represented in firm-training programmes. Large and innovative firms train more than small and non-innovative firms. Secondly, the main problem in estimating returns to training concerns the recovery of a causal effect of training net of selectivity bias, since training is very unlikely to be exogenous but may rather pick up returns from unobservable characteristics. The selection problem is recurrent when empirically evaluating a public policy. First estimations about returns to employee training focused on wage returns. At the beginning, Mincer's model of earnings (1974) has been used to relate income distribution in America to the varying amounts of on-the-job training among workers. The instrumental variable method, the control function method and the method of matching are three broad classes of alternative methods that have been developed since. The first two require some excluded instrument that determines training participation but not earnings while the matching method requires an extensive set of observable characteristics on which to match. All place strong demands on data. Some empirical studies have also used these methods to take an interest in the issue of how training affects workers' employment prospects but there are comparatively few. Heckman, Lalonde and Smith (1999) provide a rich overview of available identification and estimation strategies to examine the evidence on the effectiveness of welfare state active labor market policies. Essays in this thesis deal also with both aspects: explaining training access and evaluating benefits for recipients.

This thesis focuses on three main concerns of the economic literature on training: returns to employee training, firms' training incentives and investments efficiency.This requires to understand why and how policies to regulate the training market should be implemented. First, in light of the huge amount of training expenditures and as a public policy concern, one may be interested in knowing what can be quantitatively expected from firm-specific training. This question is also particularly interesting in the French case since the French training system (so-called "pay-or-train" system) with mandatory contributions has been clearly questioned. Cahuc and Zylberberg (2006) particularly underline how both inefficient and inequitable it is. A large part of the empirical literature has been interested in private returns to training but most of studies focus on the impact of training on wages and few are interested in evaluating the training effect on workers' mobility on the labor market (employment to unemployment transitions and job-to-job transitions). Yet, the impact of training on mobility is not necessarily obvious from a theoretical viewpoint because of labor market imperfections. The first chapter *Returns to Firm-provided Training in France: Evidence on Mobility and Wages* brings evidence on the empirical effect of formal training both on these transitions and wages. From French panel data covering the 1998-2000 period, we find that participation to a training program in 1998 strongly reduces both the probability to change firms and to become unemployed during the two succeeding years, once endogeneity of the training participation is taken into account. Both estimates from matching estimators and bivariate Probit models lead to very similar results, although the negative training impact on the employment-unemployment transition is no longer significant with the second specification. Further, on the basis of a correction of selection on observables, we find a significant and positive impact of training on wages. Finally, quantile matching estimators suggest that the wage premium remains rather flat along the wage distribution.

In the second chapter *Stricter Employment Protection and Firms' incentives to Train: The case of French Older Workers*, we use a conditional difference-in-difference estimator to identify the effect of an exogenous change in employment protection among older workers on firms' incentives to provide training. In light of theoretical considerations, this seems to be an important concern since employment protection may have an age-differentiated impact (depending on whether workers are affected or not by the firing tax) due to its anticipation. To our knowledge, no paper has already studied the effect of an employment protection specific to older workers on firms' incentives to engage in training. Laying off workers aged 50 and above, French firms have to pay a tax to the unemployment insurance system, known as the Delalande tax. In 1999, the measure was subjected to a reform that increased due taxes but that did not concern equally all firms. We find that the increase in the Delalande tax for large firms significantly raised the access rate to employer-provided training of treated workers aged 45 to 49 by 11.5 points of percentage. Further, a skill-decomposition of this effect shows that the 1999 reform only had a positive and significant effect on the training rate of less productive workers. According to our theoretical findings derived from a simple labor market model we develop in a first stage, this could result from the persistence of shocks. Indeed, if productivity shocks are persistent and jobs are highly productive, increased employment protection of older workers does not affect firms' incentives to engage in training since jobs are robust to ageing even without investment in training.

In chapter 1, we have highlighted that firm-specific training leads to decrease employment to unemployment transitions. Further, chapter 2 shows that training decisions could be influenced by employment protection because of its impact on job destruction decisions. Therefore, this suggests that job destruction and training investment decisions would be dependent. The third chapter *Inefficient Job Destructions and Training with Holdup* shows how complementary both decisions actually are. In particular, firms may have strong incentives to invest high enough in training to protect matches from idiosyncratic productivity shocks. The analysis carried out in this chapter also points out inefficiencies issues due to contract renegotiation after specific investments have been made (holdup). We use the interplay between firing and training decisions to solve the holdup problem. In particular, we show that this complementarity should lead to reexamine the instruments of economic policy used to bring back efficiency and that training subsidies turn out to be a central instrument. Firing taxes and training subsidies are mostly studied separately but here, it is definitively the combination of those two parameters that achieves efficiency. We show that both policy instruments are not unconnected because of the strong complementarity between firing and training decisions.

This chapter has been accepted for publication in ***Labour: Review of labour economics and industrial relations*** *(joint with Arnaud Chéron), December 2011, volume 25 (4), pages 397-420.*

Chapter 4 *Endogenous Job Destructions and the Distribution of Wages* does not deal with training issues but will be useful for carrying out the analysis in chapter 5. This fourth chapter starts from two recurrent stylized facts that characterize workers in OECD countries, at the aggregate level as well as inside skill groups: a log-normal-like shape of wage distributions and a negative relationship between employment to unemployment transition rates and wage deciles. The goal of this chapter is then to draw a parallel between those two empirical observations in order to highlight the role of firms' decisions about reservation productivity in the wage dispersion analysis. This has been neglected until now since existing models put the emphasis on on-the-job-search, and are usually characterized by exogenous firings. To that end, we consider a matching model with both endogenous job destructions and workers' heterogeneity across *ex ante* unobservable abilities. By considering alternative ranges of productivity shocks in numerical experiments, calibrated to French data, we show that the model can generate a hump-shaped wage distribution. This first relies on the fact that the reservation productivity of low-ability workers is high, which implies that only low-ability workers who draw a good productivity are in a position to keep their job. All else being equal, this raises the average wage of low-ability workers. Secondly, the reservation productivity of high-ability workers is low. Again, all else being equal, high-ability workers whose job has been hit by a bad productivity shock earn lower wages, which move them to the left in the wage distribution.
This chapter has been accepted for publication in ***Labour Economics*** *(joint with Arnaud Chéron), forthcoming.*

Chapter 3 highlighted the need for training subsidies in combination with firing taxes when firms face a holdup problem in their training investments and workers' heterogeneity is observable. The last chapter of this thesis *Training, Job Destruction and Wage distribution* extends this analysis considering *ex ante* unobservable workers' heterogeneity (as in chapter 4) in order to deal both with positive and normative issues. This framework generates a wage distribution, transition rates from employment to unemployment and average training amounts per worker by wage interval at the same time. Confronting model properties to real French data in numerical experiments, I show that this framework is credible enough to deal with the two sources of inefficiency that arise in such a theoretical framework. Indeed, in addition to holdup, the introduction of workers heterogeneity in a matching model with non-directed search -due to ex ante unobservable heterogeneity- implies another source of inefficiency: a composition externality in the search process arises since the composition of the group of the unemployed has an effect on the average expected value of a contact. In particular, the more unemployed workers with high abilities there are, the higher the probability to contact a high-ability worker, and hence the higher the expected return on a vacancy. The optimal labor market policy then consists in implementing both training subsidies and firing taxes as well but, in a first-best approach, both instruments should depend on the ability level.

Chapter 1

Returns to firm-provided training in France: Evidence on mobility and wages[1]

While numerous studies have provided selectivity-corrected estimates of the wage returns on training both in the US and in European countries, less is known about the impact of training on mobility on the labor market. In this paper, we estimate the impact of firm-provided training on both employment-unemployment and job-to-job transitions using French panel data covering the 1998-2000 period. We find that participating to a training session in 1998 reduces the probability to experience an employment-unemployment transition during the period and that the probability to change firms is higher among untrained workers. Additional results about the effect of training on wages indicate that training participation in 1998 increases wages by 7% in 2000 but the wage premium remains flat along the wage distribution.

JEL Classification: C14, J24

Keywords: Returns on training, labor market mobility, Propensity score matching

[1]This chapter reviews a joint work with Arnaud Chéron and Francois-Charles Wolff. I am indebted to two anonymous referees, Salima Bouayad and Xavier Joutard for their very helpful comments. I also would like to thank participants at the EPEE seminar (Evry), Journées Journées LAGV 2008, ESPE Congress 2008 and EEA Congress 2008.

1.1 Introduction

The European Heads of Government in the Lisbon Summit at the beginning of the new millennium strongly committed to make Europe by 2010 as "the most competitive and dynamic knowledge-based economy in the world". The development of high quality vocational training in Europe is a crucial part of this strategy, especially in order to improve and to adapt existing skills to the changes of technology and to promote employability[2]. Firms have a key role in this training investment process since they are the most important provider of on-the-job training. For instance, in France, which is the country under consideration in this paper, training periods were funded by firms in about 86% of cases in 1999 and around three-quarters of the training programs were reported to be at least partially initiated by firms.

Estimating wage returns to training leads to well-known measurement and estimation issues. The main problem concerns the recovery of a causal effect of training on wages net of selectivity bias, since training is very unlikely to be exogenous. Estimated returns to private-sector training may pick up returns from unobservable characteristics. Leuven and Oosterbeek (2008) find that the returns to training tend to drop to zero when comparing workers participating in training and workers who wanted to participate in training, but did not do so because of random events. While most of empirical studies focus on the impact of training on wages, very few studies are interested in evaluating the training effect on workers' mobility. This is really surprising as firms should have stronger incentives to invest in the training of their employees when the latter have no expectations to quit their current job. From a theoretical viewpoint, it is worth noticing that the impact of training on mobility is not necessarily obvious. Considering the transition from a job to another one (i.e. a change of firms), participating in a training program may have two offsetting effects. On one hand, if there are significant wage returns to training, this should reduce the probability for participants to search for new jobs. On the other hand, trained workers could also choose to behave in an opportunistic way by moving to another firm in order to receive a higher paying job with increased responsibilities[3]. These two offsetting effects depend crucially on whether the training is general or specific. In this sense and according to the hu-

[2]For comparative evidence and differences in training practices in Europe, see Bassanini et al. (2007).

[3]By definition, firms have strong incentives to hire recently trained workers since they can reap the benefits of the former training programs without supporting the costs of these programs. This refers to the so-called "poaching externality" (Pigou (1912) and more recently Acemoglu (1997)).

man capital theory (Becker (1964)) -that draws a crucial distinction between general and specific skills- the worker's post training outside option is considerably reduced if the training is specific[4], so that workers with specific skills are not actually in a position to behave in an opportunist way. However, this analysis first rests on the assumption of perfectly competitive labor markets, in which workers receive their marginal product. Due to search frictions (that rise the cost of turnover) or information asymmetries (the training firm is better informed about the training of its employees than outside employers) or else institutions, labor markets are actually imperfect. Therefore, training firms have a monopsony power and are able to capture some of the returns to training. This gives them strong incentives to finance general training. Secondly, most skills may be industry specific but also general because typically there are many firms in the same industry using similar technologies (Acemoglu and Pischke (999a)). Accordingly, even if skills are specific, trained workers are to some extent in a position to behave in an opportunist way. Recent developments in the training literature focus on all these strategic interactions between employers and employees (Acemoglu and Pischke (1998), Acemoglu and Pischke (999b)). Expectations seem more clear-cut when considering transitions from employment to unemployment. The probability to be fired should be reduced among trained workers, given their increase in skills and the training costs supported by the employer.

While knowing the mobility impact of training programs is of importance for both firm managers and public policy markers, empirical evidence remains scarce in European countries. Ok and Tergeist (2003), Budria and Pereira (2004) and Goux and Maurin (2000) are interesting recent exceptions. These authors focus on different transition probabilities on the labor market and find mixed results. Again, the difficulty is to properly account for endogeneity of the training program and recovering the counterfactual (i.e. what would have been the situation of the participants if they had not benefited from the training program) is a challenging task. The purpose of our contribution is to bring additional evidence on the empirical effect of formal training on both employment-unemployment transition and job-to-job transitions (with a change of firms in the latter case), and also on wages using French data. Training expenditures are

[4]Specific skills are defined as those which are only useful in the training firm and increase the worker's productivity only in her current job. In contrast, general skills are also useful with other employers. In fact, in a competitive labor market, firms could never recoup their investments in general skills, so they will never pay for general training.

very important in France. The total amount of expenditures related to initial and continuing vocational training was 25.9 billions of euros in 2005, about 1.5% of GDP. A particular feature of the French funding system for continuing training is the existence of mandatory contributions. All firms have to devote a specific percentage of their total wage bill to train their employees. When this is not the case, they have instead to pay a tax, leading to the so-called "pay-or-train" system[5].

To investigate the consequences of training on the labor market mobility, we use two French data sets focusing on continuous training and on labor participation, both gathered by INSEE, and covering the period from 1998 to 2000. This allows us to study the consequences of training two years after the program. We find that participation to a training session in 1998 reduces the probability to experience an employment-unemployment transition during the period and that the probability to change firms is higher for untrained than for trained workers. Finally, we find that participation in training has a significant and positive impact on wages, but the wage premium remains rather flat along the wage distribution.

The remainder of our paper is organized as follows. In the next section, we provide a brief review of European empirical works on the impact of training, with special emphasis on the worker's mobility. The data used in our study and the corresponding descriptive statistics are presented in section 4.3. We estimate the impact of training on mobility using both matching estimators and bivariate Probit models in section 1.4, while section 1.5 focuses on the wage returns to training. Finally, section 1.6 concludes the paper.

1.2 Literature review on training and mobility

Numerous empirical studies have proposed estimates of the wage returns to training. The main problem concerning the recovery of the causal effect of training on wages lies in the correction for selectivity into training. While many papers have found large returns to formal private sector

[5] In the case of private-sector companies, the contribution amount and the method of calculating it vary according to the size of the company workforce. Mandatory required contributions are equivalent to 1.6% of the total wage bill of companies with a workforce of 20 employees or more The total amount is divided in the following way: i) 0.9% for financing training plans, ii) 0.5% for financing the professionalisation measures and individual training entitlement and iii) 0.20% for financing individual training leave. These contributions are equivalent to 1.05% of the total wage bill of companies with a workforce from 10 to 19 employees, while the contribution of smaller firms (less than 10 employees) is 0.55% of their total wage bill.

training, recent studies based on exogenous variation in training participation find much smaller wage effects of training (Leuven (2005)). Conversely, few studies have focused on the training impact on professional mobility.

Using US data from the National Longitudinal Survey of Youth, Lynch (1991) estimates the effect of different types of training among young workers on the probability of leaving an employer. Estimates from Cox proportional hazard with time-varying covariates show that young people who had some formal on-the-job training are less likely to leave their employer, while those who participated in some form of off-the-job training are more likely to leave. Using the same data, Loewenstein and Spletzer (1999) analyze how job-to-job mobility between 1993 and 1994 is influenced by measures of specific and general training in 1993. General training does not influence the probability that a worker changes jobs, while specific training reduces the magnitude of workers' mobility[6].

In Europe, Ok and Tergeist (2003) find that the probability of being unemployed three years after a training session (that took place in 1994) amounts to 4% among trained workers. The same probability among workers with comparable individual characteristics who did not receive training is about twice higher (7.5%). Their analysis is based on data from Austria, Belgium, Denmark, France, Greece, Ireland, Italy, the Netherlands, Portugal and Spain. However, there is no evidence that trained workers are more likely to stay in the firm than their non-trained counterparts.

Dearden et al. (1997) investigate the relationship between mobility and training using data from the British National Child Development Survey with information between 1981 and 1991 and the UK Labor Force Survey. Receiving either work-related or employer-funded training decreases the men's probability of job switching, while the overall effect of training is not significant among women[7]. Still with data from the UK, Green et al. (2000) examine the multiple factors that affect the impact of training on mobility, mobility being measured through respondents' expectations. While training has on average no impact on mobility, training tends to reduce the likelihood of job search when training is paid by firms.

[6]Using the NLSY data and proportional hazard models, Parent () finds that on-the-job training reduces to a sizeable extent the conditional probability of leaving. However, skills acquired with previous employers are not a significant factor of increased mobility once controlling for unobserved individual heterogeneity.

[7]However, training that is specifically employer-funded appears to decrease mobility among women.

Budria and Pereira (2004, 2007) study the impact of training both on wages and mobility using pooled data from 1998 to 2000 collected in Portugal. They consider training schemes that do not refer to a particular time, so that the training activities may have been completed several years ago. Being trained does not significantly reduce the probability of entering unemployment. However, as pointed out by the authors, these findings have to be interpreted with caution since very few covariates influence the outcome under consideration. This could be due to the fact that the transition equation is poorly specified.

Finally, in France, previous evidence on mobility and training remains scarce. Using the 1993 survey on Education and Qualifications (FQP hereafter) matched with the Corporate Tax Return database, Goux and Maurin (2000) find that the impact of a training session that took place between 1989 and 1992 on the probability of changing firm between 1989 and 1993 is negative, but of small magnitude and not significantly different from zero. Blasco et al. (2008) use the 2003 FQP survey that allows a 60 months follow-up between 1998 and 2003 for each respondent. They focus on the effect of a participation in training spells (either dedicated to employed or unemployed workers) on the employment and unemployment duration distributions. Estimates from a multi-state transition model first show that past participation in training programs increases the conditional probability of return to employment. More surprisingly, they also show that participation in employment training during the previous year (a training period during an employment spell) increases the probability of exiting employment.
Compared to this study, our analysis gives the impact of a training session not only on the employment-unemployment transition, but also on the job-to-job transition. We also evaluate the impact of training on a longer term since we allow training participation to have an impact up to 24 months after completion. Finally, we provide additional results on the effect of training on wages.

1.3 Data and descriptive statistics

1.3.1 The French data

To assess the effect of training on labor market mobility and wages, we use in this study two complementary French databases. The first one is a cross-sectional survey entitled “Formation

Continue" conducted by INSEE (the French National Institute for Statistics and Economic Studies) in March 2000. It was carried out on a sample of 28667 individuals. The main interest of this survey is that it includes detailed information on training. In particular, it provides accurate information on the different training periods followed by each respondent along the life cycle. From the questionnaire, it is possible to separate three main periods: i) from graduation to February 1998, ii) from March 1998 to December 1998, and iii) from January 1999 to March 2000. The "Formation Continue" survey describes the type of experienced training using four main categories: i) training in a work situation, ii) apprenticeships, iii) work placement or training courses and iv) self-training. We have also information about the purpose of the training activity, which has to fit in one of the following categories: i) to adapt to the job, ii) to switch to another job or to get a job, iii) to obtain a diploma or a certification, iv) to execute political duties, v) personal or cultural reasons, and vi) no specific reason. When turning to the data, we choose to only consider the first motive since it was the result of 76% of the training sessions taken by employees in 1998[8]. More precisely, 64% among them have participated to a training program in order to adapt to their job as a result of a change in the content of their work. This reflects the fact that new technologies and organizations require continuing learning. In so doing, we only account for work-related training which is expected to have some influence on the labor market situation of the respondents. Furthermore, we decide to only consider firm-financed training, which was the case of 86% of the employees having participated in a training program in 1998[9]. Finally, we also know that 86% of the training sessions taken by employees in 1998 did not to lead to any recognized certification. This fact corroborates a theoretical point previously mentioned about asymmetric information. The monopsony power of the current employer is more important when the training content is not obvious for the market. Certifications could have decreased this uncertainty.

Keeping those facts in mind, it must be well understood that the predominance of a sort of training in the data does not allow us to provide the econometric estimation for different

[8]Concerning the other motives, 2.1% of the employment training sessions were followed in order to switch to another job (or to get a job), 4.8% in order to obtain a diploma, 0.5% in order to execute political duties, and 7.3% for personal or cultural reasons. Finally, 8.1% of training spells were followed without specific reasons.

[9]Concerning the other participation in training, 7% of them have been financed by workers themselves, 2% by regions, 2.5% by the State, 1% by suppliers, 1% by associations, 1% by the institution that manages the funds for training leaves and 0.5% by other sources.

categories of training. Of course, the motive 'to obtain a job' is likely to contain more general training than "to adapt to her job", but again, due to the small number of observations for other motives than "to adapt to her job", we were not able to differentiate the training between specific or general.

We merge these data with the labor force survey ("Enquête Emploi") conducted over the 1998-2000 period. This is a rotating panel since exactly one-third of the sample is dropped from the sample each year and is replaced with a new, comparable sample drawn from the current population. The size of the "Emploi" survey is about 135000 individuals who are interviewed about their situation on the labor market. The main feature of these data is that they provide detailed information over three-years for one-third of the 1998 original sample. This means that we can investigate the effect of formal training received in 1998 both in terms of labor market mobility and wages two years after the program. For each employee, the "Emploi" survey provides detailed individual characteristics including gender, age, marital status, citizenship, level of education, place of residence, years spent in the firm, type of job contract, number of worked hours, occupation and sector of activity among other covariates. In what follows, we focus on the two following variables of interest. The first one is about mobility on the labor market. We define two dummy variables respectively associated to transition from employment to unemployment and to job-to-job transition. The second outcome is the monthly wage level, expressed in euros. Another important question is about the type of sector, either private or public. In our empirical analysis, we only focus on the former case since wages are essentially fixed by the French legislation in the public sector. As a consequence, they are not necessarily responding to productivity reasons. Furthermore, mobility is very infrequent in the public sector.

We define our sample in the following way. First, we focus on the population of respondents who were working in March 1998 and consider their participation to a training spell between March 1998 and December 1998. They have thus to remain employed during this period. We get information on the various training periods that they have experienced (if any) during ten months and on their professional mobility and wages during the two following years (from March 1998 to March 2000). Secondly, we choose to select the subsample of workers aged from 18 and 60 and exclude farmers and self-employed. As part-time is not infrequent in France, we decide to account for both full-time and part-time jobs. Thirdly, we delete the few observations with missing values, mainly because of missing wages. These different selections leave us with a 'restricted' sample comprising 5107 observations. However, considering a labor market outcome two years after a training program may set some problems. In particular, it could be argued that either mobility between 1998 and 2000 or wages in 2000 are likely to be influenced by training periods experienced after 1998 by the respondent. For instance, if we observe a case of job mobility between 1998 and 2000 for a person who has no training period in 1998, we cannot rule out the possibility that the job mobility is linked to training received in 1999 or even in the early months of 2000. So, for the sake of robustness and to avoid misleading conclusions, we construct another sample where individuals have a job during all the period and construct a training variable which is equal to one when the worker has benefited from training at least once during the 1998-2000 period. This "extended" sample includes 4761 workers[10].

1.3.2 Descriptive statistics

When considering participation in a training program from April 1998 to December 1998 (our "restricted" sample), we find that 762 individuals have benefited from such training experience. The participation rate is hence 15.0% (762/5107). According to the French data, the mean duration of the training period is 41.4 hours. Furthermore, nearly 28% of these training experiences are reported to be in a work situation. We also find that the training spells in a work situation are significantly longer than that of the other types of training. The mean duration is about 52.1 hours for the former, but only 37.2 hours for the latter. As expected, we find a much higher proportion of workers having benefited from a training period when we extend the possibility of participating in such program during the period covering April 1998 to March 2000. The proportion of trained workers is about twice higher within the 'extended' sample, 32.4% (1541/4761) instead of 16%. Interestingly, we find that a significant proportion of the respondents have participated more than once in the training programs over the period. Among the 1541 recipients, 48.3% of them have benefited from training activities both in 1998 (from March to December) and in 1999-2000[11]. It thus matters to account for the fact that workers

[10]Note that the size of this 'extended' sample is lower than that of the "restricted" sample since we only focus on individuals who had a job during the whole period.

[11]The proportions of trained workers who have benefited from only one training period is respectively 19.7% in 1999 and 31.9% in the 1999-2000 period.

may have benefited from several training periods.

Our main outcome of interest is related to mobility on the labor market. As shown in Table 1.6 in appendix, the probability of observing a transition following the training period is somewhat low. The proportions of workers reporting either job-to-job transition or employment to unemployment mobility are respectively equal to 7.1% and 3.3% in the restricted sample. Interestingly, there are large differences depending on whether workers have participated in training activities in 1998 or not. Both rates of mobility are much lower among trained workers, respectively 3.7% instead of 7.7% for job-to-job transition (a decrease of 51.9%) and 1.1% instead of 3.7% for employment-to-unemployment transition (a decrease of 70.3%)[12].

While the differences in labor mobility that we observe from the data may be a consequence of the training program, they are presumably strongly affected by the non-random participation in training activities. The characteristics of the untrained and trained workers reported in Table 1 show that it is important to account for selectivity when assessing the consequences of participation in training. On average, the proportion of male workers involved in training activities is slightly higher than that of female workers (59.3% instead of 40.7%). Participation is more frequently observed among middle-aged workers (from 30 to 49 years old) and among workers having spent more than 10 years in their firm.

There are also substantial differences depending on education. The proportion of low-educated respondents (no diploma or BEPC) is for instance equal to 38.6% among untrained workers, while it is 18.5% among trained workers. Conversely, more than 32% of the participants are high-educated (undergraduate, graduate, postgraduate studies) instead of 13.2% among non-participants. Job characteristics also matter, since the participation in training activities is more frequent among permanent contracts, full-time workers, executives and intermediary professions, and among workers in large firms (more than 500 employees)[13].

To summarize, these descriptive statistics show that the propensity to participate in training activities is more important among workers endowed with high earnings generating characteris-

[12]The figures associated to job-to-job mobility are very similar when considering the extended sample, although there are fewer differences between untrained workers and trained workers when considering job-to-job transitions. Since we account for the possibility of training periods over the whole period, then some workers may have participated in training activities in their new location, i.e. after the job-to-job mobility.

[13]Among participants, 14.2% of them are working in firms with less than 20 employees and 46.9% in firms with more than 500 employees. The same figures are respectively 36.5% and 23.5% among non-participants.

tics. It thus matters to account for this positive selection into training since it is likely to bias the effect of training on both mobility and wages.

1.4 The effect of training on mobility

To address the selectivity issue, we rely on two different methods. As the participation in the training program is expected to be a function of both observable and unobservable characteristics of each worker (like a high degree of productivity or commitment), we ideally need to rely on an instrumental variable strategy. The difficulty here is to find a variable that would be strongly correlated to participation in training, but not with labour market outcomes (either mobility or wages). Given the lack of appropriate instruments in the French data, we decide to first apply matching estimators that only control for observed heterogeneity. Then, we turn to a bivariate Probit model to solve the endogeneity problem of the training variable.

1.4.1 A propensity score matching analysis of mobility

As a preliminary approach, we assume that participation in the training program is only influenced by individual characteristics which are observed from the data. This allows us to rely on matching estimators (see Heckman et al. (1998)). This method consists in building a control group of non-treated individuals whose characteristics are very similar to those of individuals of the treatment group (i.e. trained workers), and then in comparing the outcomes of the two groups. Differences in outcomes of treated and non-treated workers are attributed to the training program. The matching estimator controls for the selection bias at the entry of programs since it mimics random assignment through the construction of a control group.

Let us briefly describe the evaluation methodology. We denote by T the treatment variable which is equal to 1 when the individual has participated in a training program and to 0 otherwise. Let X be a vector of observed individual characteristics and Y be the outcome under consideration, either job-to-job mobility or employment-unemployment transition in our case: Y_1 refers to the outcome of trained workers, while Y_0 is the outcome of non-trained workers. In what follows, we focus on the average treatment effect on the treated, defined by $\Delta = E(Y_1|T = 1) - E(Y_0|T = 1)$. By definition, the outcome of non-treatment for treated workers is never observed from the data.

To estimate the counterfactual $E(Y_0|T=1)$, the key assumption when considering matching estimators is that both the treated and untreated groups have to be comparable conditionally on observed characteristics X. Once the conditional independence assumption is satisfied, then the counterfactual is $E(Y_0|T=1) = E(Y_0|T=0)$. Following Rosenbaum and Rubin (1983), we match treated and untreated workers on the basis of their propensity score. We proceed in the following way. First, we calculate for each worker the propensity score from a Probit regression explaining the probability to participate in training. The set of covariates is chosen in order to satisfy the conditional independence assumption. Then, we select the common support of the densities of the two groups. Finally, we estimate the causal effect of the training program using a Kernel matching estimator (Heckman et al. (1998))[14].

Let us first consider selection into training. To calculate the propensity score for each worker, we regress the training participation on a set of socio-economic characteristics. The different covariates introduced into the regression are gender, age, marital status (in couple versus alone), education, nationality, part-time job, temporary contract, whether the worker holds a second activity, occupational dummies, dummies for tenure, dummies for firm size, and sectoral and regional dummies. Finally, we choose to include in the Probit regression the wage level observed in March 1998, i.e. before the training spell starts. This covariate is of course expected to be highly (positively) correlated with the training assignment, but it also could pick up part of the unobserved heterogeneity related to the worker effort and motivation prior to the participation in the program[15].

The results of the propensity score matching analysis are described in Tables 1.1 and 1.2. Let us first focus on the Probit estimates of the training participation equation (table 1.1). We estimate separate regressions for the job-to job and employment-unemployment transitions as our control group is always made of immobile workers. We exclude the cases of employment-unemployment mobility when focusing on job-to-job transitions, while respondents concerned with job-to-job transitions are excluded when investigating the employment to unemployment

[14]This method uses weighted averages of all individuals in the control group to construct the counterfactual outcome. Weights depend on the distance between each individual from the control group and the participant for which the counterfactual is estimated.

[15]We are indebted to Daron Acemoglu for this suggestion. Additional results (not reported) show that very similar results are found when estimating the participation regression without the prior-training wage.

transitions. This leaves us with two samples comprising respectively 4940 and 4743 observations.

According to the French data, there are no gender differences in participation. Training activities are less frequent among older workers (above 50 years old) and among workers of foreign origin. Those who live in couple are less likely to receive training, but this effect is only significant when the focus is on mobility from employment to unemployment. As expected, we find a positive effect of the different educational dummies on the probability of having been trained in 1998. Training is much more likely among high educated workers, although there is no significant difference between under-graduated and graduated workers. Finally, the participation in a training program is more frequent among intermediary occupations and among employees. Participation in training is not affected by job seniority, having a temporary contract, working full-time or having a secondary activity. The only characteristic of the job that influences the probability of training is the size of the firm. There is a positive correlation between training activities and firm's size, and participation is much more likely when workers operate in large firms (more than 500 employees). The last finding is the positive effect of the wage level (prior to the training period) in the training equation. It is of course more profitable for firms to invest in their more able and more productive employees.

Table 1.2 includes the results of the propensity score analysis for the two transitions. Let us first focus on the case of job-to-job mobility. Under the exogeneity assumption of participation in training (unmatched estimate), we find a difference of -4.3% for the mobility rate between the treated and the control groups. Once selection into training is taken into account of the basis of observable individual characteristics, we get a lower value for the difference in mobility. The causal effect is now equal to -2.9%, but still statistically significant. Very similar findings hold for the employment-unemployment mobility. While the unmatched difference is equal to -2.8%, we get a value of -1.8% for the average effect of the treatment on the treated which is significant at the 1 percent level. So, a first conclusion drawn from the French data is that transitions both from one job to another job and from employment to unemployment are less likely among trained workers. While this pattern holds after controlling for selection into training, a shortcoming of the matching estimators is that they only account for observable characteristics to tackle the selectivity issue.

Table 1.1: Propensity score matching analysis on mobility: Selection into training (Probit estimates)

	Job-to-job mobility		Employment-to-unemployment mobility	
Constant	-4.918***	(8.76)	-5.238***	(8.88)
Female	-0.058	(0.94)	-0.042	(0.67)
Age (ref: 18-29)				
30-39	-0.051	(0.60)	-0.069	(0.79)
40-49	-0.093	(1.00)	-0.111	(1.18)
> 50	-0.367***	(3.34)	-0.409***	(3.66)
In couple	-0.084	(1.59)	-0.109**	(2.01)
No French citizenship	-0.329**	(2.27)	-0.332**	(2.21)
Education (ref: no diploma)				
CAP-BEP	0.193***	(3.05)	0.206***	(3.22)
Baccalaureate	0.329***	(3.82)	0.358***	(4.13)
Undergraduate	0.424***	(4.59)	0.437***	(4.65)
Graduate, postgraduate	0.485***	(4.04)	0.429***	(3.46)
Occupation (ref: Workers)				
Executives	0.195*	(1.73)	0.204*	(1.76)
Intermediary	0.357***	(4.72)	0.334***	(4.35)
Employees	0.194**	(2.49)	0.194**	(2.46)
Job seniority (ref: 5 years)				
6-10	0.121	(1.72)	0.098	(1.35)
11-20	0.064	(0.87)	0.055	(0.73)
> 20	0.144	(1.66)	0.119	(1.35)
Permanent contract	0.075	(0.51)	-0.017	(0.11)
Part-time job	-0.102	(1.11)	-0.094	(1.01)
Firm size (Ref: 0-19)				
20-99 employees	0.319***	(4.24)	0.323***	(4.16)
100-499 employees	0.526***	(6.98)	0.533***	(6.91)
500 employees	0.735***	(10.62)	0.758***	(10.69)
Secondary activity	-0.209	(1.27)	-0.198	(1.19)
Log wage in 1998	0.499***	(6.02)	0.539***	(6.27)
Number of obs.	4940		4743	
Log likelihood	-1793.6		-1731.6	

Source: French Training Survey "Formation Continue 2000", authors' calculations
Note: Estimates of the training selection are from Probit models. The training equations also include a set of sectoral dummies and regional dummies. Absolute value of t statistics are in parentheses, significance levels being respectively 1% (***), 5% (**) and 10% (*).

Table 1.2: Propensity score matching analysis on mobility: The effect of training on mobility

	Treated	Controls	Difference	
Job-to-job mobility				
Unmatched effect	0.037	0.080	-0.043***	(4.18)
Causal effect	0.037	0.066	-0.029**	(3.10)
Employment-to-unemployment mobility				
Unmatched effect	0.011	0.039	-0.028***	(3.89)
Causal effect	0.011	0.029	-0.018***	(3.10)

Source: French Training Survey "Formation Continue 2000", authors' calculations
Note: Absolute value of t statistics are in parentheses, significance levels being respectively 1% (***) and 5% (**).

1.4.2 A bivariate Probit specification

Although the two French databases include detailed individual characteristics, the possibility that participation in training is related to less visible factors like motivation or ability can definitely not be ruled out. Also, as we do not have matched employer-employee data, the role of the firm characteristics will not be adequately controlled for in our regression. Unfortunately, there is no good reliable instrument in our data, so that we are not able to turn to an IV analysis. However, since both the treatment variable and the labor market transition (either from job-to-job or from employment to unemployment) are binary, we are able to take the endogeneity problem into account using a recursive, simultaneous equations model.

Let us describe the corresponding framework. A first equation indicates the probability for an individual to participate in training. Denoting by T^* the latent variable associated to the training decision (it can be either negative or positive), the training equation will be explained by a set of exogenous worker's characteristics X_T in the following way:

$$T^* = X_T \beta_T + \varepsilon_T \tag{1.1}$$

where β_T is a vector of parameters to estimate and ε_T is a random perturbation (normally distributed). Now, let M^* be a latent variable associated to mobility on the labor market. A second equation indicates that mobility depends on a set of exogenous covariates X_M and on the participation into the training program T:

$$M^* = X_M\beta_M + \delta_T T + \varepsilon_M \tag{1.2}$$

where β_M is the vector of associated parameters, δ_T measures the impact of the training decision on mobility, and ε_M is an error term. While the latent variables M^* and T^* are not observed, we have some information on their observed counterpart since we know that $M = 1$ when $M^* > 0$ (and $M = 0$ otherwise) and $T = 1$ when $T^* > 0$ (and $T = 0$ otherwise).

Under the assumption that the two residuals ε_T and ε_M follow a bivariate normal distribution such that $(\epsilon_T, \epsilon_M)\ N(0, 0, 1, 1, \rho)$, with ρ the coefficient of correlation between ε_T and ε_M, equations (1.1) and (1.2) define a recursive simultaneous model which comprises two Probit equations. This bivariate Probit model can easily be estimated by full information maximum likelihood. As shown in Greene (1998), the simultaneity problem does not matter when the two dependent variables are jointly determined in the bivariate Probit specification.

A central problem when estimating such simultaneous models concerns identification. While it is often argued that exclusion restrictions are needed to identify such models[16], Wilde (2000) has shown that this condition was not necessary in the context of a two equations Probit model with one endogenous dummy regressor. Each single equation has just to include at least one varying exogenous variable in the list of covariates. Nevertheless, in that case, it should be noted that the bivariate Probit model is only weakly identified through the non-linear distribution of the two residuals, while strong identification requires inclusion of additional explanatory factors in the training participation equation.

When turning to the data, we proceed in the following way to secure identification. On the one hand, we introduce into the training equation the different covariates reported in Table 2 and add a dummy variable which is equal to one when there is a company training plan. This variable is expected to be positively correlated with participation in training at the individual level, while it should have no influence on the employment transitions. On the other hand, we introduce into the mobility equation both individual and job characteristics (except sectoral dummies) . We also control for the number of young children living in the household and for the wage level prior to the training program (if any), as individuals receiving high wages may be less tempted

[16] This is indeed the case with a standard IV specification where the dependent variable of the recursive model is continuous instead of binary.

to quit their current job.

We report in Tables 1.3 and the bivariate 1.4 Probit estimates respectively for job-to-job mobility and for transition from employment to unemployment. Before turning to the impact of training on mobility, a few comments are in order. First, with respect to the previous estimates of Table 1.2, we find that the probability to participate in a training program is significantly increased among workers in firms with a training plan. Secondly, there are differences in the determinants of the two types of mobility.

When considering the case of unemployment, only seniority and the type of contract have a significant influence. Occurrence of unemployment is much higher among workers who have spent less than 10 years in their firm and among those who have a temporary contract. A similar pattern holds for seniority and type of contract when explaining job-to-job mobility. Women and older workers are also less likely to experience such mobility, which is conversely more likely among executives. This is not really surprising as executives may benefit from higher wage opportunities when changing firm.

Let us now asses the role of training participation. As shown in Table 3, we find a negative coefficient for the endogenous participation dummy both for job-to-job and employment-unemployment transitions. However, the coefficient associated to training is only significant when considering job-to-job transitions, whilst it is not significant at any conventional level for changes in the employment status. With respect to the matching estimators, the difference could be due to the fact that we now control for unobserved heterogeneity when using the bivariate Probit specification. However, it should also be kept in mind that the number of respondents facing mobility from employment to unemployment remains low, which could (at least partly) explain the lack of significance of the training dummy in the mobility equation.

Finally, we assess the magnitude of the causal effect of training on the job-to-job transition using the bivariate estimates. In this setting, the average treatment on the treated Δ is $\Delta = Pr(M = 1|T = 1) - Pr(M = 1|T = 0)$. It is then straightforward to calculate these two conditional probabilities, given by $Pr(M = 1|T = 1) = \frac{\Phi_2(X_M\beta_M+\delta_T, X_T\beta_T, \rho)}{\Phi(X_T\beta_T)}$ and $Pr(M = 1|T = 0) = \frac{\Phi_2(X_M\beta_M, -X_T\beta_T, -\rho)}{\Phi(-X_T\beta_T)}$ by definition of the bivariate Probit model. When calculating Δ on the subsample of trained workers, we find a value of -1.1% for the training effect on job-to-job transition and of -2% on mobility to unemployment which are not so different from our previous

Table 1.3: Bivariate Probit estimate of the effect of training on job-to-job transitions

	Training		Mobility	
Constant	-4.453***	(7.77)	-0.042	(0.06)
Female	-0.071	(1.16)	-0.291***	(3.58)
Age (ref: 18-29)				
30-39	-0.030	(0.36)	0.000	(0.00)
40-49	-0.058	(0.63)	-0.199**	(2.16)
50	-0.330***	(2.99)	-0.499***	(4.13)
In couple	-0.097*	(1.84)	0.011	(0.17)
Number of children less than 6			0.069	(1.41)
No French citizenship	-0.282**	(1.96)	-0.114	(0.97)
Education (ref: no diploma)				
CAP-BEP	0.172***	(2.71)	0.008	(0.11)
Baccalaureate	0.317***	(3.68)	0.033	(0.29)
Undergraduate	0.396***	(4.27)	0.097	(0.78)
Graduate, postgraduate	0.463***	(3.86)	0.236	(1.51)
Occupation (ref: Workers)				
Executives	0.210*	(1.85)	0.369***	(2.62)
Intermediary	0.344***	(4.53)	0.152	(1.30)
Employees	0.218***	(2.77)	0.141	(0.15)
Job seniority (ref: 5 years)				
6-10	0.087	(1.22)	-0.438***	(4.56)
11-20	0.018	(0.24)	-0.722***	(5.55)
20	0.072	(0.83)	-0.699***	(4.39)
Temporary contract	0.053	(0.37)	0.646***	(5.58)
Part-time job	-0.083	(0.90)	-0.069	(0.66)
Firm size (Ref: 0-19)				
20-99 employees	0.287***	(3.80)	0.096	(1.22)
100-499 employees	0.409***	(5.37)	-0.057	(0.47)
500 employees	0.592***	(8.33)	-0.019	(0.13)
Secondary activity	-0.146	(0.91)	-0.039	(0.26)
Log wage in 1998	0.412***	(4.89)	-0.076	(0.75)
Existence of a company training plan	0.386***	(6.66)		
Endogenous training participation			-1.384***	(4.42)
Number of observations			4940	
Coefficient of correlation (t-test)			0.686	
Log likelihood			-2842.0	

Source: French Training Survey "Formation Continue 2000", authors' calculations
Note: Bivariate Probit estimate. Both equations also include a set of regional dummies and a set of sectoral dummies are introduced into the training participation equation. Absolute value of t statistics are in parentheses, significance levels being respectively 1% (***), 5% (**) and 10% (*).

Table 1.4: Bivariate Probit estimate of the effect of training on employment-to-unemployment transitions

	Training		Mobility	
Constant	-5.029***	(8.52)	-0.754	(0.91)
Female	-0.046	(0.72)	-0.036	(0.34)
Age (ref: 18-29)				
30-39	-0.043	(0.49)	0.061	(0.51)
40-49	-0.080	(0.84)	0.111	(0.83)
50	-0.384***	(3.40)	-0.122	(0.73)
In couple	-0.111**	(2.05)	-0.056	(0.58)
Number of children less than 6			-0.094	(1.08)
No French citizenship	-0.303**	(1.99)	0.030	(0.19)
Education (ref: no diploma)				
CAP-BEP	0.194***	(2.98)	-0.023	(0.22)
Baccalaureate	0.345***	(3.94)	0.185	(1.27)
Undergraduate	0.414***	(4.37)	-0.135	(0.74)
Graduate, postgraduate	0.409***	(3.28)	-0.012	(0.05)
Occupation (ref: Workers)				
Executives	0.181	(1.56)	-0.041	(0.17)
Intermediary	0.314***	(4.05)	0.109	(0.75)
Employees	0.193**	(2.42)	-0.029	(0.24)
Job seniority (ref: 5 years)				
6-10	0.072	(0.98)	-0.381***	(3.46)
11-20	0.013	(0.17)	-1.016***	(5.54)
20	0.061	(0.68)	-0.999***	(4.55)
Temporary contract	0.010	(0.07)	1.497***	(12.73)
Part-time job	-0.079	(0.84)	-0.099	(0.68)
Firm size (Ref: 0-19)				
20-99 employees	0.293***	(3.72)	-0.081	(0.69)
100-499 employees	0.432***	(5.45)	-0.074	(0.55)
500 employees	0.623***	(8.42)	-0.198	(1.24)
Secondary activity	-0.178	(1.06)	-0.218	(0.91)
Log wage in 1998	0.487***	(5.57)	-0.051	(0.39)
Existence of a company training plan	0.389***	(6.97)		
Endogenous training participation			-0.617	(1.22)
Number of observations	4743			
Coefficient of correlation (t-test)	0.115			
Log likelihood	-2223.0			

Source: French Training Survey "Formation Continue 2000", authors' calculations
Note: Bivariate Probit estimate. Both equations also include a set of regional dummies and a set of sectoral dummies are introduced into the training participation equation. Absolute value of t statistics are in parentheses, significance levels being respectively 1% (***), 5% (**) and 10% (*).

matching results (respectively -2.9% and -1.8%)[17].

1.5 Additional results on wages

The data also allow us to investigate the wage effect of training. According to the human capital theory, training increases worker's productivity. Accordingly, if training is mainly specific, wages growth only depend on the productivity growth and the mobility between jobs is small. In contrast, if training is mainly general, workers' outside options increase (as productivity increases), leading either to a wage increase in the same firm i.e. without mobility or to a wage increase in another firm i.e. following a change of employer.

Nevertheless, according to the new literature on training following Acemoglu and Pischke (999b), wage growth can be small even in case of general training due to the monopsony power of the training firm: the marginal increase in the worker's productivity is not fully reflected in the best market opportunity due to market imperfections. As suggested, wage growth may also be influenced by job mobility. Consequently, we restrict our attention to the subsample of workers not concerned by mobility during the two years following training participation. Given the lack of suitable instruments to control for unobserved heterogeneity, we focus in what follows on matching estimators to assess the wage effect of participation to a training program in 1998.

We begin with a brief description of wages distribution of trained and untrained workers. As shown in Table 1.5, there are large differences in the mean wage between the two groups. On average, respondents who have participated in a training program benefit from earnings which are 31.8% higher in 2000 than those of the untrained workers. Although the wage gap is large, part of the difference is expected to stem from differences in individual characteristics between untrained and trained workers. A disaggregated analysis by gender further indicates that the wage gap due to training is much lower among men than among women, respectively 24.4% instead of 39.8%. This difference may be linked to the selectivity of women into both the labor market and training activities.

When evaluating the returns to training participation, we again rely on a propensity score

[17]When considering the bivariate Probit coefficients, we find that the average treatment on the treated is respectively given by 6.62-7.72=-1.1% in the case of job-to-job transition, while it is 1.8-3.77=-1.97% in the case of employment-unemployment mobility.

Table 1.5: Matching estimates of training participation (in 1998) on wages

Monthly wage	Men		Women		All	
in 2000 (log)	Unmat.	Matched	Unmat.	Matched	Unmat.	Matched
With wage in 1998 as control						
Treated	7.416	7.411	7.193	7.192	7.323	7.323
Control	7.172	7.353	6.795	7.118	7.005	7.256
Difference	0.244***	0.058***	0.398***	0.074**	0.318***	0.067***
Abs. t-value	(12.26)	(2.62)	(11.37)	(2.38)	(15.81)	(3.46)
Without wage in 1998 as control						
Treated	7.416	7.415	7.193	7.192	7.323	7.323
Control	7.172	7.326	6.795	7.069	7.005	7.216
Difference	0.244***	0.089***	0.398***	0.123***	0.318***	0.107***
Abs. t-value	(12.26)	(4.04)	(11.37)	(3.97)	(15.81)	(5.58)

Source: French Training Survey "Formation Continue 2000", authors' calculations.
Note: Kernel matching estimates, significance levels being respectively 1% (***) and 5% (**). The list of covariates included in the training participation equation is described in Table 1.1.

analysis and on Kernel matching estimators. We introduce in the selection equation the list of covariates considered in the mobility analysis (gender, age, marital status, education, nationality, part-time job, temporary contract, secondary activity, occupation, tenure, firm size, sector, region). As the training decision is likely to be influenced by unobservable characteristics as well, we also include former training participation (before March 1998) as a proxy for the general attitude of individuals towards training. We estimate two different set of regressions, one with the wage level in 1998 and one without[18]. A lower matched difference is expected in the former case since the wage level before training should pick up part of the unobserved heterogeneity such as motivation. Results from the matching estimates are in Table 1.5. Several comments are in order.

First, we find a much lower value for the wage effect once selection into training is controlled for (on the basis of observable characteristics). The average effect of the treatment on the treated is equal to 6.7% when the wage level in 1998 is introduced in the participation equation (first specification). By comparison, the unmatched difference was more than four times higher (31.8%). Secondly, as expected, we find a much large matched difference when the wage level before the training program is not included in the participation equation (10.7% instead of 6.7%). Since the wage level in 1998 is correlated with unobserved characteristics of the worker

[18]We do not report the Probit estimates explaining participation into training since our results are very similar to those described in section 1.4.

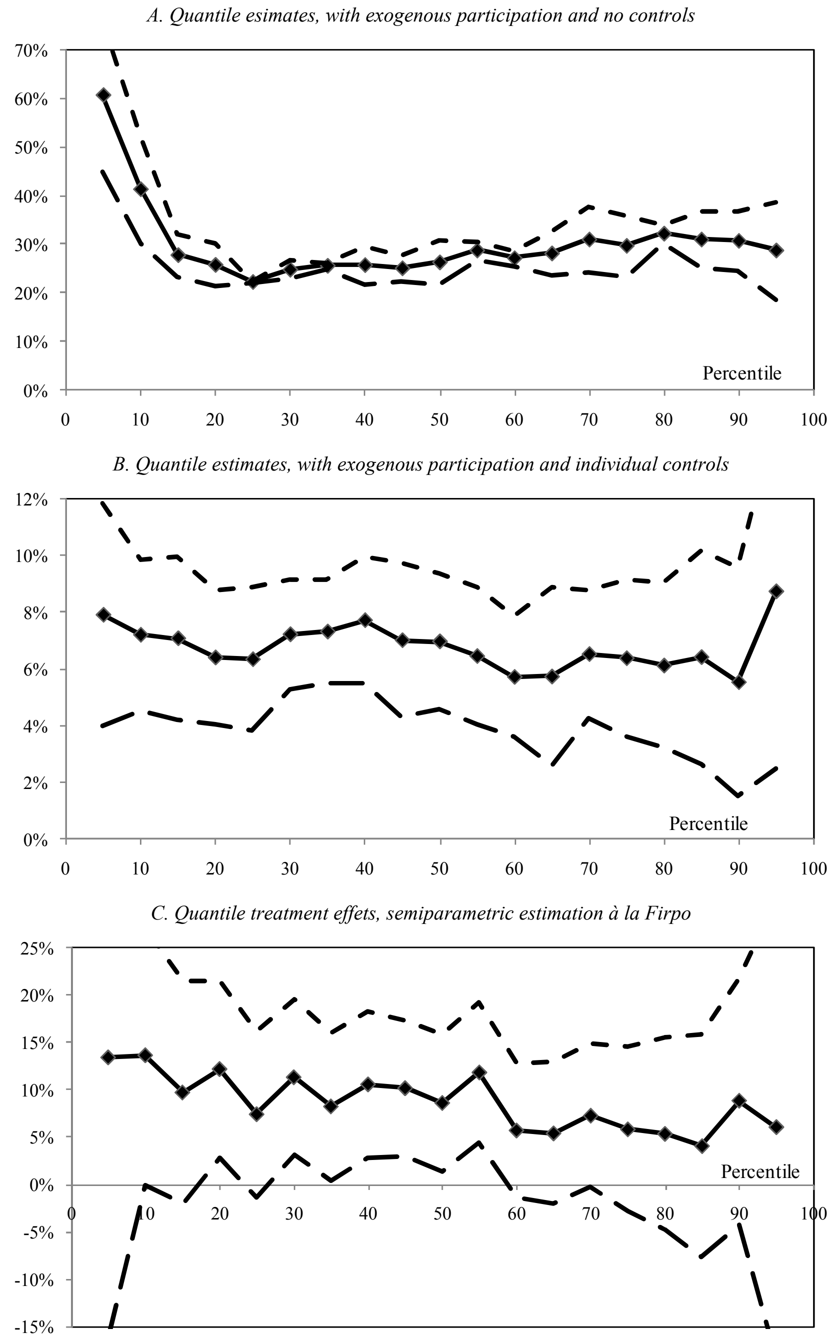
before participation in the program (like motivation or ability), then the returns to training will be overestimated once the unobserved factors are partially controlled for. Thirdly, we highlight higher returns to training among women (7.4%) than among men (5.8%)[19]. These results suggest that training participation leads to a higher productivity for non-job switchers, suggesting that the training is mainly general. Nevertheless, the small number of workers who have switch a job after training completion does not allow us to have a lot to say about the monopsony power of training firms.

As recent studies in program evaluation have suggested that it matters to learn not only about the average treatment effect but also on the tails of the outcome distribution, we turn to a quantile regression framework (Koenker and Bassett (1978)). Conditional mean regressions (OLS regressions), which indicate the mean effect of the impact of a covariate, fail to describe its distributional impact. However, as suggested in Chernozhukov and Hansen (2005), the effect of training activities on the low tail of the wage distribution may be of more interest for public policy makers than the effect of participation on the mean of the wage distribution. Their estimates from instrumental variable quantile regressions show that in the US, the percentage impact of the training program on earning is quite stable along the distribution. Let us begin with a descriptive analysis of the wage gap between trained and untrained workers along the earnings distribution. As shown in Panel A of Figure 1.1, we find a large percentage increase in wages in the low earnings quantiles, which declines as one moves to the upper quantiles of the distribution. Above the 20th percentile, the difference in earnings remains rather flat and is comprised between 25 and 30 percent. Then, we add a set of individual characteristics in the quantile regressions. As expected, we now find a much lower earnings difference between non-participants and participants. According to Panel B of Figure 1.1, the gap varies between 6% and 8% along the earnings distribution and is fairly flat . Nevertheless, the moderate positive and significant effect of training on earnings quantiles does not account for the endogeneity of training status.

Since there is no relevant instrument in the data to control for the selection into the treatment, we again assume that selection to treatment is exclusively based on observables (the unconfound-

[19] As shown in Table 1.5, we get a larger gap on the basis of the matched estimators when the wage level in 1998 is excluded from the regression, the matched returns being respectively equal to 8.9% among men and 12.3% among women.

Figure 1.1: Returns to training along the wage distribution, with confidence intervals



Source: French Training Survey "Formation Continue 2000", authors' calculations.

ness assumption) and rely on the efficient semiparametric estimator recently proposed by Firpo (2007). Estimation of quantile treatment effects is implemented using a two-step procedure, with first a non-parametric estimation of the propensity score and then a computation of the difference between quantiles for the treated and for the control individuals. The corresponding estimates are in Panel C of Figure 1.1. Two comments are in order.

On the one hand, with respect to the previous quantile estimates with exogenous participation, we now find slightly higher values for the training benefits, at least in the first part of the distribution. The wage gap between non-participants and participants, which is for instance equal to 14% at the 10th percentile, then declines when moving to the upper quantiles. Above the 60th percentiles, the percentage impact of the training program varies between 5 and 7 percent. On the other hand, it is also clear that the quantile treatment effects parameters are most often imprecisely estimated. The training benefits are not significantly different from zero both in the lower part (till the 20th percentile) and in the upper part (above the 60th percentile) of the earnings distribution.

1.6 Conclusion

In this paper, we have focused on the consequences of training participation on labor market mobility using French data collected between 1998 and 2000. Considering both job-to-job mobility and transition from employment to unemployment, our estimates show that the participation in a training program in 1998 reduces the probability either to change firms or to become unemployed during the two succeeding years. The magnitude of mobility, which is somewhat low (around 7% from job-to-job transitions and 3% from employment to unemployment transitions), is divided by about two once endogeneity of the training participation is taken into account. Interestingly, our estimates from both matching estimators and bivariate Probit models lead to very similar results, although the negative training impact on the employment-unemployment transition is no longer significant with the second specification.

We also provide additional results on the wage returns to training (participation in 1998). On the basis of a correction of selection on observables, we find a value of 7% for these returns. Quantile matching estimators suggest that the returns to training remain rather flat along the wage distribution. As they stand, the wage effects have to be interpreted as a upper bound since the matching estimates are likely to be biased because of unobserved heterogeneity. The difficulty here is that the data do not provide any convincing instrument. A correction for selectivity into training on unobservable characteristics could clearly lead to smaller wage effects of the participation in training programs.

Finally, several extensions of our empirical analysis could be considered. First, it would be worthwhile to consider the impact of training on a longer period of time. For instance, the marginal protective effect of training may be more important among older than younger workers. Secondly, the impact of training on labor mobility may be affected by both the type and the total duration of the training program. Finally, the characteristics of the firms and the way employers select their employees who benefit from the training program are also expected to influence job-to-job and employment-unemployment transitions. The use of longitudinal matched employer-employee data could shed light on these issues which are left for future work.

Table 1.6: Description of the sample

	Restricted sample with training in 1998			Extended sample with training from 1998 to 2000		
	No T.	T.	All	No T.	T.	All
Outcomes						
Mean wage in 2000 (log)	6.997	7.321	7.046	6.969	7.261	7.058
Percentile						
10th	6.462	6.899	6.531	6.413	6.819	6.574
25th	6.819	7.042	6.819	6.785	7.001	6.835
50th	7.001	7.279	7.042	6.973	7.224	7.041
75th	7.246	7.553	7.329	7.224	7.512	7.329
90th	7.553	7.860	7.607	7.512	7.799	7.629
Mobility						
From job to job	0.077	0.037	0.071	0.077	0.061	0.072
From empl. to unempl.	0.037	0.011	0.033	-	-	-
Explanatory variables						
Gender						
Male	0.565	0.593	0.569	0.568	0.585	0.573
Female	0.435	0.407	0.431	0.432	0.415	0.427
In couple						
No	0.337	0.357	0.340	0.328	0.344	0.333
Yes	0.663	0.643	0.660	0.672	0.656	0.667
Age						
18-29	0.144	0.135	0.143	0.134	0.142	0.136
30-39	0.341	0.369	0.345	0.334	0.372	0.346
40-49	0.332	0.367	0.337	0.337	0.351	0.342
> 50	0.183	0.131	0.175	0.195	0.135	0.176
Job seniority						
5 years	0.369	0.276	0.355	0.354	0.296	0.336
6-10 years	0.219	0.232	0.222	0.218	0.240	0.225
11-20 years	0.235	0.268	0.239	0.245	0.254	0.248
> 20 years	0.177	0.224	0.184	0.183	0.210	0.191
Education						
No diploma - BEPC	0.386	0.185	0.356	0.410	0.227	0.354
CAP-BEP	0.376	0.327	0.369	0.385	0.341	0.372
Baccalaureate	0.106	0.167	0.115	0.092	0.165	0.114
Undergraduate	0.086	0.193	0.102	0.072	0.166	0.102
Graduate	0.046	0.128	0.058	0.040	0.101	0.058
Nationality						
French	0.935	0.982	0.942	0.931	0.974	0.944
Others	0.065	0.018	0.058	0.069	0.026	0.056
Contract						
Permanent	0.950	0.970	0.953	0.967	0.967	0.967
Temporary	0.050	0.030	0.047	0.033	0.033	0.033
Type of job						
Full time	0.841	0.904	0.851	0.838	0.891	0.855
Part-time	0.159	0.096	0.149	0.161	0.109	0.145
Occupation						
Executives	0.073	0.188	0.090	0.063	0.164	0.093
Intermediary	0.163	0.327	0.188	0.149	0.280	0.189
Employees	0.280	0.236	0.273	0.274	0.265	0.272
Workers	0.484	0.249	0.449	0.514	0.291	0.446
Firm size						
0-19 employees	0.362	0.142	0.328	0.389	0.183	0.326
20-99 employees	0.210	0.172	0.205	0.207	0.203	0.206
100-499 employees	0.193	0.218	0.197	0.186	0.224	0.197
> 500 employees	0.235	0.468	0.270	0.215	0.390	0.271
Number of obs.	4345	762	5107	3308	1452	4761

Source: French Training Survey "Formation Continue 2000", authors' calculations.

Chapter 2

Stricter employment protection and firms' incentives to train: The case of French older workers[1]

From French data, this paper uses a difference-in-differences approach combined with propensity score matching to identify the effect of an exogenous change in employment protection among older workers on firm's incentives to provide training. Laying off workers aged 50 and above, French firms have to pay a tax to the unemployment insurance system, known as the Delalande tax. In 1999, the measure was subjected to a reform that increased due taxes but that did not concern equally all firms. We find that this exogenous shock to employment protection for older workers substantially rises firms' incentives to train the 45-49 age group of workers. This result confirms predictions of the simple labor market model we develop in a first stage.

JEL Classification: J14, J24, J26

Keywords: Older workers, employment protection, firms' training incentives.

[1]This chapter reviews a joint work with Pierre-Jean Messe. I am grateful to Francois Langot for thoughtful comments and suggestions. I also thank Francois-Charles Wolff and Arnaud Chéron. I have received helpful comments on earlier drafts of this article from participants at the CNRS thematic school "Public Policies Evaluation" 2011, JMA Meeting 2011, Journées LAGV 2011, EALE Conference 2011 and seminar participants at EPEE (Evry) and GAINS (Le Mans), with a special mention to Jean-Olivier Hairault.

2.1 Introduction

It is widely known that employment rates in OECD countries considerably differ in age due to large age-differences in labor market flows[2]. In particular, employment rates of older workers are low before the retirement age. This is a common characteristic to OECD countries, whatever the retirement age in force. Faced with these low employment rates of older workers, some countries have experimented with specific older worker employment protection in the form of higher firing taxes and subsidies on hiring (see OECD [2006]). In Belgium, Finland, France, Japan, Korea and Norway, it is indeed more costly for firms to lay off older workers because of longer notice periods or higher severance pay[3]. Specific older worker employment protection should foster long-term relationships between older workers and employers[4]. For instance, Schnalzenberger and Winter-Ebmer (2009) show that an age-specific firing tax caused a substantial reduction in layoffs for older workers in Austria. A similar regulation for France has been analyzed by Behaghel, Crépon and Sédillot (2008). The authors show that the most stringent schedule of this tax following the 1999 reform (change that is under consideration in this paper) led to decrease sizeably layoffs of older workers in large firms.

In this paper, we examine a case of specific older worker employment protection and look at its effect on firms' training incentives, which obviously raises specific age issues. Chéron, Hairault and Langot (2011) argue that the shorter distance to retirement (known as the "horizon effect") is the key point for understanding the economics of older workers employment. This view is supported by empirical evidence on micro-data (see Hairault, Langot and Sopraseuth (2010)). Our work absolutely fits with these concerns. It is widely known indeed that training incidence is a function of age (Bassanini et al. (2007)). This is most related to a distance to retirement issue

[2]The hump-shaped age-dynamic of employment in OECD countries reflects the age-dynamic of labor market flows, characterized by U-shaped inflow rates to unemployment (firing rates) and age-decreasing hiring rates.

[3]To compensate for age discrimination, governments in most European countries have also put specific inactivity and disability programs in place that provide generous substitution incomes until retirement. Finally, some countries have experimented with specific subsidies to increase the likelihood for older workers to find a job (UK, USA).

[4]The general conclusion reached in the large literature on employment protection legislation is that employment protection measures do not have a significant impact on steady-state employment, but are likely to influence the dynamics of employment (see Young (2003) for a review). More precisely, with fewer job terminations and less job creation, EPL is known to reduce inflows into unemployment and outflows from employment, while also lowering outflows from unemployment and inflows into employment. However, this indirect negative effect of employment protection on the overall employment rate does not concern older workers as their hiring rate is very low. Therefore, only the direct effect on firing rates matters regarding older workers.

since the present value of net returns to human capital investments in older employees is lower due to the shorter period during which both employees and employers can reap the benefits of these investments. From a natural experiment in The Netherlands, Montizaan, Cörvers and De Grip (2010) show that a decrease in pension rights postpones expected retirement and then increases participation in training courses among older employees (although exclusively for those employed in large organizations). From an equilibrium search model supported by an estimation based on French data, Khaskhoussi and Langot (2008) show that a short distance to the retirement age explains the low investment in training of elderly.

In this paper, using individual data, we try to properly identify the effect of stricter employment protection among older workers on firms' incentives to engage in firm specific-skills. Specifically, we study the impact of the 1999 French Delalande tax change. Since its introduction in 1987[5], French firms have to pay a tax to the unemployment insurance system laying off workers aged 50 and above, known as the Delalande tax[6]. The amount of the tax is proportional to the worker's gross wage at the time of layoff. Since 1992, firms are exempted from the tax for workers hired after the age of 50 if they are laid off later on. It is only due if the worker is employed under a permanent contract and only the private sector is concerned. The 1999 change resulted in an increase in the tax schedule for firms with more than 50 workers. This rise in the tax was implemented in a context of rapidly growing employment that benefited all categories of workers, except older unemployed workers. Table 2.1 shows how the amount of the tax has varied after the reform.

Table 2.1: Delalande tax schedule according to the age of the laid off worker (monthly gross wage)

		Worker's age								
		50	51	52	53	54	55	56-57	58	59
Jan. 1993-Dec. 1998	All firm sizes	1	1	2	2	4	5	6	6	6
Since Jan. 1999	More than 50 employees	2	3	5	6	8	10	12	10	8
	Less than 50 employees	1	1	2	2	4	5	6	6	6

Source: Behaghel, Crépon and Sédillot (2008), legislative texts.
Notes: For each age group, the table displays the tax due by the firm to the unemployment insurance system if it lays a worker off. The tax is a function of previous wages, and is stated in months of growth wages.

[5]Since January 2008, the Delalande tax no longer exists.
[6]The threshold-age was 55 in 1987 but was lowered to 50 after the 1992 reform.

As the Delalande tax increases the firing cost of workers aged 50 and above, it comes to an age-increasing firing tax. Therefore, we have to account for the fact that employment protection may have an age-differentiated impact. In particular, Chéron et al. (2007) study the effect of introducing an additional tax when laying off older workers (near retirement age) by extending the theory of job creation and job destruction to account for a finite working life-time[7]. As far as our paper is concerned, two conclusions can be drawn. First, the authors show that the introduction of the tax reduces firings of workers concerned by the tax while, on the contrary, the firing probability of workers who are below from the threshold-age of the tax increases. Indeed, the value of job continuation in the latter case is reduced because of the expected firing tax due in the last period of working life (in case of layoff), while retirement allows firms to avoid it. Anticipating the tax, firms increase layoffs before being subject to the tax while it is in their best interest to keep older workers on working in the last period of working life. Second, the authors provide a quantitative analysis of the "Delalande Tax". Higher firing costs after 55[8] are found to lead to better employment protection for the 55-59 years-old (who benefit from an increase in the employment rate of 4.2 points) but to negatively affect the employment rate for workers aged 45-54 (decreases by 2.7 points). These findings are consistent with the study of Behaghel et al. (2008). A higher level of the tax indeed deterred firms from laying older workers off.

The 1999 Delalande tax change led to increase firing taxes differences in age terms by tightening employment protection of workers employed in large firms and aged 50 and above. To our knowledge, no paper has already studied the effect of an employment protection specific to older workers on firms' incentives to engage in training. In light of these theoretical considerations, this seems to be an important concern. The approach and the key results we obtain can be summarized as follows. First, we develop a simple model of the labor market with both endogenous firing and training decisions, and where dismissals in the last period of working life (above a threshold-age) are subject to a firing tax. The impact of this tax on firm's incentives to engage in training particularly depends on the way it affects firing decisions in the last period as firms anticipate this tax. In particular, job destructions of older workers decrease while firings

[7]The equilibrium of such models is typically featured by increasing (decreasing) firing (hiring) rates with age, and a hump-shaped age-dynamics of employment.

[8]The calibration in Chéron et al. (2007) is based on the period before the 1992 tax reform. Therefore, the threshold-age above which firms are liable for the Delalande tax is 55 in their study while it is 50 in ours.

of younger workers rise. Accordingly, we show that the firing tax has no effect on older workers training but only may rise the training incidence of younger workers. The reason is that the age-specific firing cost plays on the future expected returns but older workers will be retired in the next period. Further, we show that firms' incentives to engage in training (for workers below the threshold-age) also depend on the arrival rate of idiosyncratic productivity shocks and on the initial job productivity: the lower the persistence of shocks, the higher firms' incentives. This comes from a complementarity effect between job destructions and training investments (see Chéron and Rouland (2011a)): training investments ensure a higher job value by increasing workers' productivity, which in turn reduces the risk of layoff that would become too costly at the last period of working life due to the tax. Conversely, in the event of persistent productivity shocks, this complementarity effect that leads to protect matches against future bad productivity shocks no longer matters. The effect of the firing tax on training only rests on a simple productivity effect that determines whether the tax affects the probability that the job will be robust to ageing. Therefore, age-specific firing costs may have no impact on training incidence, particularly considering high productive jobs for which the job is initially robust. Accordingly, the lower the initial productivity, the more likely the tax has an effect on firms' training incentives.

Empirical estimations on French data allow us to quantify these effects. In particular, we take advantage of the change in the Delalande tax schedule in 1999 to propose a reliable identification strategy based on the exogenous increase in the costs of laying older workers off. Indeed, we study employment protection reform in a case where the law explicitly treats workers differently depending on the firm size they work in. In particular, we use a difference-in-differences (DiD hereafter) approach combined with propensity score matching to compare older workers training rates in small and large firms, before and after the reform. By doing so, we are able to consistently estimate the average treatment effect on the treated, eliminating time-invariant biases between the treated sample and the comparison group sample due to mismatch related to firm size and differences in the measurement of the dependent variable. Once observable and unobservable factors are controlled for, we find a substantial effect of stricter employment protection on firms' incentives to train workers, but only significant for the 45-49 age group. In particular, the training rate of these workers is increased by 11.5 points of percentage in large relative to small firms after the reform. As expected, the effect is not significant for workers aged between 50 and

54. Finally, we show that the treatment effect appears to be greater among less productive jobs, suggesting that the evolution of technology is not so rapid.

The paper proceeds as follows. In section 5.2, we develop the theoretical model. Section 4.3 presents data and associated descriptive statistics. Section 2.4 presents the identification strategy and the results obtained through DiD specifications. Finally, section 2.5 concludes.

2.2 Qualitative analysis: A simplified theoretical model

2.2.1 Model environment

We study the theoretical implications of older workers employment protection on firms' incentives to engage in training in the following simplified environment. We consider a two-period, discrete time model in which older workers exit the labor market at the exogenous age T, perfectly known by employers. The last period of working life before retiring is denoted $T-1$ and the next to last period $T-2$. Apart from age $i \in [T-2, T-1]$, there is no other heterogeneity across workers. The economy is in steady-state and we do not allow for any aggregate uncertainty.

A productive unit is the association of one worker and one firm who are already matched[9]. The productivity of a worker is the sum of a random component ε and a deterministic one $y^i(k_i)$, derived from training investments in firm-specific skills during both periods k_i[10]. Lastly, at any age, workers earn an exogenous wage b.

The time of events and of decisions is as follows. First, at the beginning of the period, an idiosyncratic productivity shock may hit jobs at Poisson rate λ. If that is the case, a new job productivity ϵ is drawn in the general distribution $G(\epsilon)$ with $\epsilon \in [0, \overline{\varepsilon}]$, and the firm has then no choice but either to continue production or to terminate the job for a zero return and dismisses the worker. Dismissals of workers aged $T-2$ implies no specific cost while a firing cost F has to be paid when a firm fires a worker of age $T-1$. In this way, we account for a specific older workers

[9]As we only focus on training and firing decisions of firms, we do not account for the hiring process. Therefore and for the sake of simplicity, we consider associations already productive.

[10]The additive form of the output of the match we assumed between an endogenous component ($y^i(k_i)$) and another exogenous one (ε) clearly simplifies calculations but also fits the usual definition of training. Usually, training is considered as a way to improve workers' skills. Without training, workers are still able to produce but at lower productivity levels. To mention only a few, Lechthaler (2009) and Belot, Boone and van Ours (2007), within the framework of endogenous human capital and productivity shocks, consider an additive form of the output of the match as well.

employment protection, experimented in many OECD countries[11]. Firms decide to close down any jobs which productivity is below an (endogenous) productivity threshold denoted $R^i(k_i)$ that depends on the invested amounts in training. The job destruction rate is then determined by $\lambda G(R^i(k_i))$. Second, firms decide on the investment in firm-specific skills k_i that determines both the human capital of workers for the period $y^i(k_i)$ and their overall human capital level[12]. There is no human capital depreciation and skills are all allowed to accumulate between the two periods. Precisely, $y^{T-2} = y^{T-2}(k_{T-2})$ and $y^{T-1} = y^{T-1}(k_{T-2}, k_{T-1})$.

2.2.2 Firms' decisions

Firing decision For a firm, the intertemporal value of a filled job depends both on the worker's human capital $y^i(k_i)$ and on the idiosyncratic component ε. We denote this value by $J^i(k_i, \varepsilon)$. Further, there is no future expected profit in the value of $J^{T-1}(k_i, \epsilon)$ as the worker will retire in T. We assume that firms pay all the training cost $C(k_i)$[13]. Corresponding Bellman equations for $T-1$ and $T-2$ satisfy, respectively:

$$J^{T-1}(k_{T-2}, k_{T-1}, \varepsilon) = y^{T-1}(k_{T-2}, k_{T-1}) + \varepsilon - b - C(k_{T-1}) \tag{2.1}$$

$$\begin{aligned} J^{T-2}(k_{T-2}, \varepsilon) = & \underbrace{y^{T-2}(k_{T-2}) + \varepsilon - b - C(k_{T-2})}_{\text{instantaneous profit}} \\ & \underbrace{+ \beta\lambda \left[\int_{R^{T-1}(k_{T-2}, k_{T-1})}^{\overline{\varepsilon}} J^{T-1}(k_{T-2}, k_{T-1}, x) dG(x) - FG(R^{T-1}(k_{T-2}, k_{T-1})) \right]}_{\text{expected profit in } T-1 \text{ if new shock}} \\ & + \underbrace{\beta(1-\lambda) max\{J^{T-1}(k_{T-2}, k_{T-1}, \epsilon), -F\}}_{\text{expected profit in } T-1 \text{ if the shock lasts}} \end{aligned} \tag{2.2}$$

Integrating by parts $\int_{R^{T-1}(k_{T-2},k_{T-1})}^{\overline{\varepsilon}} J^{T-1}(k_{T-2}, k_{T-1}, x) dG(x)$ in equation (2.2) leads to $\int_{R^{T-1}(k_{T-2},k_{T-1})}^{\overline{\varepsilon}} (1- G(x))dx - F\left[1 - G(R^{T-1}(k_{T-2}, k_{T-1}))\right]$. Therefore, equation (2.2) comes to :

[11] We could also have formalized a specific older worker employment protection by considering firing costs due at any age but a higher tax for the last period of working life.

[12] The function $y^i(k_i)$ is supposed strictly increasing and concave, with $y(0) = 0$.

[13] with $C'(0) = 0$, $C'(k_i) > 0$ and $C''(k_i) = 0$.

$$
\begin{aligned}
J^{T-2}(k_{T-2},k_{T-1},\varepsilon) &= y^{T-2}(k_{T-2}) + \varepsilon - b - C(k_{T-2}) \\
&+ \beta\lambda \left[\int_{R^{T-1}(k_{T-2},k_{T-1})}^{\overline{\varepsilon}} (1-G(x))dx - F \right] \\
&+ \beta(1-\lambda)max\{J^{T-1}(k_{T-2},k_{T-1},\epsilon) + F, 0\} \qquad -\beta(1-\lambda)F
\end{aligned} \tag{2.3}
$$

The endogenous job destruction rule leads to a reservation productivity $R^{T-1}(k_{T-2}, k_{T-1})$ in $T-1$ defined by $J^{T-1}(k_{T-2}, k_{T-1}, R^{T-1}) = -F$ as a firing cost has to be paid at this period. On the other hand, the reservation productivity $R^{T-2}(k_{T-2}, k_{T-1})$ in $T-2$ is defined by $J^{T-2}(k_{T-2}, R^{T-2}) = 0$ as dismissals do not imply any cost. Therefore, the firing cost F has an opposite effect on $R^{T-1}(k_{T-2}, k_{T-1})$ and on $R^{T-2}(k_{T-2}, k_{T-1})$. It leads to reduce the job destruction flow in $T-1$ while it increases it in $T-2$. As argued by Chéron et al. (2011), the reason is that it is in the best interest of firms in $T-1$ to wait for the imminent retirement age that allows them to not to be subject to the tax. This is the labor-hoarding effect of firing costs. On the opposite, in $T-2$, firms anticipate the future firing tax and increase dismissals. This is the perverse anticipation effect of age-specific firing costs. And the higher the firing cost, the more extensive these effects. In addition, training investments improve job tenure increasing future productivity gains. In particular, the higher the option value of filled jobs (expected gains in the future) depending on the training investments, the weaker the job destructions.

$$
R^{T-1}(k_{T-2},k_{T-1}) = -y^{T-1}(k_{T-2},k_{T-1}) + b + C(k_{T-1}) - F \tag{2.4}
$$

$$
\begin{aligned}
R^{T-2}(k_{T-2},k_{T-1}) &= -y^{T-2}(k_{T-2}) + b + C(k_{T-2}) + \beta F \\
&- \beta\lambda \int_{R^{T-1}(k_{T-2},k_{T-1})}^{\overline{\varepsilon}} (1-G(x))dx \\
&- \beta(1-\lambda)max\{J^{T-1}(k_{T-2},k_{T-1},R_{T-2}) + F, 0\}
\end{aligned} \tag{2.5}
$$

Training investment decision Firms choose how many firm-specific skills they invest in, in order to maximize the net expected value of a filled job. It follows that the investment decision in $T-2$ is stated as:

$$\underset{k_i \geq 0}{max}\ J^i(k_i, \varepsilon) \quad \Longrightarrow \quad C'(k_i) = J_1(k_i, \varepsilon) \tag{2.6}$$

In this way, firms decide on the sum they invest in specific training so that the expected marginal return on investment is equaled to its marginal cost. The marginal return depends on present and future expected profits. In $T-1$, the expected value of a filled job only depends on the instantaneous profit as workers retire in T. But the expected profit in $T-1$ is also determined by the training investment carried out in $T-2$. Therefore, let denote $k_{T-1} = \kappa(k_{T-2})$[14] the optimal investment decision rule in $T-1$ conditionally to k_{T-2} that results from:

$$C'(k_{T-1}) = y_2^{T-1}(k_{T-2}, k_{T-1}) \tag{2.7}$$

If the job productivity drawn in $T-2$ lasts in $T-1$ (with probability $1-\lambda$), the job also does if $J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) \geq -F$, which occurs with probability $P(J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon)+F \geq 0)$. This is the probability that jobs will be robust to ageing since the job destruction rule changes in $T-1$ following the introduction of the firing tax[15]. By increasing productivity in $T-1$, the training investment in $T-2$ affects not only the expected job value in $T-1$, but also this probability. Therefore, $\underset{k_{T-2} \geq 0}{max}\ J^{T-2}(k_{T-2}, \kappa(k_{T-2}), \varepsilon)$ leads to $C'(k_{T-2}) = J_1(k_{T-2}, \kappa(k_{T-2}), \varepsilon) + J_2(k_{T-2}, \kappa(k_{T-2}), \varepsilon)$, that is to:

$$\begin{aligned} C'(k_{T-2}) &= y'_{T-2}(k_{T-2}) + \beta\lambda \frac{\partial \int_{R^{T-1}(k_{T-2}, \kappa(k_{T-2}))}^{\overline{\varepsilon}} (1-G(x))dx}{\partial k_{T-2}} \\ &+ \beta(1-\lambda) \left\{ \frac{\partial P(J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F \geq 0)}{\partial k_{T-2}} \left[J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F \right] \right\} \\ &+ \beta(1-\lambda) \left\{ \frac{\partial J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon)}{\partial k_{T-2}} \left[P(J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F \geq 0) \right] \right\} \end{aligned} \tag{2.8}$$

[14] With $\kappa'(k_{T-2}) \geq 0$.
[15] $J^{T-1}(k_{T-2}, \kappa(k_{T-2}), R^{T-1}) = -F$ instead of $J^{T-2}(k_{T-2}, \kappa(k_{T-2}), R^{T-2}) = 0$.

First, from equation (2.1), $\frac{\partial J^{T-1}(k_{T-2},\kappa(k_{T-2}),\epsilon)}{\partial k_{T-2}}$ is equivalent to $y_1^{T-1}(k_{T-2},\kappa(k_{T-2})) + \kappa'(k_{T-2})\left[y_2^{T-1}(k_{T-2},\kappa(k_{T-2})) - C'(k_{T-2})\right]$. Let denote A this expression. It is the net marginal return in $T-1$ on the training investment carried out in $T-2$ and is strictly positive. Second, from the Leibniz rule, differentiating $\int_{R^{T-1}(k_{T-2},\kappa(k_{T-2}))}^{\overline{\varepsilon}}(1-G(x))dx$ with respect to k_{T-2} leads to: $A\left[1 - G(R^{T-1}(k_{T-2},\kappa(k_{T-2})))\right]$. Accordingly, equation (2.8) can be rewritten as:

$$\begin{aligned} C'(k_{T-2}) &= y'_{T-2}(k_{T-2}) \\ &\quad + A\left\{\beta\lambda\left[1 - G(R_{T-1}(.))\right] + \beta(1-\lambda)\left[P(J^{T-1}(k_{T-2},\kappa(k_{T-2}),\epsilon) + F \geq 0)\right]\right\} \\ &\quad + \beta(1-\lambda)\left[J^{T-1}(k_{T-2},\kappa(k_{T-2}),\epsilon) + F\right]\left[\frac{\partial P(J^{T-1}(k_{T-2},\kappa(k_{T-2}),\epsilon) + F \geq 0)}{\partial k_{T-2}}\right] \end{aligned} \tag{2.9}$$

2.2.3 The effect of the age-specific firing cost on firms' training incentives

Considering the effect of F on k_{T-2}, we explore two different cases separately, according to the value of the rate λ. This helps us to identify mechanisms at work.

No persistence of shocks ($\lambda = 1$) Combining equations (2.9) and (2.4) implies:

$$\Phi(k_{T-2},F) \equiv -C'(k_{T-2}) + y'_{T-2}(k_{T-2}) + \beta A\left[1 - G(R^{T-1}(.))\right] = 0 \tag{2.10}$$

Differentiating equation (2.10) leads to $\frac{\partial k_{T-2}}{\partial F} = -\frac{\Phi_2(k_{T-2},F)}{\Phi_1(k_{T-2},F)}$ where $\Phi_1(k_{T-2},F)$ is necessarily negative to get an interior solution. Therefore, the impact of F on k_{T-2} only depends on the sign of $\Phi_2(k_{T-2},F)$:

$$\begin{aligned} \Phi_2(k_{T-2},F) &= \beta A\left[\frac{\partial\left[1 - G(R^{T-1}(.))\right]}{\partial F}\right] \\ &\equiv \beta A g(R^{T-1}(k_{T-2},\kappa(k_{T-2}))) > 0 \end{aligned} \tag{2.11}$$

The impact of F (due in $T-1$) on k_{T-2} rests on the way the tax affects the probability the job does not terminate after a productivity shock. When there is no persistence of shocks for sure, an increase in the firing cost clearly rises firms' incentives to engage in workers' training in $T-2$. This comes from a complementarity effect between training and firing decisions[16]. In particular, firms have strong incentives to protect matches against bad productivity shocks when a new shock will definitely hit the job and while dismissals in the next period are subject to the tax. The training investment increases the worker's productivity and therefore the (intertemporal) job value, which in turns reduces the risk of layoff that would become too costly at the next period because of the tax. It is worth noticing that firms' training incentives concerns both bad and high productive jobs (determined by the value of ϵ) as the value of the new shock in $T-1$ is completely independent of its value in $T-2$. Nevertheless, this unambiguous impact of F on training is not independent of the degree of persistence of the i.i.d. shock.

Possible persistence of shock ($0 < \lambda < 1$) Combining equations (2.9) and (2.4) implies:

$$\begin{aligned}\Phi(k_{T-2}, F) \equiv & -C'(k_{T-2}) + y'_{T-2}(k_{T-2}) \\ & + A\left\{\beta\lambda\left[1 - G(R_{T-1}(.))\right] + \beta(1-\lambda)\left[P(J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F \geq 0)\right]\right\} \\ & + \beta(1-\lambda)\left[J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F\right]\left[\frac{\partial P(J^{T-1}(k_{T-2}, \kappa(k_{T-2}), \epsilon) + F \geq 0)}{\partial k_{T-2}}\right] \\ & = 0 \end{aligned} \tag{2.12}$$

Again, the derivative of k_{T-2} has the same sign as $\Phi_2(k_{T-2}, F)$:

[16] Chéron and Rouland (2011a) show that job destructions and training investments are highly complementary since firms have strong incentives to invest in training to protect matches from idiosyncratic productivity shocks. Expected productivity gains due to training investments rise the job tenure, which in turn encourages firms to invest more.

$$
\begin{aligned}
\Phi_2(k_{T-2}, F) = \beta\lambda A g(R^{T-1}(.)) + \beta(1-\lambda)A\left[\frac{\partial P(J^{T-1}(.) + F \geq 0)}{\partial F}\right] \\
+ \beta(1-\lambda)\left[\frac{\partial P(J^{T-1}(.) + F \geq 0)}{\partial k_{T-2}}\right] \\
+ \beta(1-\lambda)\left[\frac{\partial^2 P(J^{T-1}(.) + F \geq 0)}{\partial k_{T-2}\partial F}\right]\left[J^{T-1}(.) + F\right]
\end{aligned} \tag{2.13}
$$

The last term on the right-hand side is equal to zero as the training investment in $T-2$ affects the probability $P(J_{T-1}(k,\varepsilon) + F \geq 0)$ through the accumulation of human capital, which does not depend on the tax level. Therefore, when shocks may be persistent between periods, the effect of the firing cost in $T-1$ on the training investment in $T-2$ not only depends on the positive effect of the tax on the probability the job does not terminate after a productivity shock (first term of the right-hand side), but also on the way the training investment itself and the firing cost affect the probability that the job will be robust to ageing (second and third term of the right-hand side). This second impact is not clearly stated. In particular, the tax may have no effect on this probability that the job will be robust to ageing considering high productive jobs for which the job is initially robust (the job value is higher than $-F$ from $T-2$). The lower the idiosyncratic productivity drawn in $T-2$, the more likely the tax has an effect on this probability. Overall, the lower the persistence of shocks (i.e. the higher λ), the higher firms' incentives to engage in training in $T-2$ to protect matches against bad productivity shocks. But, for low values of λ, the job productivity has to be low in order to encourage firms to engage in training. The impact of F on k_{T-2} actually depends on the value of λ and on the initial job productivity.

The 1999 Delalande tax change leads to increase firing taxes differences in age terms by rising the firing costs of workers aged 50 and above. Following the theoretical predictions we developed, training rates of workers affected by the reform (50 and above) are not likely to increase following the tax change. The tax change only may have an effect on the training incidence of workers below the threshold-age. But all depend on the evolution of technology and of the initial job productivity. In the event of rapid change of technology, training investments ensure a higher job value by increasing workers' productivity, which in turn reduces the risk of layoff that would become too costly after 50. Training investments help to protect matches against future bad productivity shocks. Conversely, in the event of persistent idiosyncratic productivity, the lower the initial productivity, the higher firms' training incentives. Next sections empirically quantify these effects.

2.3 Data and descriptive statistics

2.3.1 Data description

To assess the effect of the Delalande tax reform on firm-provided training, we use in this study two complementary French databases. The first one is a cross-sectional survey entitled "Formation Continue" conducted by INSEE (the French National Institute for Statistics and Economic Studies) in March 2000. It was carried out on a sample of 28667 individuals. The main interest of this survey is that it includes detailed information on training. In particular, it provides accurate information on the different training periods followed by each respondent along the life cycle. From the questionnaire, it is possible to separate three main periods: i) from exit to school to February 1998, ii) from March 1998 to December 1998 (training in 1998 hereafter), and iii) from January 1999 to March 2000 (training in 1999 hereafter). The "Formation Continue" survey gives information about the financing organization. We decide to only consider firm-financed training as the Delalande tax reform affects firms. This was the case for about 81% (80%) of the older workers having participated in a training program in 1998 (1999). The survey also describes the type of experienced training using four categories: i) training in a work situation, ii) apprenticeships, iii) work placement or training courses and iv) self-training. Furthermore, we have information about the purpose of the training activity, which has to fit in one of the following categories: i) to adapt to the job, ii) to switch to another job or to get a job, iii) to obtain a diploma or a certification, iv) to execute political duties, v) personal or cultural reasons, and vi) no specific reason. When turning to the data, we choose to only consider the first motive since it was the result of 75% (82%) of the firm-financed training sessions received by employees in 1998 (1999).

We merge these data with the 1998 and 1999 waves of the French Labor Force Survey ("Enquête Emploi"). This is a rotating panel since exactly one-third of the sample is dropped from the sample each year and is replaced with a new, comparable sample drawn from the current population. The size of the Enquête Emploi is about 135000 individuals who are yearly interviewed about their situation on the labor market. The main feature of these data is that they provide detailed information over two years for two-third of the 1998 original sample. This means that we can investigate firms training decisions before and after the reform. Finally, for each respondent, the Enquête Emploi contains detailed information about socio-demographic individual characteristics, as well as job and firm characteristics.

2.3.2 Sample selection and descriptive statistics

We define our sample in the following way. First, we exclude farmers and self-employed as well as individuals working in the public sector since layoffs are very infrequent in the public sector. We focus on the population of respondents aged from 45 to 54 in March 1998[17]. Furthermore, we restrict our analysis on men to control for the distance to retirement that determines the number of contributive years from graduation[18]. As the distance to retirement is expected to have a strong impact on firm training decisions, we consider only men in our sample. Dropping the few missing values (mainly because of missing firm sizes) and selecting only workers who were working both in 1998 and in 1999[19], we have a total set of 1000 observations for each year. We consider the participation of these individuals to a firm-financed training spell while employed between March 1998 and December 1998 for the pre-reform period and between January 1999 and March 2000 for the post-reform period.

Some descriptive statistics about the sample we use are provided in table 2.7 in appendix.

[17]Workers aged 55 are so close to the retirement age that we expect the reform will have no effect on their access to firm-financed training ("horizon" effect).

[18]Following Hairault et al. (2010), distance to retirement is captured by the difference between the current age and the retirement age. Considering the French pension system, the retirement age can be approximated by the required number of contributive years to get the full pension rate: the full pension age which is exogenous to the labor market status. The distance to retirement for an individual is then equal to the full pension age minus her current age. However, if a person enters the job market at a very young age, she cannot retire before the eligibility age for full pension (60 years old) even though she has accumulated the required number of contributive quarters before this age. In this case, the retirement age is then set at 60 and the distance to retirement is 60 minus the current age. While unemployment episodes in the French system are included in the number of contributive periods, this proxy for the retirement age does not take into account non-continuous careers due to maternity leaves and family commitments. The retirement age is then only relevant for male, which implies to consider only male workers in our sample selection.

[19]This selection allows us to use a balanced panel for estimations. However, less than 1% of individuals employed in 1998 were fired in 1999.

First, focusing either on large or small firms, there are very few differences between the pre- and the post-reform period, except for the distance to retirement. Beyond these differences, comparing individuals employed in firms with 50 workers or more and those employed in firms with less than 50 workers is of particular interest to us. Not surprisingly, the probability that the firm has a training plan is overwhelming in large related to small firms (depending on the year, about 73% against 27%)[20]. Similarly, we observe that job seniority tends to be much higher for individuals working in large firms than for those working in small firms. Furthermore, the sectoral composition strongly depends on firms' size as well. For instance, the building sector represents barely 6% of all the jobs in large firms while it is about 20% in small ones. Large firms are also characterized by the predominance of the tertiary industry while the manufacturing sector is the greatest in small firms. Lastly, regarding the wage distribution, workers in large firms have on average better paid jobs than workers in small ones (about 1800 euros a month against 1500 euros)[21].

As these variables may strongly matter in explaining the access to firm-financed training, the observed deviations make both groups not comparable. We have then to estimate the causal effect of the change in the Delalande tax schedule on workers' training rate by accounting for the differences in the distribution of covariates between both groups. Following Heckman, Ichimura and Todd (1998) and Blundell and Costa Dias (2000, 2004), we exploit the panel aspect of the data using a conditional DiD approach.

2.4 Quantitative analysis: Identification strategy and results

The goal of the paper is to measure the impact of stricter employment protection among older workers on firms' incentives to provide training. We exploit a discontinuity in the Delalande tax reform: in 1999, the legislation led to an increase in the tax for firms with 50 workers or more while the tax remained unchanged for firms with fewer than 50 employees. The treatment is a

[20]This result is in line with Bassanini et al. (2007) and Montizan et al. (2010) for instance, who show that training incidence is much higher among larger organizations.

[21]Table 2.8 in appendix gives the difference in means between both groups for each observable characteristic. This confirms results in table 2.7. Thus, the difference in the probability of having a company training plan between large and small firms exceeds 0.46 for each year and for both samples for instance. Besides, jobs in building or tertiary sector are over-represented in small organizations, while jobs in the industrial sector represent around one third of all the jobs in large firms.

unexpected one-time change in government policy and applied almost equally to all members of the treatment group[22]. The one-time nature of the change makes it easy to select specific pre- and post-treatment points in time. Consequently, we choose to use a DiD approach combined with propensity score matching for our evaluation[23].

The basic intuition of the DiD approach is to study the impact of some "treatment" on the recipients, comparing the difference in average performance of the eligible group pre- and post-treatment relative to the performance of some control group pre- and post-treatment. More formally, let $P_{i,t}$ be a dummy variable equal to 1 if worker i has participated to a firm-financed training session at time t, with $t \in \{1998; 1999\}$. Treatment and control group are identified by the dummy variable T_i, such that $T_i = 1$ if the worker i is employed in a large firm (i.e. a firm with more than 50 employees)[24]. A set of covariates $X_{i,t}$ assumed to affect significantly the access rate to firm-provided training is also included. In this way, we include common training determinants such as gender, marital status, occupation, education, nationality, job seniority, existence of a training plan in the firm. Finally, following Hairault et al. (2010), we also include the distance to retirement in the set of regressors.

We aim at estimating the following linear probability model:

$$E(P_{i,t} = 1) = \beta X_{i,t} + U_{it} \quad \text{if } t = 1998 \tag{2.14}$$

$$E(P_{i,t} = 1) = \beta X_{i,t} + \alpha_i T_i + U_{it} \quad \text{if } t = 1999 \tag{2.15}$$

where $U_{i,t}$ is the error term assumed to be normally distributed with a mean 0. In equation (2.15), α_i measures the effect of the change in the tax schedule on the access rate to firm-provided training of each individual i. As shown by Blundell and Costa Dias (2000), the individual-specific component of the treatment effect may differ between the treatment and the control group of individuals, making the identification of the average effect of the treatment more difficult. In this

[22]As reported in table 2.1, even though the due tax doubles in most ages, the tax reform is not strictly equally applied to all workers according to their age since it also trebles in cases. The rise is also less significant in oldest ages. However, given the sizeable tax reform, the effect is never insignificant, so that evaluating its impact on the whole group is not a problem.

[23]Of course, DiD method and natural experiments are not the only way to evaluate the effect of a treatment. See Blundell and Costa Dias (2000) for a review of non-experimental methods for the evaluation of social programmes.

[24]Workers employed in large firms but hired after 50 years old are included in the control group, together with workers employed in small firms.

setting, a DiD approach allows us to recover the average effect of the treatment on the treated individuals (ATT effect hereafter) under certain conditions. The DiD estimator can be stated as:

$$\begin{aligned}\alpha_{DiD} &= [E(P_{i,99} = 1|T = 1) - E(P_{i,99} = 1|T = 0)] \\ &- [E(P_{i,98} = 1|T = 1) - E(P_{i,98} = 1|T = 0)]\end{aligned} \tag{2.16}$$

In addition, we consider the following decomposition of the error term $U_{i,t}$:

$$U_{i,t} = \phi_i + \theta_t + \mu_{i,t}$$

where ϕ_i stands for an individual-specific effect constant over time. θ_t represents a common time effect (or common macro effect) and $\mu_{i,t}$ is a temporary individual specific effect. Substituting equations (2.14) and (2.15) into (2.16), the DiD estimator can be expressed in the following way:

$$\begin{aligned}\alpha_{DiD} &= E(\alpha_i|T_i = 1) \\ &+ \beta[E(X_{i,99}|T_i = 1) - E(X_{i,99}|T_i = 0) + E(X_{i,98}|T_i = 0) - E(X_{i,98}|T_i = 1)] \\ &+ [E(\mu_{i,99}|T_i = 1) - E(\mu_{i,98}|T_i = 1) + E(\mu_{i,98}|T_i = 0) - E(\mu_{i,99}|T_i = 0)]\end{aligned} \tag{2.17}$$

The first term in the right-hand side represents the ATT effect. The second term stands for the difference in means of covariates across groups (i.e. treatment and control groups) for each year. The last term indicates the unobserved temporary individual-specific component of the error term. It is worth noting that the DiD estimator allows to remove unobservable individual-specific effects constant over time and common time effects. However, the second and the third term must equal to 0 in order the DiD estimator to provide a consistent estimate of the ATT effect. Thus, the DiD estimator is based on the identifying assumption that, in absence of the treatment, the average outcome for the treated would have experienced the same variation as the average outcome for the untreated (such that without treatment $\alpha_{DiD} = 0$): this is the "time invariance" assumption. Formally, this identifying assumption writes:

$$E(P_{i,99} = 0|T = 1, X_{1,99}) - E(P_{i,98} = 0|T = 1, X_{1,98}) \quad (2.18)$$

$$= E(P_{i,99} = 0|T = 0, X_{0,99}) - E(P_{i,98} = 0|T = 0, X_{0,98}) \quad (2.19)$$

In the next part, we estimate the ATT effect of the change in the Delalande tax schedule on the access rate to firm-provided training, assuming time invariance through a DiD approach. Then, we relax this assumption by re-estimating the ATT effect through a DiD regression combined with a propensity score matching procedure. Finally, we try to identify workers' characteristics for which the ATT is significant and positive.

2.4.1 A DiD approach

We first aim at checking whether the tax change had an effect on firms' incentives to provide training. As in Kugler and Pica (2008), we then estimate the following linear probability model to control for the possibility that higher training rates are the result of changing characteristics of workers:

$$E(P_{i,t} = 1|X_{i,t}, T_i, \tau_t) = \beta X_{i,t} + \delta_1 T_i + \delta_2 \tau_t + \alpha_{DiD}(T_i * \tau_t) + c_i + U_{i,t} \quad (2.20)$$

where c_i is an individual effect and τ_t a dummy variable that takes the value of 1 from 1999 (*ie* after the reform) and zero otherwise. The interaction term between the large firm dummy and the post-reform dummy captures the effect of interest. α_{DiD} then identifies the causal effect of treatment under the identifying assumption (2.18) resulting in $E[u_{i,t}|P_{i,t}] = 0$.

Table 2.2 reports marginal effects of the linear probability model using equation (2.20). Column (1) reports results for the whole sample of workers aged between 45 and 54 while column (2) gives results for workers aged between 45 and 49. First, they show a large and statistically significant raise in training access in large relative to small firms after the reform was introduced. Thus, increasing employment protection of older workers through the tax reform leads to rise by 8.3 (12.7) percentage points the training rate of workers aged 45-54 (45-49) in large firms. The treatment effect is stronger for the 45-49 age group than for the 45-54, which just goes to prove that 50 is a fateful threshold that determines firms' firing and training decisions. Besides, the training access rate only significantly rises for workers aged between 45 and 49. This result confirms theoretical predictions of section 5.2.

The positive and significant ATT effect for workers aged between 45 and 49 might result from the implementation of other reforms than the change in the Delalande tax schedule. For instance, in 1999, the French government introduced the 35-hour workweek regulation but all French firms did not sign an agreement on working time reduction at the same time. As shown in Aeberhardt et al. (2011), the signing date of such a regulation strongly depends on firms size. In particular, the signing date was earlier for large firms than for small firms. Consequently, this could have a differentiated impact on firms' training incentives between treated and control groups of observations. We suggest a simple test to check whether the effect of working time reduction on training rates would differ across groups. In the null hypothesis, the effect would be the same between both groups and would be removed by the DiD approach. Therefore, the DiD estimator would be significant only for workers aged 45-49, in line with our theoretical results. In the alternative hypothesis, the effect of the switch to the 35-hour workweek would differ across groups and the DiD estimator should be significant for all cohorts of workers. To perform this test, we estimate the same linear probability model as in equation (2.20) without any selection on age, including workers aged between 30 and 54. Results are presented in table 2.3. Only interaction terms α_{DiD} are reported. This table shows that the training access rate only significantly rises for workers aged between 45 and 49. Therefore, the null hypothesis cannot be rejected. The effect of the switch to the 35-hour workweek is similar between groups and is removed by the DiD approach.

Finally, table 2.2 also gives information on the determinants of workers' access to firm-financed training sessions. Results show that there are neither nationality nor sectoral differences in participation. There is also no significant differences between job seniority. But there is a positive effect of occupation on the probability of having been trained. Thus, training is much more frequent among executives and intermediaries. As expected, the existence of a training plan in the firm raises significantly enrolments in training. The higher the wage, the higher the training incidence as well. The last finding is the positive effect of the distance to retirement. That is for sure all the more profitable for firms to train older workers that they are likely to remain employed for a long time[25]

2.4.2 A DiD matching strategy

Time invariance assumption implies that the average training propensities for workers employed in large firms would have experienced the same variation as the ones for workers employed in small ones, had they worked in small firms as well. To be plausible, this assumption then requires that being employed in a large firm is similar to working in a small one. However, given differences between both groups highlighted in tables 2.7 and 2.8, one may not be confident with the time invariance assumption. Therefore, to account for the differences in the distribution of covariates between the treated and the control group, we implement a conditional difference-in-differences estimator (CDiD), as suggested by Heckman, Ichimura and Todd (1998) and Blundell et al. (2000, 2004). This method combines a propensity score matching approach with DiD such that, at each period, a counterfactual outcome for workers employed in large firms if they were working in a small one is estimated semiparametrically. This technique enables us to relax, relative to standard DiD, the linear assumption when controlling for observables and to control for unobservables exploiting the panel dimension of the data. The matching procedure makes the distribution of covariates across groups comparable by building a suitable sample control group. Besides, Smith and Todd (2005) show that the DiD matching estimator performs the best among nonexperimental matching based estimators.

As it is the only age-group for which the Delalande tax change had an effect, we focus in this section only on workers aged between 45 and 49. We first build a correct sample counterpart for the missing information on the treated outcomes, had they not been treated. This leads to re-establish the conditions of an experiment with a total random assignment into treatment, by matching each treated observation with a similar individual of the control group on the basis of some observable variables. In a second step, we estimate the ATT effect using a DiD regression and weighting non treated observations according to their closeness to the treated ones in terms of a set of covariates X.

[25] All in all, these results are in line with Chéron, Rouland and Wolff (2008) who estimate the impact of firm-training on mobility and wages in France.

Table 2.2: Results from a DiD estimation

	(1)		(2)	
Variables	Coefficient	Std. Err.	Coefficient	Std. Err.
Large firms	0.042	(0.029)	0.036	(0.041)
Post-reform	0.046	(0.024)	0.016	(0.037)
Large firms * Post-reform	0.083***	(0.029)	0.127***	(0.045)
Intercept	-0.149**	(0.06)	0.004	(0.146)
In couple	-0.016	(0.031)	-0.05	(0.05)
No French citizenship	-0.071	(0.054)	-0.007	(0.085)
Education (ref: no diploma)				
CAP-BEP	-0.004	(0.023)	0.032	(0.035)
Baccalaureate	-0.003	(0.04)	0.052	(0.058)
College degree	-0.038	(0.042)	0.073	(0.068)
Distance to retirement	0.01***	(0.003)	-0.005	(0.085)
Job seniority (ref: ≤ 5)				
6-10 years	0.004	(0.033)	0.044	(0.048)
11-20 years	-0.02	(0.031)	0.011	(0.044)
More than 20 years	0.008	(0.029)	0.031	(0.041)
Occupation (ref: workers)				
Executives	0.099**	(0.033)	0.11	(0.042)
Intermediary	0.092***	(0.027)	0.14***	(0.039)
Employees	0.053	(0.04)	0.101	(0.057)
Existence of a training plan	0.128***	(0.024)	0.129***	(0.035)
Part-time job	-0.016	(0.059)	-0.048	(0.098)
Sector (ref: Building)				
Industry	0.016	(0.035)	0.005	(0.053)
Services	0.054	(0.035)	0.061	(0.054)
Wages quartiles (ref: 1st quartile)				
2nd quartile	0.056**	(0.026)	0.064*	(0.037)
3rd quartile	0.087***	(0.029)	0.102**	(0.042)
4th quartile	0.092**	(0.036)	0.036	(0.054)
Number of observations	2000		982	
R^2	0.13		0.14	
Pearson's Coefficient	0.322		0.326	

Significance levels: *: 10% **: 5% ***: 1%

Source: French Training Survey "Formation Continue 2000" and Labor Force Survey (waves 1998 & 1999)

Lecture: Column (1) gives results for the whole sample while column (2) gives results for the restricted sample on age, including only workers between 45 and 49. The interaction term between "large firms" and "post-reform" measures the DiD.

Table 2.3: Results from a DiD estimation, without any selection on age

	Coefficient	Std. Err.	Number of obs.
Age group			
30-34 years old	0.021	(0.044)	1060
35-39 years old	0.043	(0.041)	1158
40-44 years old	0.02	(0.044)	1145
45-49 years old	0.127∗∗∗	(0.043)	982
50-54 years old	0.05	(0.042)	945

Significance levels: ∗∗∗ : 1%

Note: Only interaction terms α_{DiD} of DiD are reported.

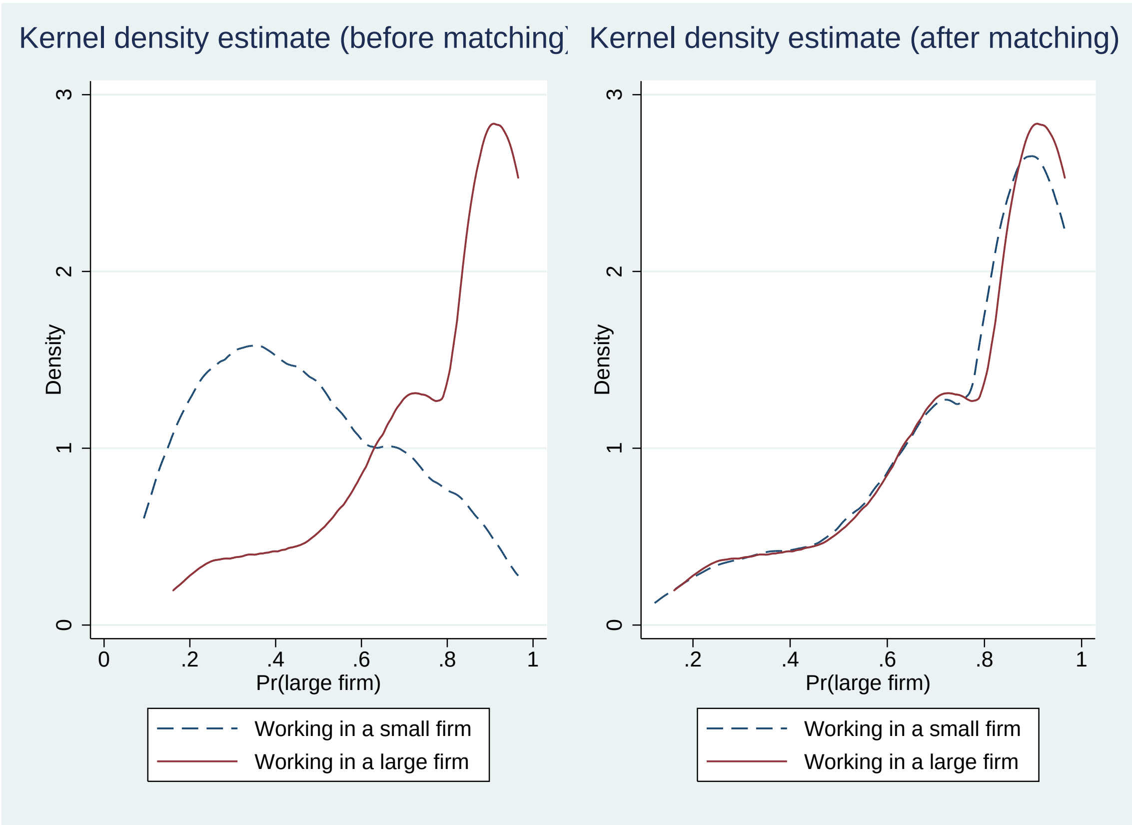
As before, T_i is the dummy variable equal to one if the agent i is employed in a firm with more than 50 workers. Following Rosenbaum and Rubin (1983), matching is usually carried out on the propensity to participate as a function of observable characteristics X: $e(X) = P(T_i = 1|X_i)$, which is the propensity score. The usual assumption required to estimate what would be the average probability of being trained of workers employed in large firms if they were working in a small one is the conditional independence assumption:

$$E(P_{i99} = 1|T_i = 0) - E(P_{i98} = 1|T_i = 0) \perp T_i|e(X_i) \tag{2.21}$$

We use a probit model to estimate the propensity score, that is the probability of working in a large firm depending on observable covariates. These ones should ideally include all important variables influencing this probability. The propensity score matching proved to be successful since the goodness of fit of the probits is high: on average, they correctly predict the treatment status in approximately 78% of the cases. Results of the probit estimates are reported in Appendix (Table 2.9). Not surprisingly, some observables such as job seniority or the presence of a training plan in the firm strongly affect the probability of working in a firm with 50 workers or more. Workers employed in the industrial sector are also more likely to work in a large firm.

Propensity score matching can be successful concerning the conditioning on observable characteristics only if the estimated propensity scores of workers employed in large and small firms overlap sufficiently. We implemented a common support requirement which led to the discarding of sixteen cases that were outside the common support region. Finally, after matching, all observable characteristics should be balanced between workers employed in a large firm and matched comparison observations. This is illustrated below in Figure 2.1, which reports the kernel density estimates of the propensity scores for workers employed in large firm and those employed in small ones. The matching procedure allows to make the distribution of covariates across groups comparable.

Figure 2.1: Common support of the propensity scores



In a second step, we estimate the ATT effect by using a DiD regression and implementing a weight function W_{ij} in the sample of workers employed in a small firm, relative to the predicted propensity score $e(X)$ of each individual i. We apply kernel matching estimators. The counterfactual outcome is then estimated on the basis of a weighted average of all workers employed in a small firm j. Denoting by α_{DiDM} the DiD matching estimator, we can write:

$$\begin{aligned}\alpha_{DiDM} = \Sigma_i[(E(P_{i,99} = 1|T_i = 1) - E(P_{i,98} = 1|T_i = 1))\\ - \Sigma_j W_{ij}(E(P_{j,99} = 1|T_i = 0) - E(P_{j,98} = 1|T_i = 0))]\end{aligned} \tag{2.22}$$

where W_{ij} is the weight placed on comparison observation j for individual i.

Once we make observations comparable between treated and control groups, we find that the change in the EPL led to an increase in the access rate to employer-provided training of treated individuals by 11.5 points of percentage[26], which is very similar to the effect we estimated using a simple DiD approach.

2.4.3 A different treatment effect across skill groups

Results from empirical sections have highlighted that costlier firing taxes for workers above 50 and employed in large firms rise firms' training incentives, but only the 45-49. According to the theoretical predictions from section 5.2, one may wonder whether these firing taxes had a different impact among the 45-49 depending on the skill level. Indeed, we have shown that, in the event of persistent productivity shocks, the effect of the firing tax on training rests on a simple productivity effect that determines whether the tax affects the probability that the job will be robust to ageing. Considering high productive jobs for which the job is initially robust, the tax will have no effect on training incidence. The lower the initial productivity, the more likely the tax has an effect on firms' training incentives. Accordingly, we can expect that the 1999 Delalande tax change only had an effect on the training incidence of less productive workers. Conversely, in the event of rapid evolution of technology, all workers maintained in employment are likely to face higher training rates following the tax change, whatever their initial skill level. This comes from a complementarity effect between job destructions and training investments. Therefore, we should not find any significant differences across skill groups.

One may assume that the less productive workers are also the less educated, those whose earnings are the lowest, or else workers who have the less important jobs. Therefore, we address these concerns by performing new estimations based on equation (2.20) and decomposing by occupation, wages quartiles and education[27]. Again, the interaction term between the large firm dummy and the post-reform dummy captures the effect of interest and identifies the causal effect of treatment. Results are presented in tables 2.4 to 2.6. They show that the treatment effect

[26] The corresponding standard error is (0.041), which means that the estimate is significant at the 5% level.

[27] For instance, we regress a first time equation (2.20) for executives, a second time for intermediaries, a third time for employees and a fourth time for workers. Estimations are then repeated for each education levels and each occupation dummies.

tends to be stronger for the less productive jobs, suggesting that the evolution of technology is not so rapid. High-productive workers do not become bad from one day to another. Therefore, training investments allow firms to increase productivity of less productive workers who have been kept on working so that their job will be robust to ageing. More productive workers do not need training since the expected value of their job plus the firing tax is strictly positive from the beginning. Precisely, the tax reform leads to rise significantly the training propensity of workers in blue-collar jobs and employed in large firms (+14.5 percentage points) while it did not affect significantly training incidences of others categories of workers. Workers whose earnings belong to the second quartile are the only one to have benefited from the tax reform (+8 percentage points)[28] and the 1999 tax change only had a significant impact on workers who have no diploma.

Table 2.4: Treatment effect on training participation in large firms: differences among skill groups (1)

Occupation	Executives	Intermediaries	Employees	Workers
Interaction term	0.085	0.146	0.108	0.145**
	(0.166)	(0.12)	(0.236)	(0.064)
Number of obs.	139	259	76	508

Significance level: ** : 5%
Note: Only interaction terms α_{DiD} of DiD are reported. Other control variables are: marital status, distance to retirement, job seniority, nationality, wages quartiles, education levels and sectoral dummies.

Table 2.5: Treatment effect on training participation in large firms: differences among skill groups (2)

Wages quartiles	1st	2nd	3rd	4th
Interaction term	-0.014	0.271***	0.08	0.155
	(0.077)	(0.102)	(0.127)	(0.136)
Number of obs.	258	251	256	217

Significance levels: * : 10% ** : 5%
Note: Only interaction terms α_{DiD} of DiD are reported. Other control variables are: marital status, distance to retirement, job seniority, nationality, education levels and sectoral and occupation dummies.

2.5 Conclusion

This paper investigates, both theoretically and empirically, the effect of stricter employment protection among older workers on firm's training incentives. First, we develop a simple model

[28] The lowest wages are quite dissociated from productivity because of minimum wage. This may explain why the effect of the tax change is not significant for workers among the first quartile of wages.

Table 2.6: Treatment effect on training participation in large firms: differences among skill groups (3)

Education	No diploma	CAP-BEP	Baccalaureate	College degree
Interaction term	0.152*	0.122	0.145	0.144
	(0.085)	(0.083)	(0.193)	(0.192)
Number of obs.	328	443	92	119

Significance level: * : 10%

Note: Only interaction terms α_{DiD} of DiD are reported. Other control variables are: marital status, distance to retirement, job seniority, nationality, wages quartiles and sectoral and occupation dummies.

with finite working life-time, endogenous job destruction and firm's training investment. We show that age-specific employment protection affects firms' incentives to engage in training only for the unprotected age group (below the threshold-age). This comes from a complementarity effect between training and job destruction. Since the expected separation cost is higher, firms have strong incentives to invest in training to protect matches against bad productivity shocks. However, we argue that the complementarity effect matters only if the job is likely to be hit by an idiosyncratic shock at the next period. Consequently, the effect of age-specific employment protection on firms' training incentives strongly depends on the persistence of shocks. If there is no persistence, the layoff tax unambiguously increases the training incidence of workers below the threshold-age. Conversely, in the event of persistent productivity shocks, this effect is no longer clearly stated and may depend on the initial productivity of the job. If the job is highly productive and therefore robust to ageing even without investment in training, stricter employment protection on older workers does not affect firms' incentives to engage in training.

We confront these theoretical predictions to French data, exploiting a change in the Delalande tax schedule in 1999 that concerns only firms employing 50 workers or more. We implement a conditional difference-in-difference estimator to remove selection bias into treatment on observables, individual specific effect constant over time and macro effects common to both groups. We find that the increase in the Delalande tax for large firms significantly raised the access rate to employer-provided training of treated workers aged 45 to 49 by 11.5 points of percentage. Further, a skill-decomposition of this effect shows that the 1999 reform only had a positive and significant effect on the training rate of less productive workers. According to our theoretical findings, this could result from the persistence of shocks.

As shown by Picchio and van Ours (2011), a better access to on-the-job training has an effect on the employability of workers, even for older workers. Therefore, the authors suggest to introduce age-specific subsidies or layoff taxes to stimulate job training and to retain the employability of older workers. Nevertheless, our results show that specific employment protection does not lead to increase firms' incentives to engage in older workers' training due to their shorter distance to retirement. Therefore, looking at the effect of age-specific training subsidies on firm-provided training to older workers could be an interesting issue for future work.

Further, beyond older workers, it is also an important concern for policy makers to worry about the employability of low-skilled workers. In this respect, we have shown that age-specific firing taxes led firms to direct their training effort on less productive workers just below the threshold-age of the tax. Alternatively, it could be worth comparing this positive effect with the impact of training subsidies on firm's incentives to train, decomposing by skill level of workers. We could expect that such subsidies would be used for workers who already have a strong labor market position, which would be of limited interest.

Table 2.7: Description of the sample, before and after the reform (in shares)

	Pre-reform		Post-reform	
	Large firms	Small firms	Large firms	Small firms
Training rate	0.208	0.083	0.329	0.119
Age				
45-49	0.565	0.501	0.455	0.413
50-54	0.435	0.499	0.545	0.587
Marital status				
In couple	0.894	0.878	0.906	0.864
Living alone	0.106	0.122	0.094	0.136
Nationality				
French	0.969	0.956	0.969	0.956
Others	0.031	0.044	0.031	0.044
Education				
No diploma	0.368	0.377	0.368	0.377
CAP-BEP	0.405	0.446	0.405	0.446
Baccalaureate	0.092	0.078	0.092	0.078
College degree	0.135	0.100	0.135	0.100
Distance to retirement				
5-10 years	0.316	0.377	0.404	0.479
10-15 years	0.524	0.51	0.565	0.493
More than 15 years	0.16	0.113	0.031	0.028
Job seniority				
Less than 5 years	0.114	0.388	0.100	0.360
6-10 years	0.111	0.172	0.114	0.191
11-20 years	0.202	0.186	0.189	0.197
More than 20 years	0.573	0.255	0.595	0.249
Existence of a company training plan				
Yes	0.736	0.271	0.736	0.271
No	0.264	0.729	0.264	0.729
Occupation				
Executives	0.175	0.133	0.180	0.144
Intermediary	0.274	0.258	0.279	0.238
Employees	0.078	0.064	0.078	0.058
Workers	0.473	0.546	0.463	0.560
Sector				
Industry	0.590	0.263	0.595	0.269
Building	0.059	0.197	0.061	0.197
Services	0.351	0.540	0.344	0.535
Type of job				
Full-time	0.978	0.967	0.978	0.970
Part-time	0.022	0.033	0.022	0.030
Monthly net wage	1786.9	1500.9	1807.9	1500.4
Number of observations	639	361	639	361

Source: French Labor Force Surveys (1998 & 1999) and French Training Survey (2000).

Table 2.8: Differences in means of covariates between the treated and the control group

	1998	1999
In couple	0.015	0.042**
	(0.021)	(0.020)
Distance to retirement		
5-10 years	-0.061*	-0.075**
	(0.031)	(0.033)
10-15 years	0.015	0.072**
	(0.033)	(0.033)
More than 15 years	0.046**	0.004
	(0.023)	(0.011)
Job seniority		
Less than 5 years	-0.274***	-0.260***
	(0.026)	(0.025)
6-10 years	-0.061***	-0.077***
	(0.022)	(0.023)
11-20 years	0.016	-0.007
	(0.026)	(0.026)
More than 20 years	0.318***	0.345***
	(0.031)	(0.031)
Training plan	0.464***	0.464***
	(0.029)	(0.029)
Occupation		
Executives	0.042*	0.036
	(0.024)	(0.025)
Intermediary	0.016	0.04
	(0.029)	(0.029)
Employees	0.015	0.020
	(0.017)	(0.017)
Workers	-0.073***	-0.096***
	(0.033)	(0.033)
Sector		
Industry	0.327***	0.326***
	(0.031)	(0.031)
Building	-0.137***	-0.136***
	(0.021)	(0.020)
Services	-0.190***	-0.190***
	(0.032)	(0.032)
Wages quartiles		
1st quartile	-0.174***	-0.162***
	(0.028)	(0.028)
2nd quartile	-0.015	-0.024
	(0.028)	(0.028)
3rd quartile	0.093***	0.092***
	(0.028)	(0.029)
4th quartile	0.096***	0.094***
	(0.028)	(0.028)

Significance levels: *: 10% **: 5% ***: 1%
Note: Only covariates for which the null hypothesis of equality of means at a 10% level can be rejected are reported.

Source : Labor Force Survey and Training Survey (waves 1998 and 1999)

Table 2.9: Estimation of the propensity score for the 45-49 years

Variable	Coefficient	Std. Err.
In couple	-0.076	(0.210)
No French citizenship	0.004	(0.354)
Education (ref:No diploma)		
CAP-BEP	-0.291*	(0.154)
Baccalaureate	-0.201	(0.253)
College degree	-0.077	(0.283)
Distance to retirement (ref: 10-15 years)		
More than 15 years	0.141	(0.156)
Job seniority (ref: 5 years or less)		
6-10 years	0.15	(0.21)
11-20 years	0.412**	(0.193)
More than 20 years	0.579***	(0.175)
Occupation (ref: workers)		
Executives	-0.116	(0.291)
Intermediary	-0.134	(0.184)
Employees	0.516**	(0.254)
Existence of a training plan	1.032***	(0.137)
Sector (ref: building)		
Industry	1.062***	(0.223)
Services	0.311	(0.221)
Part-time job	0.197	(0.426)
Wage quartiles (ref:1st quartile)		
2nd quartile	0.076	(0.196)
3rd quartile	0.188	(0.198)
4th quartile	0.17	(0.246)
Intercept	-0.963*	(0.32)
Number of observations	542	
Pseudo R^2	0.256	

Source: Labor Force Survey (wave 1998) and Training Survey
Significance levels: *: 10% **: 5% ***: 1%

Chapter 3

Inefficient Job Destructions and Training with Holdup[1]

This paper develops an equilibrium search model with endogenous job destructions and where firms decide at the time of job entry how much to invest in match-specific human capital. We first show that job destruction and training investment decisions are strongly complementary. It is possible that there are no firings at equilibrium. Further, training investments are confronted to a holdup problem making the decentralized equilibrium always inefficient. We show therefore that both training subsidies and firing taxes must be implemented to bring back efficiency.

JEL Classification: E24, J41

Keywords: Training, job destruction, holdup, efficiency

[1]This chapter reviews a joint work with Arnaud Chéron, forthcoming in *Labour: Review of labour economics and industrial relations* (2011). I thank the anonymous referee for very helpful comments.

3.1 Introduction

The normative analysis of the link between investment in specific human capital and labor market outcomes dates back to Becker's contribution: within the context of standard competitive theory, workers will not pay for specific training but firms will. However, as Becker (1962) also pointed out, firms might let workers share in the returns (and the costs as well) to reduce both inefficient turnover and investments[2]. The sharing decision supposes that the worker and the firm write a non-renegotiable contract specifying a fixed wage, set in such a way to maximize the expected total surplus. But sharing the costs is possible only if training investments can be preceded by contract negotiations specifying that workers agree to take part in the costs through lower wages. Otherwise, only firms pay all the costs leading to a holdup problem since they cannot get all the returns on their investment[3]. This finally results in under-investments and the decentralized equilibrium is always inefficient.

New developments focus on interactions between employers and employees within the framework of labor market imperfections (see Acemoglu and Pischke (1998) for a survey). In particular, when wages are determined by an ex-post bargaining, contracts are not enforceable, and there is potential for holdups: the Nash assumption implies that a fraction of the expected investment cost, that the firm saves when the worker stays in the match, is actually captured by the worker through a higher wage. As the equilibrium is then always inefficient, there is room for labor market policy. For instance, Acemoglu and Shimer (1999) and Sato and Sugiura (2003) consider ex-ante investments that take place before production begins. Acemoglu and Shimer (1999) analyze the potential for holdup in case of physical capital investments. Sato and Sugiura (2003) consider workers investments in general human capital and investigate the effects of labor market policies both on human capital accumulation and on the holdup problem. Chéron (2005) adds match-specific costs in the standard matching model that can be only partially protected from holdup. This allows the author to re-examine welfare effects of a decrease in equilibrium unemployment.

In this paper, we extend those works to account for endogenous job destructions. More

[2]Hashimoto (1981) first formalized Becker's sharing conjecture in a model with transaction costs related to post-investment uncertainty. Leuven and Oosterkeek (2001) then rigorously considered the role of uncertainty in this model.

[3]See Malcomson (1997) for a survey on "holdup" theory.

precisely, our model is characterized by endogenous hiring and firing decisions and by training investment decisions of firms in specific human capital. We first emphasize the interplay between job destructions and training. In particular, we show that job destructions and training investments are highly complementary since firms have strong incentives to invest in training to protect matches from idiosyncratic productivity shocks. Expected productivity gains due to training investments rise the job tenure, which in turn encourages firms to invest more. Therefore, there might be no firings at equilibrium.

Second, this paper shows how the efficient allocation can be reached in this framework where a holdup problem may arise. In particular, we focus on a holdup problem that results from an "insider wage structure" (Mortensen and Pissarides (1999))[4]. Assuming a training investment at the time of job entry (wasted if the negotiation fails) typically comes to introduce a fixed job creation cost. Reducing the expected job value, holdup then results in an excess of job destructions at equilibrium and on the contrary in a lack of training investments. Hence, we show that it is optimal to implement both training subsidies and firing taxes to achieve efficiency.

More generally, several papers have studied the positive link between employment protection and training investments. Fostering long-term employment, employment protection may promote investments in human capital since longer-lasting employment will increase the expected returns to training. The empirical contributions of Bishop (1991) and Pierre and Scarpetta (2006) put the emphasis on the fact that firms react to strict employment protection by investing more in training. From a natural experiment in France, Messe and Rouland (2011) show that stricter employment protection for older workers rises firms incentives to train them. In a theoretical perspective, Fella (2005) explains that large enough conditional termination penalties improve employers investments in general training if the latter is not directly contractible. The reason is that separation payments ensure firms to capture a positive share of the return as training is vested in the worker on separation. Closer to our paper are Lechthaler (2009) and Belot, Boone and van Ours (2007) who investigate the impact of firings costs on equilibrium unemployment and welfare in a matching model with training. Lechthaler (2009) considers firms investments in general training. Therefore, inefficiencies stem from the fact that training firms do not take

[4]Pissarides (2009) gives some empirical arguments supporting this wage setting: the author shows that fixed job creation costs (paid before the Nash-bargaining of wages) can raise the volatility of unemployment over business cycles, as found in data.

into account that fired workers are more productive in their following relationships as well. Firing costs are useful as mean to raise training investments. Holdup is another potential source of inefficiency, as explored in Belot et al. (2007) who focus on workers investments in firm-specific knowledge. Firing costs work as a commitment device of the employer and workers react with higher investments in firm-specific knowledge. In comparison with those contributions, we consider in this paper specific training investments provided by firms to their workers and show that employment protection is not enough to cover inefficiencies due to holdup in a context of firms investments in specific training. In addition to firing costs, we stress the need for training subsidies to restore social efficiency. Firing taxes and training subsidies are mostly studied separately but here, it is definitively the combination of those two parameters that achieves efficiency. We show that both policy instruments are not unconnected because of the strong complementarity between firing and training decisions. Therefore, our paper suggests that the right design of firing costs should account for the fact that training investments are sub-optimal.

The paper is organized as follows. The next section presents the decentralized partial equilibrium with firstly a two-tier (efficient) wage contract and secondly, an insider wage structure. We then examine the optimal design of labor market policy. A last section concludes.

3.2 Labor market equilibrium

3.2.1 Model environment and labor market flows

We consider a continuous-time equilibrium search model at steady state with endogenous job destruction. The population of workers is a continuum mass. Workers look for jobs and are randomly matched with employers looking for workers to fill vacant units of production. A productive unit is the association of one worker and one firm. Matches are randomly formed according to a constant return to scale matching function $M(V,U)$ that gives the number of hirings (the job creation flows) as a function of the number of vacancies V and the number of unemployed workers U. Each worker matches with a firm with probability $\theta q(\theta) \equiv \frac{M(v,u)}{u}$ where $q(\theta) \equiv \frac{M(v,u)}{v}$ defines the probability to fill a vacancy for a firm, and with $\theta \equiv \frac{v}{u}$ the labor market tightness.

The time of events and of decisions is as follows. First, at the time of match formation, firms decide on the investment in continuing training[5] k, resulting in the human capital of workers $y(k)$[6]. The firm and its worker then bargain over the wage. We assume that the human capital level of a worker is determined at the entry into the job and is constant for all the job tenure. Training investments increase the output of the worker only if she stays with the training firm. In this way, training is assumed to be specific in Becker's (1962) sense as it is fully lost on separation [7]. The productivity of a worker is the sum of a random component ε and a deterministic one $y(k)$[8]. The random component is related to shocks that occur at Poisson rate λ, and where the cdf is $G(\varepsilon)\ \forall\ \varepsilon \in [0, \bar{\varepsilon}]$.

Second, a new value of ε is randomly drawn from its distribution. The worker and the firm then bargain over a new wage if there is a positive surplus to share. In the opposite case, they optimally separate. Job destructions arise when ε falls below an endogenous threshold that depends on the invested amount in continuing training. We denote this threshold $R(k)$. We assume that jobs can also be destroyed exogenously at rate s in the form of voluntary quits of workers[9], so that $s + \lambda G(R(k))$ gives the overall destruction rate of a job matched with a worker who has received a training investment k.

Lastly, whenever an idiosyncratic shock arrives, the firm either accounts for this new value of ε in a new wage negotiation or destroy the job for a zero return.

3.2.2 Firms decisions

Hiring and firing decisions For a firm, the intertemporal value of a filled job depends both on worker's human capital $y(k)$ and on the idiosyncratic component ε. We denote this value

[5]Continuing training refers to training that occurs after leaving school.

[6]The function $y(k)$ is assumed strictly increasing and concave, with $y(0) = 0$.

[7]Using data from the International Adult Literacy Survey, O'Connell (1999) reports (i) that employed adults are more likely than unemployed adults to participate in training, (ii) that employers are by far the the most common financial sponsors of training and (iii) that participation in job-related training is substantially higher than participation in training undertaken for personal or other reasons. All in all, most of the training sessions are enrolled while employed and are not only firms-financed but also job-related, making them apparently more specific. This may support our choice of training modelling.

[8]The additive form of the output of the match we assumed between an endogenous component ($y(k)$) and another exogenous one (ε) clearly simplifies calculations but also fits the usual definition of training. Usually, training is considered as a way to improve workers' skills. Without training, workers are still able to produce but at lower productivity levels. To mention only a few, Lechthaler (2009) and Belot, Boone and van Ours (2007), within the framework of endogenous human capital and productivity shocks, consider an additive form of the output of the match as well.

[9]We will put the emphasis on the role played by this assumption in the section devoted to the labor market equilibrium analysis.

by $J(k, \varepsilon)$. Following Mortensen and Pissarides (1994), we assume that all new jobs are created with maximum productivity, $\overline{\varepsilon}$. We also assume that firms pay all the training cost $C(k)$[10] at the time of match formation (before the wage bargaining). The asset value of a vacancy then writes:

$$rV = -c + q(\theta)[\overline{J}(k) - C(k) - V]$$

with r the discount factor, $c \geq 0$ the flow cost of recruiting a worker and where the corresponding Bellman equations for new and continuing jobs respectively satisfy:[11]

$$r\overline{J}(k) = y(k) + \overline{\varepsilon} - \overline{w}(k) + \lambda \int_{R(k)}^{\overline{\varepsilon}} J(k, x) dG(x) - (\lambda + s)\overline{J}(k)$$
$$rJ(k, \varepsilon) = y(k) + \varepsilon - w(k, \varepsilon) + \lambda \int_{R(k)}^{\overline{\varepsilon}} J(k, x) dG(x) - (\lambda + s)J(k, \varepsilon)$$

where $w(k, \varepsilon)$ denotes the real wage.

As firms open vacancies until all rents from vacant jobs are exhausted, endogenous job creation satisfies the condition:

$$\frac{c}{q(\theta)} = \overline{J}(k) - C(k) \tag{3.1}$$

Job creation entails both a recruiting cost c, proportional to the probability to fill a vacancy, and a training cost $C(k)$ depending on the invested amount.

In turn, the endogenous job destruction rule $J(k, \varepsilon) \leq 0$ leads to a reservation productivity $R(k)$ defined by $J(k, R(k)) = 0$ and such as:

$$R(k) = -y(k) + w(k, R(k)) - \lambda \int_{R(k)}^{\overline{\varepsilon}} J(k, x) dG(x) \tag{3.2}$$

On one hand, the higher the wage, the higher the reservation productivity $R(k)$, and hence

[10] with $C'(0) = 0$, $C'(k) > 0$ and $C''(k) > 0$.

[11] The upper bar refers to new jobs.

the higher the job destruction flow. On the other hand, the higher the option value of filled jobs (expected gains in the future) depending on the training investment, the weaker the job destructions. It follows that a firm may be able to afford to lose instantaneous profit, waiting for future gains that may compensate for. In addition, given the wage, training investments improve job tenure increasing both present and future productivity gains.

Human capital investment decision Firms choose how much specific training they invest in order to maximize the net expected value of a filled job. It follows that the investment decision is stated as:

$$\underset{k \geq 0}{max}\ \overline{J}(k) - C(k) \quad \Longrightarrow \quad C'(k) = \overline{J_1'}(k)$$

In this way, firms decide on the sum they invest in specific training so that the expected marginal return on investment is equaled to its marginal cost. The marginal return particulary depends on the relation between the bargained wage and the investment level, and hence on the potential holdup problem.

3.2.3 Equilibrium with a two-tier wage structure

The Nash wage bargaining

Wages are determined by a Nash bargaining. The firm and the worker share the global surplus generated by a job according to their bargaining power. But, following Mortensen and Pissarides (1999), we first assume that the wage structure that arises when firms are liable for hiring costs (a training cost here) is a two-tier one. On one hand, the initial wage reflects the fact that workers share in the initial hiring (training) cost by accepting a lower wage. On the other hand, renegotiated wages subsequent to match productivity shocks no longer include training costs since they are sunk.

Standard function values of employed and unemployed workers are respectively given by:

$$r\overline{W}(k) = \overline{w}(k) + \lambda \int_{R(k)}^{\overline{\varepsilon}} W(k,x)dG(x) + \lambda G(R(k))U - \lambda \overline{W}(k) + s\big[U - \overline{W}(k)\big]$$
$$rU = z + \theta q(\theta)[\overline{W}(k) - U]$$

where z is home production. The global surplus of a new match $\overline{S}(k) = \overline{J}(k) - C(k) + \overline{W}(k) - U$ is split as follows:

$$\overline{J}(k) - C(k) = (1-\gamma)\overline{S}(k) \quad \text{and} \quad \overline{W}(k) - U = \gamma \overline{S}(k)$$

From the Nash-bargaining rule of the surplus of new matches $(1-\gamma)\left[\overline{W}(k) - U\right] = \gamma\left[\overline{J}(k) - C(k)\right]$ and the job creation condition, we then derive the following expression for the starting wage[12]:

$$\overline{w}(k) = (1-\gamma)z + \gamma\left[y(k) + \overline{\varepsilon} + c\theta\right] - \gamma(r+s+\lambda)C(k)$$

Then, from the Nash-bargaining rule of the surplus generated by a job after a random change in ε, $(1-\gamma)\left[W(k,\varepsilon) - U\right] = \gamma\left[J(k,\varepsilon)\right]$, and the job creation condition, renegotiated wages write:

$$w^C(k,\varepsilon) = (1-\gamma)z + \gamma\left[y(k) + \varepsilon + c\theta\right]$$

Firstly, wages are a weighted average of the reservation wage of the worker (first term in the right-hand side of $\overline{w}(k)$) and secondly, of the productivity and recruitment costs the firm saves (second term). The last term of stating wages reflects the fact that workers agree to share the training cost with the firm. This is in line with the definition of complete contracts: anyone benefiting from an investment must pay one's share of the cost.

The labor market equilibrium

Definition 1. *A labor market equilibrium with a two-tier wage structure is characterized by a triplet* $\{\theta^C, R^C(k), k^C\}$ *solving*[13]*:*

[12] See appendix 3.5.1 for details on derivation.
[13] Again, see appendix 3.5.1 for details on derivation.

$$\frac{c}{q(\theta)} = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\bar{\varepsilon} - R(k)\right] - (1-\gamma)C(k) \tag{3.3a}$$

$$R(k) = -y(k) + z + \left(\frac{\gamma}{1-\gamma}\right)c\theta - \frac{\lambda}{r+\lambda+s}\int_{R(k)}^{\bar{\varepsilon}}[1-G(x)]dx \tag{3.3b}$$

$$C'(k) = \left(\frac{1}{r+s+\lambda G(R(k))}\right)y'(k) \tag{3.3c}$$

Proposition 1. *Let* $C(k) = k,\ y(k) = k^{\alpha},\ G(x) = x,\ x \in [0, \bar{\varepsilon}]$*.*
If $r + s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$*, a unique equilibrium* $\{\theta^C, R^C(k), k^C\}$ *with firings exists.*

Proof. See appendix 3.5.4. □

Corollary 1. *Considering* $r \to 0$*, if* $s = 0$*, it comes that* $R^C(k) = 0$ *and* $k = \left(\frac{\alpha}{r}\right)^{\frac{1}{2-\alpha}}$

Proof. Condition $r + s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ can never be achieved for $s = 0$. □

If there are no exogenous breakups ($s = 0$), firms can definitely reap all the benefits of their training investment. Therefore, firms have strong incentives to invest high enough in training to protect matches from idiosyncratic productivity shocks. Training investment, improving both present and future productivity, increases the job value and then reduces the productivity threshold. Consequently, the probability that the random component of productivity falls below this threshold, as well as the probability that the match endogenously closes, are both smaller. Both probabilities are all the more low than the training investment is high. Anticipating that a higher training investment leads to increase job tenure, firms are finally encouraged to invest more. At the limit, a substantial training investment leads to a so tiny threshold that the job is never destroyed.

However, this is valid as long as there is no risk for the firm to lose the benefits of the training investment. On the opposite, if the exogenous probability of breakups is high enough (in the form of voluntary quits), ie. if s satisfies Proposition 1, training investment is then low enough so that endogenous firings exist at equilibrium[14]. This actually points out how complementary job destruction and training investment decisions are.

[14]For $r \to 0$, $\alpha = 0.1$ and $\lambda = 0.1$, the condition $r + s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is true for $s \geq 0.0417$ whereas for $\alpha = 0.5$ (and again for $\lambda = 0.1$), the condition is true for $s \geq 0.2627$. In other words, the quit rate of workers

3.2.4 Equilibrium with insider wage

The Nash wage bargaining

Effects of training investments on wages and job destructions are highly dependent on the wage setting game. To stress that point, we now derive the partial equilibrium with an insider wage structure as proposed by Mortensen and Pissarides (1999). When firms support hiring costs (such as a training cost), a natural holdup problem may arise. Indeed, new workers have an incentive to renegotiate immediately after been hired as training investments require continuing relationships to be efficient. Starting wages are then not credible. Therefore, second tier wages apply initially as well as subsequent to any shock to match productivity ("insider wage"). The ex-post bargaining process increases employees' threat point. Demanding a higher wage, workers capture some of the rents created by the training cost without paying for, leading finally to a holdup problem.

The global surplus generated by a continuing job $S(k, \varepsilon) = J(k, \varepsilon) + W(k, \varepsilon) - U$ is now split as follows:

$$J(k, \varepsilon) = (1 - \gamma)S(k, \varepsilon) \quad \text{and} \quad W(k, \varepsilon) - U = \gamma S(k, \varepsilon)$$

As firms have to pay the training cost in both cases of success and failure of the wage bargaining, wages write:

$$w(k, \varepsilon) = (1 - \gamma)z + \gamma\big[y(k) + \varepsilon + c\theta\big] + \gamma\theta q(\theta)C(k)$$

The last term of the right-hand side does not appear in renegotiated wages of the two-tier wage structure and refers to holdup. It depends on the investment level and rises the bargained wage: if the negotiation fails, the firm will have to pay another training cost $C(k)$ when it meets a new worker, an event that takes place at rate $\theta q(\theta)$. So, staying in the match, the worker enables the firm to save the expected cost $\theta q(\theta)C(k)$ and wages increase by a fraction γ of that saving by the Nash assumptions. Workers are all the more in a position to threaten firms than

must be at least 4.17%. According to US data from the Department of Labor (Job Openings and Labor Turnover Survey - JOLTS), the total annual quit rate between 2001 and 2008 fluctuates between 22.6% (for 2008) and 27.6% (for 2001). This means that the condition so that an interior solution exists is typically satisfied.

the probability to find a job is high. The holdup problem becomes then stronger.

The labor market equilibrium

Definition 2. *A labor market equilibrium with wage bargaining is characterized by a triplet* $\{\theta^I, R^I(k), k^I\}$ *solving:*

$$\frac{c}{q(\theta)} = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\bar{\varepsilon} - R(k)\right] - C(k) \tag{3.4a}$$

$$R(k) = -y(k) + z + \left(\frac{\gamma}{1-\gamma}\right)c\theta + \frac{\gamma}{1-\gamma}\theta q(\theta)C(k) - \frac{\lambda}{r+\lambda+s}\int_{R(k)}^{\bar{\varepsilon}}[1-G(x)]dx \tag{3.4b}$$

$$C'(k) = \left(\frac{1-\gamma}{r+s+\lambda G(R(k))+\gamma\theta q\theta}\right)y'(k) \tag{3.4c}$$

Proposition 2. *Let* $C(k) = k,\ y(k) = k^{\alpha},\ G(x) = x,\ x \in [0, \bar{\varepsilon}]$.
$r+s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ *is a sufficient condition for a unique equilibrium* $\{\theta^I, R^I(k), k^I\}$ *with some firings to exist.*

Proof. See appendix 3.5.4. □

To get back to Corollary 1, holdup counteracts the positive effects of the training investment on the job tenure by increasing wages. Thereby, the risk firms face to lose all or a part of the training investment rises. Holdup reduces then the incentives firms have to invest highly in training.

3.3 Optimal labor market policy

3.3.1 The efficient allocation

We derive the optimal allocation by maximizing steady-state output with respect to the labor market tightness $\theta^{\star}$, the reservation productivity $R^{\star}(k)$ and the training investment $k^{\star}$. The problem of the planner is stated as follows:

$$\underset{\{\theta,R(k),k\}}{max} \int_0^\infty e^{-rt} \left[\overline{y} + uz - c\theta u - \theta q(\theta) u C(k)\right] dt \tag{3.5}$$

subject to the evolution of u and $\overline{y}$:

$$\dot{u} = (1-u)\left[\lambda G(R(k)) + s\right] - u\theta q(\theta)$$

$$\dot{\overline{y}} = u\theta q(\theta)\left[\overline{\varepsilon} + y(k)\right] + \lambda(1-u)\left[1 - G(R(k))\right] y(k) + \lambda(1-u)\int_{R(k)}^{\overline{\varepsilon}} \varepsilon dG(\varepsilon) - (\lambda + s)\overline{y}$$

Definition 3. *Defining* $\eta(\theta) = \frac{-\theta q'(\theta)}{q(\theta)}$*, the efficient labor market allocation is then characterized by a triplet* $\{\theta^\star, R^\star(k), k^\star\}$ *solving:*

$$\frac{c}{q(\theta^\star)} = \left(\frac{1-\eta(\theta^\star)}{r+\lambda+s}\right)\left[\overline{\varepsilon} - R^\star(k)\right] - (1-\eta(\theta^\star))C(k^\star) \tag{3.6a}$$

$$R^\star(k) = -y(k^\star) + z + \left(\frac{\eta(\theta^\star)}{1-\eta(\theta^\star)}\right) c\theta^\star - \frac{\lambda}{r+\lambda+s}\int_{R^\star(k)}^{\overline{\varepsilon}} [1-G(x)]dx \tag{3.6b}$$

$$C'(k^\star) = \left(\frac{1}{r+s+\lambda G(R^\star(k))}\right) y'(k^\star) \tag{3.6c}$$

Proposition 3. *Let* $C(k) = k,\ y(k) = k^\alpha,\ G(x) = x,\ x \in [0, \overline{\varepsilon}]$.
If $r + s > \alpha^{\frac{1}{2-\alpha}} \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$*, a unique efficient allocation* $\{\theta^\star, R^\star(k), k^\star\}$ *with some firings exists.*

Proof. See appendix 3.5.4. □

Property 1. *The Hosios condition* $\gamma = \eta(\theta)$ *does not achieve equilibrium efficiency in the insider wage case whereas it does in the two-tier wage structure.*

Proof. Straightforward by comparing expressions of $\{\theta^\star, R^\star(k), k^\star\}$ in Definition 5 to $\{\theta^C, R^C(k), k^C\}$ in Definition 1 and to $\{\theta^I, R^I(k), k^I\}$ in Definition 2. □

Equilibrium choices are not efficient due to holdup that reduces the expected job value. In order to recognize the mechanisms that make the equilibrium inefficient, we explore separately the free entry condition of firms (3.4a), the reservation productivity (3.4b) and the investment decision condition of firms (3.4c). This helps us to consider labor policies that remove the distortions.

First, we examine the investment decision condition of firms. Evaluating (3.4c) at $R^I(k) = R^\star(k)$, we obtain the investment level under the assumption that the reservation productivity takes the optimal value $k^I|_{R(k)=R^\star(k)}$. Comparing $k^I|_{R(k)=R^\star(k)}$ with $k^\star$ leads to the following lemma.

Lemma 1. *The investment decision condition of firms generates under-investments:* $k^I|_{R(k)=R^\star(k)} < k^\star$.

Proof. See appendix 3.5.4. □

Training investments are not enough at equilibrium. With an insider wage structure, not only workers do not contribute to the training cost but they also capture some of the returns on the training investment (holdup). Therefore, firms under-invest in training their workers because they have to pay all the training cost while they get only a fraction of the gains. Thereby, the job value is lowered by holdup.

Second, we address the reservation productivity of firms (3.4b). Evaluating (3.4b) at $k^I = k^\star$ and $\theta^I = \theta^\star$ gives the reservation productivity under the assumptions that the investment level and the market tightness are both optimal $R^I(k)|_{k=k^\star,\theta=\theta^\star}$. Comparing $R^I(k)|_{k=k^\star,\theta=\theta^\star}$ with $R^\star(k)$ gives the following lemma.

Lemma 2. *Under the Hosios condition, the productivity reservation of firms generates too many job destructions:* $R^I(k)|_{k=k^\star,\theta=\theta^\star} > R^\star(k)$.

Proof. See appendix 3.5.4. □

There are too many job destructions at equilibrium. Firms close endogenously too many jobs at equilibrium because holdup, increasing wages, rises the productivity threshold. Therefore, productivity gains induced by the training investments are not enough to improve job tenure.

Finally, we examine the free entry condition of firms (3.4a). Evaluating (3.4a) at $k^I = k^\star$ and $R^I(k) = R^\star(k)$ gives the market tightness under the assumptions that the investment level and the reservation productivity are both optimal $\theta^I(k)|_{k=k^\star,R(k)=R^\star(k)}$. Comparing $\theta^I(k)|_{k=k^\star,R(k)=R^\star(k)}$ with $\theta^\star(k)$ gives the following lemma.

Lemma 3. *Under the Hosios condition, the free entry condition generates too little labor market tightness:* $\theta^I(k)|_{k=k^\star,R(k)=R^\star(k)} < \theta^\star(k)$.

Proof. See appendix 3.5.4. □

Firms do not post enough vacant jobs at equilibrium. Here again, the inefficiency comes from the contract type that allows workers to capture some of the rents following the training investment without contributing to its cost (holdup). As shown previously, this rises firms' reservation productivity, which in turn decreases both the job tenure and the expected job value. Further, the expected job value is also reduced since firms have to bear all the training cost. Firms then post too few vacant jobs compared to what would be optimal.

3.3.2 Restoring efficiency

This last section investigates the way to restore the optimality of equilibrium choices. Job destruction decisions and training investment decisions are strongly complementary: a fraction γ of the expected training cost $\theta q(\theta)C(k)$, that the firm saves when the worker stays in the match, is captured by the worker through the wage bargaining (holdup). This rises the productivity threshold, leading finally to an excess of job destructions. Firing taxes F can be implemented to reach the efficient level of job destructions, together with training subsidies T get at the time of match formation in order to lower the training cost.

As the training subsidy decreases the training cost, the value of a vacancy is now such as:

$$rV = -c + q(\theta)\left[\overline{J}(k) - C(k) + T - V\right]$$

The free entry condition implies $V = 0$ and then:

$$\frac{c}{q(\theta)} = \overline{J}(k) - C(k) + T \tag{3.7}$$

The reservation productivity $R(k)$ is now defined by $J(k, R(k)) = -F$. In the context of an insider wage structure, the surplus sharing rule is now such that $W(a,\varepsilon) - U = \gamma S(a,\varepsilon)$ and $J(a,\varepsilon) + F = (1-\gamma)S(a,\varepsilon)$ where $S(a,\varepsilon) = J(a,\varepsilon) + F + W(a,\varepsilon) - U$. We therefore derive the following wage expression:

$$w^P(k,\varepsilon) = (1-\gamma)z + \gamma\left[y(k) + \varepsilon + c\theta\right] + \gamma(r+s+\theta q(\theta))F + \gamma\theta q(\theta)\left[C(k) - T\right]$$

On the one hand, the training subsidy reduces the training cost (last term of the right-hand side). But, on the other hand, workers are now in a position to capture also a fraction γ of the firing taxes the firm saves if the negotiation does not fail or when the worker quits voluntary the firm (second term of the right-hand side).

Proposition 4. *Assuming $\gamma = \eta(\theta)$, the optimal labor market policy with training subsidies and firing taxes $\{T, F\}$ solves*[15]*:*

$$\begin{cases} T &= \gamma C(k^\star) + F \\ \\ F &= \frac{\theta^\star q(\theta^\star)}{r+s}\gamma C(k^\star) \end{cases}$$

where $k^\star$ solves the optimal allocation.

Proof. See appendix 5.6.2. □

Firstly, firing taxes depend on the expected value of the distortion on job destructions, ie. holdup that increases wages. While $\gamma\theta q(\theta)C(k)$ defines the instantaneous value of this distortion, $1/[r+s]$ defines the discount factor that depends not only on the interest but also

[15]Similarly, an appropriate combination of training subsidies and unemployment benefits can also be implemented to reach the efficient allocation. On one hand, as the excess of job destructions comes from holdup that rises wages, negative unemployment benefits can be used to reduce outside options of workers and thus wages. On the other hand, workers do not contribute to the training cost but get a fraction of the gains of the training investment (through the productivity gains). Then, training subsidies can be used to share the training cost between both parties. More precisely, unemployment benefits have to cover the fraction of the expected training cost that the firm saves when the worker stays in the match and that the latter gets while bargaining over the wage (holdup). In this way, unemployment benefits remove the distortion on job destructions. About training subsidies, they have to cover the fraction of the training cost that the worker should have born as she share in the returns. Training subsidies depend on worker's bargaining power that determines the fraction of the gains she gets. See appendix 3.5.3 for details.

on the probability of voluntary quits: with probability s, firms will not have to pay firing taxes. Therefore, the higher the probability the worker quits voluntary the firm, the lower the distortion, hence firing taxes.

Secondly, training subsidies result from the distortion on invested amounts in training: workers get a fraction γ of the expected training cost that firms save when they stay in the match. Then, the optimal training subsidy integrates this distortion plus the negative incidence of the firing tax on job creations.

3.4 Conclusion

This paper mainly emphasized that there exists a strong complementarity between firing and training decisions. This has first allowed to establish under which conditions positive firing decisions occur at equilibrium. We have then stressed the need for both firing taxes and training subsidies in order to restore equilibrium efficiency when a holdup problem arises. More generally, our work shows that the interplay between firing and training decisions should lead to reexamine the instruments of economic policy used to bring back efficiency. In this way, training subsidies turn out to be a central instrument.

The need for training subsidies could also be relevant within frameworks where inefficient job destructions do not arise as a result of holdup. For instance, they could be important in the presence of (unobservable) heterogeneity of workers that results in too many job destructions and too few training investments for low-skilled workers at equilibrium.

We have assumed that human capital is purely match-specific for convenience. This is a strong assumption and skills are practically neither purely general nor specific. But, considering general human capital as a part of worker's productivity as well would imply to consider *ex ante* heterogenous workers: since accumulation (depreciation) of general human capital depends on the length of employment (unemployment) spells, we should determine the steady-state of the distribution of general human capital. However, the complementarity between both firms decisions would still stand since firms invest in training to protect from negative shocks. Further, the problem of training underinvestments would be stronger with both general and specific human capital because of the transferability of skills. Not only workers would benefit from the investment without bearing the cost (holdup), but also future employers ("poaching externality", Acemoglu (1997) or Lechthaler (2009)). The interaction between firing taxes and training subsidies we have emphasized would be even more relevant. Looking at the complementarity between investments in general and specific human capital would be an interesting issue for future work.

3.5 Appendix

3.5.1 Equilibrium equations under two-tier wage contract

Wage setting First, from the Nash-bargaining rule of the surplus from new matches $(1-\gamma)\left[\overline{W}(k)-U\right]=\gamma\left[\overline{J}(k)-C(k)\right]$, entry wages write $\overline{w}(k)=(1-\gamma)z+\gamma\left(y(k)+\overline{\varepsilon}\right)-\gamma(r+s+\lambda)C(k)-\gamma\theta q(\theta)C(k)+\gamma\theta q(\theta)\overline{J}(k)$. From this Nash-bargaining rule and the job creation condition $\frac{c}{q(\theta)}=\overline{J}(k)-C(k)$, entry wages finally solve:

$$\overline{w}(k)=(1-\gamma)z+\gamma\left[y(k)+\overline{\varepsilon}+c\theta\right]-\gamma(r+s+\lambda)C(k)$$

Then, from the Nash-bargaining rule of the surplus generated by a job after a random change in ε, $(1-\gamma)\left[W(k,\varepsilon)-U\right]=\gamma\left[J(k,\varepsilon)\right]$, and the job creation condition, renegotiated wages write:

$$w^C(k,\varepsilon)=(1-\gamma)z+\gamma\left[y(k)+\varepsilon+c\theta\right]$$

Job creations As $J(k,R(k))=0$ in $(r+\lambda+s)\left[\overline{J}(k)-J(k,R(k))\right]=\overline{\varepsilon}-R(k)-\overline{w}(k)+w(k,R(k))$ and using wage expressions, it follows $(r+\lambda+s)\overline{J}(k)=(1-\gamma)\left[\overline{\varepsilon}-R(k)\right]+\gamma(r+s+\lambda)C(k)$.

The job creation condition $\frac{c}{q(\theta)}=\overline{J}(k)-C(k)$ finally leads to derive job creation equation:

$$\frac{c}{q(\theta)}=\left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\overline{\varepsilon}-R(k)\right]-(1-\gamma)C(k)$$

Job destructions Replacing the renegotiated wage expression in the reservation productivity threshold $R(k)=-y(k)+w(k,R(k))-\lambda\int_{R(k)}^{\overline{\varepsilon}}J(k,x)dG(x)$ leads to the job destruction equation:

$$R(k) = -y(k) + z + \frac{\gamma}{1-\gamma} c\theta - \frac{\lambda}{r+\lambda+s} \int_{R(k)}^{\bar{\varepsilon}} [1 - G(x)]\, dx$$

since integrating by parts $\lambda \int_{R(k)}^{\bar{\varepsilon}} J(k,x)dG(x) = \lambda \int_{R(k)}^{\bar{\varepsilon}} J'(k,x)[1-G(x)]dx$, with $J(k,x) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[x - R(k)] + \gamma C(k)$.

Then, it turns: $\frac{\partial R(k)}{\partial k} = -\left(\frac{r+\lambda+s}{r+\lambda G(R(k))+s}\right) y'(k)$.

Training investment As $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\bar{\varepsilon} - R(k)] + \gamma C(k)$ at equilibrium, the training investment level is determined by:

$$\max_k \overline{J}(k) - C(k) \quad \Leftrightarrow \quad \max_k \left(\frac{1-\gamma}{r+\lambda+s}\right)[\bar{\varepsilon} - R(k)] - (1-\gamma)C(k)$$

where the first order condition implies $C'(k) = -\left(\frac{1}{r+s+\lambda}\right)\frac{\partial R(k)}{\partial k}$. The training equation is finally defined by $C'(k) = \left(\frac{1}{r+s+\lambda G(R(k))}\right) y'(k)$.

Assuming both $r \to 0$ and $\gamma = \eta(\theta)$, let remark that the general equilibrium with a two-tier wage structure is first-best efficient.

3.5.2 Equilibrium equations with insider wage contract

Job creations As $J(k, R(k)) = 0$ in $(r+\lambda+s)\left[\overline{J}(k) - J(k,R(k))\right] = \bar{\varepsilon} - R(k) - w(k,\bar{\varepsilon}) + w(k, R(k))$ and using the wage expression, it follows $(r+\lambda+s)\overline{J}(k) = (1-\gamma)[\bar{\varepsilon} - R(k)]$. Equation (5.2) leads then to derive job creation equation:

$$\frac{c}{q(\theta)} = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\bar{\varepsilon} - R(k)\right] - C(k)$$

Job destructions We derive job destruction equation from reservation productivity (5.1) and wage expression $w(k, R(k))$. By integrating by parts, it comes that $\int_{R(k)}^{\bar{\varepsilon}} J(k,x)dG(x) = \int_{R(k)}^{\bar{\varepsilon}} J'(k,x)[1-G(x)]dx$. Furthermore, noticing that $J'(k,x) = \frac{1-\gamma}{r+\lambda+s}$ as $J(k,x) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[x - R(k)]$, it follows that $\int_{R(k)}^{\bar{\varepsilon}} J'(k,x)[1-G(x)]dx = \frac{1-\gamma}{r+\lambda+s}\int_{R(k)}^{\bar{\varepsilon}}[1-G(x)]dx$. Job destruction equation is finally defined by:

$$R(k) = -y(k) + z + \left(\frac{\gamma}{1-\gamma}\right) c\theta + \frac{\gamma}{1-\gamma}\theta q(\theta)C(k) - \frac{\lambda}{r+\lambda+s}\int_{R(k)}^{\overline{\varepsilon}} [1-G(x)]dx$$

Training investment level As $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)]$, the training investment level is determined by:

$$\max_k \overline{J}(k) - C(k) \quad \Leftrightarrow \quad \max_k \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)] - C(k)$$

First order condition implies $C'(k) = -\left(\frac{1-\gamma}{r+\lambda+s}\right)\frac{\partial R(k)}{\partial k}$,
with $\frac{\partial R(k)}{\partial k} = \left(\frac{r+\lambda+s}{r+\lambda G(R(k))+s}\right)\left[-y'(k) + \left(\frac{\gamma}{1-\gamma}\right)\theta q(\theta) C'(k)\right]$,
and then $C'(k) = \left(\frac{1-\gamma}{r+s+\lambda G(R(k))+\gamma\theta q(\theta)}\right) y'(k)$.

3.5.3 Equilibrium equations with labor market policy (and insider wage)

Implementing firing taxes and training subsidies

Wage setting First, the surplus from a match, $S(k,\varepsilon) = J(k,\varepsilon) + F + W(k,\varepsilon) - U$ shared such as $W(k,\varepsilon) - U = \frac{\gamma}{1-\gamma} J(k,\varepsilon)$, implies $w(k,\varepsilon) = (1-\gamma)z + \gamma[y(k)+\varepsilon] + (1-\gamma)\theta q(\theta)[W(k,\overline{\varepsilon}) - U] + \gamma(r+s)F$.

Then, from the job creation condition $\frac{c}{q(\theta)} = \overline{J}(k) - C(k) + T$ and using the Nash-bargaining rule of the surplus, wages finally write:

$$w^P(k,\varepsilon) = (1-\gamma)z + \gamma[y(k)+\varepsilon+c\theta] + \gamma(r+s+\theta q(\theta))F + \gamma\theta q(\theta)[C(k) - T]$$

Job creations Endogenous job destruction rule now such as $J(k,\varepsilon) < -F$ leads to a reservation productivity $R(k)$ defined by $J(k,R(k)) = -F$. Therefore, it follows $(r+s+\lambda)\left[\overline{J}(k) - J(k,R(k))\right] = \overline{\varepsilon} - R(k) - w^P(k,\overline{\varepsilon}) + w^P(k,R(k))$. Using the wage expression it comes that:

$$(r+s+\lambda)\left[\overline{J}(k) - J(k,R(k))\right] = (1-\gamma)[\overline{\varepsilon} - R(k)] \quad \Leftrightarrow \quad \overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)] - F$$

Job creations equation is then defined by:

$$\frac{c}{q(\theta^P)} = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\overline{\varepsilon} - R^P(k)\right] - C(k^P) + T - F$$

Making job creations optimal when Hosios condition holds implies:

$$\frac{c}{q(\theta^P)} = \frac{c}{q(\theta^*)} \quad \Leftrightarrow \quad T = \gamma C(k^P) + F$$

Job destructions The reservation productivity $R(k)$ defined by $J(k, R(k)) = -F$ is such as:

$$R(k) = -y(k) + w^P(k, R(k)) - (r+s)F - \lambda \int_{R(k)}^{\overline{\varepsilon}} \left[J(k,x) + F\right] dG(x)$$

As mentioned before, $J(k, R(k)) = -F$ implies $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\overline{\varepsilon} - R(k)\right] - F$. It then comes that $\lambda \int_{R(k)}^{\overline{\varepsilon}} \left[J(k,x) + F\right] dG(x) = \left(\frac{\lambda(1-\gamma)}{r+\lambda+s}\right) \int_{R(k)}^{\overline{\varepsilon}} \left[x - R(k)\right] dG(x)$. Integrating by parts this term and replacing the wage expression finally leads to the job destruction equation:

$$\begin{aligned} R^P(k) = &-y(k^P) + z + \left(\frac{\gamma}{1-\gamma}\right) c\theta + \left(\frac{\gamma}{1-\gamma}\right) \theta q \theta \left[C(k^P) - T + F\right] - (r+s)F \\ &- \frac{\lambda}{r+\lambda+s} \int_{R^P(k)}^{\overline{\varepsilon}} [1 - G(x)] dx \end{aligned}$$

Given that $T = \gamma C(k) + F$, making job destructions optimal when Hosios condition holds implies:

$$R^P(k) = R(k)^* \quad \Leftrightarrow \quad F = \left(\frac{\gamma \theta q(\theta)}{r+s}\right) C(k^P)$$

The job destruction equation at equilibrium is then defined by:

$$R^P(k) = -y(k^P) + z + \left(\frac{\gamma}{1-\gamma}\right) c\theta - \frac{\lambda}{r+\lambda+s} \int_{R^P(k)}^{\overline{\varepsilon}} [1 - G(x)] dx$$

Then it turns $\frac{\partial R^P(k)}{\partial k^P} = -\left(\frac{r+\lambda+s}{r+\lambda G(R(k))+s}\right) y'(k^P)$.

Training investment level As $T = \gamma C(K) + F$ and as $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)] - F$ at equilibrium, the training investment level is determined by:

$$\max_k \overline{J}(k) - C(k) + T \quad \Leftrightarrow \quad \max_k \left(\frac{1-\gamma}{\lambda+s}\right)\left[\overline{\varepsilon} - R^P(k)\right] - (1-\gamma)C(k)$$

where the first order condition implies $C'(k^P) = -\left(\frac{1}{r+\lambda+s}\right)\frac{\partial R^P(k)}{\partial k^P}$ with $\frac{\partial R^P(k)}{\partial k^P} = -\left(\frac{r+\lambda+s}{r+\lambda G(R^P(k))+s}\right)y'(k^P)$.

Training equation is finally defined by $C'(k^P) = \left(\frac{1}{r+\lambda G(R^P(k))+s}\right)y'(k^P)$.

Implementing unemployment benefits and training subsidies

Wage setting As unemployment benefits b rise the reservation wage of the worker, the value of a unemployed worker satisfies now: $rU = b + z + \theta q(\theta)\left[W(k,\varepsilon) - U\right]$. With training subsidies T and from the Nash-bargaining rule of the surplus and the job creation condition $\frac{c}{q(\theta)} = \overline{J}(k) - C(k) + T$, wages finally write:

$$w^B(k,\varepsilon) = (1-\gamma)(z+b) + \gamma\left[y(k) + \varepsilon + c\theta\right] + \gamma\theta q(\theta)[C(k) - T]$$

Job creations As $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)]$, job creation equation is defined by:

$$\frac{c}{q(\theta^B)} = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\overline{\varepsilon} - R^B(k)\right] - C(k^B) + T$$

Making job creations optimal when Hosios condition holds implies:

$$\frac{c}{q(\theta^B)} = \frac{c}{q(\theta^*)} \quad \Leftrightarrow \quad T = \gamma C(k^B)$$

Job destructions With unemployment benefits, the reservation productivity $R(k)$ defined by $J(k, R(k)) = 0$ is such as:

$$R^B(k) = -y(k^B) + (b+z) + \left(\frac{\gamma}{1-\gamma}\right)c\theta + \left(\frac{\gamma}{1-\gamma}\right)\theta q\theta\left[C(k^B) - T\right] - \frac{\lambda}{r+\lambda+s}\int_{R^B(k)}^{\overline{\varepsilon}}[1-G(x)]dx$$

Given that $T = \gamma C(k)$, making job destructions optimal when Hosios condition holds implies:

$$R^B(k) = R(k)^* \quad \Leftrightarrow \quad b = -\gamma\theta q(\theta)C(k^B)$$

The job destruction equation at equilibrium is then defined by:

$$R^B(k) = -y(k^B) + z + \left(\frac{\gamma}{1-\gamma}\right) c\theta - \frac{\lambda}{r+\lambda+s}\int_{R^B(k)}^{\overline{\varepsilon}} [1 - G(x)]dx$$

Then it turns $\frac{\partial R(k)^B}{\partial k} = -\left(\frac{r+\lambda+s}{r+\lambda G(R(k))+s}\right) y'(k)$.

Training investment level As $T = \gamma C(K)$ and as $\overline{J}(k) = \left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - R(k)]$ at equilibrium, the training investment level is determined by:

$$\max_k \overline{J}(k) - C(k) + T \quad \Leftrightarrow \quad \max_k \left(\frac{1-\gamma}{\lambda+s}\right)[\overline{\varepsilon} - R(k)] - (1-\gamma)C(k)$$

where the first order condition implies $C'(k) = -\left(\frac{1}{r+\lambda+s}\right)\frac{\partial R(k)}{\partial k}$ with $\frac{\partial R(k)}{\partial k} = -\left(\frac{r+\lambda+s}{r+\lambda G(R(k))+s}\right) y'(k)$.

Training equation is finally defined by $C'(k^B) = \left(\frac{1}{r+\lambda G(R(k))+s}\right) y'(k^B)$.

3.5.4 Proofs of propositions

Two-tier equilibrium existence proof

On one hand, combining equations (3.3b) and (3.3c) of the optimal allocation implies $R(k) + \left[\frac{\alpha}{r+s+\lambda R(k)}\right]^{\frac{\alpha}{1-\alpha}} - \frac{1}{2}\left(\frac{\lambda}{r+\lambda+s}\right)\left[(1-\overline{\varepsilon})^2 - (1-R(k))^2\right] = z + \left(\frac{\gamma}{1-\gamma}\right)c\theta$.

Therefore, $\frac{dR(k)}{d\theta} = \left(\frac{c\gamma}{1-\gamma}\right)\left[\frac{(r+\lambda+s)(1-\alpha)}{(1-\alpha)[r+s+\lambda R(k)] - \lambda(r+\lambda+s)\alpha^{\frac{1}{1-\alpha}}[r+s+\lambda R(k)]^{\frac{1}{1-\alpha}}}\right]$. Job destruction equation is then an upward-sloping curve in the reservation productivity-tightness space if the denominator of the second term on the right-hand side is positive, namely if $r + s + \lambda R(k) > \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}\alpha^{\frac{1}{2-\alpha}}$. It is then straightforward to see that $r + s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

On the other hand, from equation (3.3a) and (3.3c), it follows that

$$\frac{dR(k)}{d\theta} = \left[\frac{cq'(\theta)}{(1-\eta(\theta))(q(\theta))^2}\right]\left[\frac{(r+\lambda+s)(1-\alpha)}{(1-\alpha) - \lambda(r+\lambda+s)\alpha^{\frac{1}{1-\alpha}}[r+s+\lambda R(k)]^{\frac{-1}{1-\alpha}-1}}\right]$$

As $q'(\theta) < 0$, the job creation equation is a downward-sloping curve if $r + s + \lambda R(k) > \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}} \alpha^{\frac{1}{2-\alpha}}$. Again, $r + s > \alpha^{\frac{1}{2-\alpha}} \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

Finally, there exists a unique equilibrium if the intersection of job destruction and job creation curves corresponds with both θ and $R(k)$ positive. This is the case if the job destruction curve evaluated for $\theta = 0$ is below the job creation curve also evaluated for $\theta = 0$, namely if:

$$
\begin{aligned}
&2[R(k) - z][r + s + \lambda R(k)]^{\frac{1}{1-\alpha}+\frac{\alpha}{1-\alpha}}[\bar{\varepsilon} - R(k)] + 2\alpha^{\frac{\alpha}{1-\alpha}}[\bar{\varepsilon} - R(k)][r + s + \lambda R(k)]^{\frac{1}{1-\alpha}} \\
&- \lambda\alpha^{\frac{1}{1-\alpha}}[r + s + \lambda R(k)]^{\frac{\alpha}{1-\alpha}}[(1 - \bar{\varepsilon})^2 - (1 - R(k))^2] > 0
\end{aligned}
$$

As $z < R(k)$ and $\bar{\varepsilon} > R(k)$, this inequality is true.

Insider equilibrium existence proof

On one hand, combining equations (3.4b) and (3.4c) of the decentralized equilibrium implies

$$
\begin{aligned}
&\underbrace{R(k) - \frac{1}{2}\left(\frac{r+\lambda}{s+\lambda}\right)\left[(1-\bar{\varepsilon})^2 - (1-R(k))^2\right]}_{\psi(R(k))} \\
&+ \underbrace{\left[\frac{\alpha(1-\gamma)}{r+s+\lambda R(k)+\gamma p(\theta)}\right]^{\frac{\alpha}{1-\alpha}} - \frac{\gamma p(\theta)}{1-\gamma}\left[\frac{\alpha(1-\gamma)}{r+s+\lambda R(k)+\gamma p(\theta)}\right]^{\frac{1}{1-\alpha}}}_{\varphi(R(k),\theta)} \\
&= z + \frac{\gamma}{1-\gamma}c\theta
\end{aligned}
$$

Therefore, $\frac{dR(k)}{d\theta} = \frac{\frac{\gamma}{1-\gamma}c - \varphi_2'(R(k),\theta)}{\psi'(R(k)) + \varphi_1'(R(k),\theta)}$ with $\psi'(R(k)) > 0$, $\varphi_1'(R(k),\theta) < 0$ and $\varphi_2'(R(k),\theta) < 0$.

First, $\frac{\gamma}{1-\gamma}c - \varphi_2'(R(k),\theta)$ is positive as $\varphi_2'(R(k),\theta) < 0$. Job destruction equation is then an upward-sloping curve in the reservation productivity-tightness space if $\psi'(R(k)) + \varphi_1'(R(k),\theta) > 0$ $\Leftrightarrow r + s + \lambda R(k) + \gamma p(\theta) > \alpha^{\frac{1}{(2-\alpha)}}(1-\gamma)^{\frac{\alpha}{(2-\alpha)}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$. It is then straightforward to see that $r + s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

On the other hand, from equation (3.4a) and (3.4c), it follows that:

$$\frac{dR(k)}{d\theta} = \left(\frac{r+\lambda+s}{(q(\theta))^2}\right) \times \left[\frac{(q(\theta))^2[\alpha(1-\gamma)]^{\frac{1}{1-\alpha}}\gamma p'(\theta)\left[r+s+\lambda R(k)+\gamma p(\theta)\right]^{\frac{-1}{1-\alpha}-1} + cq'(\theta)(1-\alpha)(1-\gamma)}{(1-\alpha)(1-\gamma) - \lambda(r+\lambda+s)[\alpha(1-\gamma)]^{\frac{1}{1-\alpha}}\left[r+s+\lambda R(k)+\gamma p(\theta)\right]^{\frac{-1}{1-\alpha}-1}}\right]$$

As $q'(\theta) < 0$ and $p'(\theta) > 0$, the job creation equation is a downward-sloping curve either if the numerator is positive and the denominator negative (first case), or if the numerator is negative and the denominator positive (second case). But the numerator is clearly negative if $\gamma = 0$, suggesting that the second case is the most likely. The job creation equation is then a downward-sloping curve if the denominator is positive, which implies $r + s + \lambda R(k) + \gamma p(\theta) > \alpha^{\frac{1}{(2-\alpha)}}(1-\gamma)^{\frac{\alpha}{(2-\alpha)}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$. Again, it is then straightforward to see that $r+s > \alpha^{\frac{1}{2-\alpha}}\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

Finally, there exists a unique equilibrium if the intersection of job destruction and job creation curves corresponds with both θ and $R(k)$ positive. This is the case if the job destruction curve evaluated for $\theta = 0$ is below the job creation curve also evaluated for $\theta = 0$, namely if:

$$\begin{aligned} &2[R(k)-z](1-\gamma)\left[\bar{\varepsilon}-R(k)\right]\left[s+\lambda R(k)\right]^{\frac{1}{1-\alpha}+\frac{\alpha}{1-\alpha}} \\ &+ 2[\alpha(1-\gamma)]^{\frac{\alpha}{1-\alpha}}(1-\gamma)\left[\bar{\varepsilon}-R(k)\right]\left[s+\lambda R(k)\right]^{\frac{1}{1-\alpha}} \\ &- \lambda[\alpha(1-\gamma)]^{\frac{1}{1-\alpha}}\left[s+\lambda R(k)\right]^{\frac{\alpha}{1-\alpha}}\left[(1-\bar{\varepsilon})^2-(1-R(k))^2\right] > 0 \end{aligned}$$

As $z < R(k)$ and $(1-\bar{\varepsilon})^2 - (1-R(k))^2 < 0$ (since $\bar{\varepsilon} > R(k)$), this inequality is true.

Optimum existence proof

On one hand, combining equations (3.6b) and (3.6c) of the optimal allocation implies $R(k) + \left[\frac{\alpha}{r+s+\lambda R(k)}\right]^{\frac{\alpha}{1-\alpha}} - \frac{1}{2}\left(\frac{\lambda}{r+\lambda+s}\right)\left[(1-\bar{\varepsilon})^2-(1-R(k))^2\right] = z + \left(\frac{\eta(\theta)}{1-\eta(\theta)}\right)c\theta$.
Therefore, $\frac{dR(k)}{d\theta} = \left(\frac{c\eta(\theta)}{1-\eta(\theta)}\right)\left[\frac{(r+\lambda+s)(1-\alpha)}{(1-\alpha)[r+s+\lambda R(k)]-\lambda(r+\lambda+s)\alpha^{\frac{1}{1-\alpha}}[r+s+\lambda R(k)]^{\frac{1}{1-\alpha}}}\right]$. Job destruction equation is then an upward-sloping curve in the reservation productivity-tightness space if the denominator of the second term on the right-hand side is positive, namely if $r + s + \lambda R(k) >$

$\left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}} \alpha^{\frac{1}{2-\alpha}}$. It is then straightforward to see that $r+s > \alpha^{\frac{1}{2-\alpha}} \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

On the other hand, from equation (3.6a) and (3.6c), it follows that:

$$\frac{dR(k)}{d\theta} = \left[\frac{cq'(\theta)}{(1-\eta(\theta))(q(\theta))^2}\right] \left[\frac{(r+\lambda+s)(1-\alpha)}{(1-\alpha) - \lambda(r+\lambda+s)\alpha^{\frac{1}{1-\alpha}} \left[r+s+\lambda R(k)\right]^{\frac{-1}{1-\alpha}-1}}\right]$$

As $q'(\theta) < 0$, the job creation equation is a downward-sloping curve if $r+s+\lambda R(k) > \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}} \alpha^{\frac{1}{2-\alpha}}$. Again, $r+s > \alpha^{\frac{1}{2-\alpha}} \left[\frac{\lambda(r+\lambda+s)}{1-\alpha}\right]^{\frac{1-\alpha}{2-\alpha}}$ is a sufficient condition.

Finally, there exists a unique equilibrium if the intersection of job destruction and job creation curves corresponds with both θ and $R(k)$ positive. This is the case if the job destruction curve evaluated for $\theta = 0$ is below the job creation curve also evaluated for $\theta = 0$, namely if:

$$\begin{aligned} &2[R(k)-z][r+s+\lambda R(k)]^{\frac{1}{1-\alpha}+\frac{\alpha}{1-\alpha}}[\bar{\varepsilon} - R(k)] + 2\alpha^{\frac{\alpha}{1-\alpha}}[\bar{\varepsilon} - R(k)][r+s+\lambda R(k)]^{\frac{1}{1-\alpha}} \\ &- \lambda\alpha^{\frac{1}{1-\alpha}}[r+s+\lambda R(k)]^{\frac{\alpha}{1-\alpha}}[(1-\bar{\varepsilon})^2 - (1-R(k))^2] > 0 \end{aligned}$$

As $z < R(k)$ and $\bar{\varepsilon} > R(k)$, this inequality is true.

Proof of lemma 1

Lemma 1 *The investment decision condition of firms generates underinvestments ($k^I|_{R(k)=R^\star(k)} < k^\star$).*

Proof. Assume $R^I(k) = R^\star(k) = \overline{R(k)}$. Then, as $\frac{r+s+\lambda G(\overline{R(k)})+\gamma\theta^I q(\theta^I)}{1-\gamma} > r+s+\lambda G(\overline{R(k)})$, (3.4c) and (3.6c) imply that

$$\frac{r+s+\lambda G(\overline{R(k)}) + \gamma\theta^I q(\theta^I)}{1-\gamma} \frac{C'(k^I)}{y'(k^I)} = r+s+\lambda G(\overline{R(k)}) \frac{C'(k^\star)}{y'(k^\star)}$$

Therefore, it must be that $C'(k^I)$ is smaller than $C'(k^\star)$), and thus $k^\star$ exceeds k^I.

Proof of lemma 2

Lemma 2 *Under the Hosios condition, the productivity reservation of firms generates too many job destructions* ($R^I(k)|_{k=k^\star,\theta=\theta^\star} > R^\star(k)$).

Proof. Assume $k^I = k^\star = \overline{k}$ and $\theta^I = \theta^\star = \overline{\theta}$. Then, as $y(\overline{k}) + z + \frac{\gamma}{1-\gamma}c\theta + \frac{\gamma}{1-\gamma}\theta q(\theta)C(\overline{k}) > y(\overline{k}) + z + \frac{\eta(\theta)}{1-\eta(\theta)}c\theta$, (3.4b) and (3.6b) imply that:

$$R(k^I) + \frac{\lambda}{r+\lambda+s}\int_{R(k)}^{\overline{\varepsilon}}[1-G(x)]dx - \left[y(\overline{k}) + z + \frac{\gamma}{1-\gamma}c\theta + \frac{\gamma}{1-\gamma}\theta q(\theta)C(\overline{k})\right]$$
$$= R(k^\star) + \frac{\lambda}{r+\lambda+s}\int_{R^\star(k)}^{\overline{\varepsilon}}[1-G(x)]dx - \left[y(\overline{k}) + z + \frac{\eta(\theta)}{1-\eta(\theta)}c\theta\right]$$

Therefore, $R(k^\star)$ must be smaller than $R(k^I)$ to ensure this equality.

Proof of lemma 3

Lemma 3 *Under the Hosios condition, the free entry condition generates too small labor market tightness* ($\theta^I(k)|_{k=k^\star,R(k)=R^\star(k)} < \theta^\star(k)$).

Proof. Assume $k^I = k^\star = \overline{k}$ and $R^I(k) = R^\star(k) = \overline{R(k)}$. Then, as $\left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - \overline{R(k)}] - C(\overline{k}) < \left(\frac{1-\eta(\theta)}{r+\lambda+s}\right)[\overline{\varepsilon} - \overline{R(k)}] - (1-\eta(\theta))C(\overline{k})$, (3.4a) and (3.6a) imply that

$$\left[\left(\frac{1-\gamma}{r+\lambda+s}\right)[\overline{\varepsilon} - \overline{R(k)}] - C(\overline{k})\right]\frac{q(\theta^I)}{c} = \left[\left(\frac{1-\eta(\theta)}{r+\lambda+s}\right)[\overline{\varepsilon} - \overline{R(k)}] - (1-\eta(\theta))C(\overline{k})\right]\frac{q(\theta^\star)}{c}$$

Therefore, θ^I is smaller than $\theta^\star$ when $\gamma = \eta(\theta)$.

Chapter 4

Endogenous Job Destructions and the Distribution of Wages[1]

This paper considers a matching model with both idiosyncratic productivity shocks that hit jobs at random and heterogeneity of workers according to *ex ante* unobservable abilities. We argue that firms' decisions about reservation productivity can help explain the shape of wage distributions. This is shown from numerical experiments, calibrated to French data, by considering alternative ranges of productivity shocks.

JEL Classification: J31, J63

Keywords: Wage dispersion, job destruction, workers' heterogeneity

[1]This chapter reviews a joint work with Arnaud Chéron, forthcoming in *Labour Economics* (2011). I am grateful to David Jaeger (the Editor) and two anonymous referees for thoughtful comments and suggestions. I have also received helpful comments on earlier drafts of this article from participants at the PET Conference 2009, EEA Congress 2009, EALE Conference 2009 and Journées LAGV 2010, with a particular mention to Bruno Decreuse.

4.1 Introduction

The theory of equilibrium unemployment with matching and endogenous job destructions (Mortensen and Pissarides (1994)) has become an extensively-used framework both to address empirical facts of the labor market dynamics and to provide important insights into the design of labor market policies. Despite recent debates about the empirical relevance of the Nash-bargaining of wages (see Shimer (005a) and Hall (005a)), this framework undoubtedly helps explain stylized facts characterizing labor market flows (Cole and Rogerson (1999)), unemployment dynamics (Pissarides (2009)) and real business cycle features (Andolfatto (1996), Merz (1995) or Chéron and Langot (2004)). This framework is also well suited to show how employment protection, hiring subsidies or labor taxes can be used to improve welfare (see among many others Millard (1996), Mortensen and Pissarides (1999), Pissarides (2000) or more recently Chéron, Hairault and Langot (2011)).

Since the end of the 1990s, another strand of the search-matching literature has focused on wage dispersion, considering on-the-job search rather than endogenous firing decisions. Burdett and Mortensen (1998) stressed the role of search frictions within an on-the-job search background in generating wage dispersion despite having homogenous workers and firms. Subsequent work by Bontemps, Robin and van der Berg (1999), Bontemps, Robin and van der Berg (2000), Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006) mainly emphasized the effect of market frictions in combination with heterogenous productivities of both jobs and workers' abilities as a way to fit the distribution of wages. Wage dispersion is usually assumed to arise from on-the-job-search and existing models are usually characterized by exogenous firings. In contrast, this paper aims at showing that firms' firing decisions in the context of idiosyncratic productivity shocks can help explain the shape of the wage distribution.

Our paper also makes an empirical contribution. While it is widely known that OECD countries are characterized by hump-shaped wage distributions (see figure 3b in Jolivet, Postel-Vinay and Robin (2006)), little attention has been paid to the shape of employment to unemployment transition rates according to workers' position in the wage distribution, and therefore to the potential implication of those transitions in explaining wage dispersion. To give further insights into this issue, we consider the French experience, which shows recurrent stylized facts at the aggregate level as well as inside skill groups: a log-normal-like shape of wage distributions and a negative relationship between employment to unemployment transition rates and wage deciles (see section 4.3).

Drawing a parallel between these empirical observations serves as guideline for the construction of a simple labor market model which has to deliver such quantitative outcomes. In particular, we consider a job creation-job destruction model in line with Mortensen and Pissarides (1994) extended to account for heterogenous workers with *ex ante* unobservable abilities. There are two kinds of heterogeneity into the model: (i) each firm-job pair is hit by idiosyncratic productivity shocks, and (ii) each worker differs according to her ability. As the latter is assumed to be *ex ante* unobservable by firms, we consider non-directed search[2]. The lower bound of a productivity shock below which a job is closed down is determined by a reservation productivity of firms that obviously depends on worker'ability, only observed *ex post*. This endogenous reservation productivity is the key decision in the model since it determines how the combination of the exogenous distributions of shocks and abilities leads to generate the endogenous distribution of wages. We use numerical experiments to emphasize the explanatory power of firms' reservation productivity decision rule. We then argue that the model can mimic the decreasing relationship between employment to unemployment transition rates and wage deciles, and this helps explain the shape of wage distributions, as long as we allow for a sufficiently high range of productivity shocks. This result occurs despite considering a conventional Pareto distribution of abilities that would unambiguously imply strictly decreasing wage densities in case of exogenous separation rates.

The remaining of the paper is organized as follows. The next section presents the framework. The third one gives a description of the data set and computed statistics. The fourth section deals with computation experiments. The last section concludes.

[2]The introduction of workers heterogeneity in a matching model with non-directed search clearly raises (in)efficiency issues: the Hosios condition no longer achieves efficiency. Such theoretical issues have been examined by Shimer and Smith (2001), Albrecht and Vroman (2002), Blažquez and Jansen (2008) anfd Chéron, Hairault and Langot (2011) for instance. Yet, as our model does not add any new interesting insights about that point, and because our focus is above all related to the shape of wage distributions, efficiency issues are not addressed in this paper.

4.2 Model

4.2.1 Assumptions and labor market flows

We consider a continuous-time matching model in steady state with endogenous job creations and destructions. Workers are heterogenous due to unobservable ability a along the interval $[\underline{a}, \overline{a}]$, with $F(a)$ the exogenous cumulative distribution function of abilities. When a firm opens a job vacancy, it knows the distribution function of abilities but does not *ex ante* observe the ability of the contacted worker. This ability is revealed once the worker has been hired.

Each firm has one job. The productivity of the job/firm depends not only on the worker's ability, but also on a job specific random component. The idiosyncratic productivity shock, denoted ε, is realized at the time of job creation and occurs according a Poisson rate λ where $G(\varepsilon)$ is the cdf, $\forall \ \ \varepsilon \in [\underline{\varepsilon}, 1]$[3]. The overall productivity of the job is therefore given by $\varepsilon + a$. Further, the productivity threshold $R(a)$ determines the lowest productivity value a firm will accept to maintain a job. This value obviously depend on workers' ability. Accordingly, the overall job destruction rate is given by $\lambda G(R(a)) + s$, where s is an exogenous rate of separation[4].

Following Dolado, Jansen and Jimeno (2009) or Chéron, Hairault and Langot (2011) among others, we consider heterogenous workers in the context of a non-directed search process[5]. More precisely, we assume that firms cannot *ex ante* direct their search toward (unobservable) workers' ability. An aggregate matching function $M(v, u)$ then determines the number of hirings, where v and u denote the number of vacancies and unemployed workers, respectively. The matching function is increasing and concave in both arguments. Accordingly, the contact rate for each worker is given by $\theta q(\theta) \equiv \frac{M(v,u)}{u}$, where $q(\theta) \equiv \frac{M(v,u)}{v}$, and $\theta \equiv \frac{v}{u}$ is the labor market tightness. The transition rate from unemployment to employment for a worker with ability a is therefore given by $\theta q(\theta)[1 - G(R(a))]$.

Lastly, denoting $u(a)$ the number of unemployed workers with ability a and defining $f(a) \equiv F'(a)$, equilibrium labor market flows in steady state imply:

[3]Alternative values of $\underline{\varepsilon}$ will be considered in numerical experiments in section 4.4 to highlight the role of endogenous reservation productivity in explaining the shape of the wage distribution.

[4]Introducing s and letting it vary from zero to positive values will allow us in section 4.4 to assess the quantitative importance of endogenous job destructions. This indeed allows for a positive unemployment rate when $\underline{\varepsilon} = 1$.

[5]Chéron et al. (2011) consider age-differentiated workers but age discrimination is not allowed, while there are two types of workers differentiated by their level of education in the economy of Dolado et al..

$$u(a)\theta q(\theta) = [s + \lambda G(R(a))]\,[f(a) - u(a)] \quad \forall a \in [\underline{a}, \overline{a}]$$

The overall unemployment rate is written: $u = \int_{\underline{a}}^{\bar{a}} u(a)da$.

4.2.2 Firing and hiring behaviors

The value of a filled job is assumed to be defined by:

$$rJ(a,\varepsilon) = a + \varepsilon - w(a,\varepsilon) + \lambda \int_{R(a)}^{1} J(a,x)dG(x) - (s+\lambda)\,J(a,\varepsilon)$$

with r the interest rate and where $w(a,\varepsilon)$ stands for the wage.

It is not in the best interest of firms to keep workers on working if the job value is negative, i.e. $J(a,\varepsilon) \leq 0$.[6] Therefore, the threshold value for productivity $R(a)$ satisfies $J(a, R(a)) = 0$ and is positively related to wages but negatively to the labor hoarding value of the job since a new productivity is drawn at rate λ from the set $[\underline{\epsilon}, 1]$:

$$R(a) = -a + w(a, R(a)) - \lambda \int_{R(a)}^{1} J(a,x)dG(x) \tag{4.1}$$

The recruiting policy is determined by the expected average value of the job once filled. But as firms cannot *ex ante* target hirings among heterogenous workers and as a particular productivity shock is drawn once a worker is contacted, the vacancy decision depends both on abilities' and on productivity shocks' distributions. The value of a vacancy is therefore defined as follows:

$$rV = -c + q(\theta) \int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right) \left\{ \int_{R(a)}^{1} [J(a,\varepsilon) - V]\, dG(\varepsilon) \right\} da$$

with $c \geq 0$ the flow cost of recruiting a worker.

A standard free entry condition (such that the value of vacancies vanishes in equilibrium) then determines the labor market tightness θ and implies that the expected recruitment cost equalizes *ex ante* the expected value of job creation:

[6] Actually, due to Nash bargaining of wages, this separation rule is also optimal from workers' point of view.

$$\frac{c}{q(\theta)} = \int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right) \left\{ \int_{R(a)}^{1} J(a,\varepsilon) dG(\varepsilon) \right\} da \tag{4.2}$$

4.2.3 Wage setting

We consider the conventional assumption of Nash-bargaining of wages[7]. Firms and workers share the global surplus generated by a job according to their relative bargaining power: $S(a,\varepsilon) = J(a,\varepsilon)+W(a,\varepsilon)-U(a)$, where workers' values of unemployment and employment are respectively given by:

$$\begin{aligned} rU(a) &= z + \theta q(\theta) \int_{R(a)}^{1} \left[W(a,x) - U(a)\right] dG(x) \\ rW(a,\varepsilon) &= w(a,\varepsilon) + \lambda \int_{R(a)}^{1} \left[W(a,x) - W(a,\varepsilon)\right] dG(x) \\ &\quad + \left[s + \lambda G(R(a))\right] \left[U(a) - W(a,\varepsilon)\right] \end{aligned}$$

The standard Nash-sharing rule is:

$$W(a,\varepsilon) - U(a) = \gamma S(a,\varepsilon)$$

where γ stands for the bargaining power of workers. The following expression for the wage can be then derived[8]:

$$w(a,\varepsilon) = (1-\gamma)z + \gamma(a+\varepsilon) + \gamma\theta q(\theta) \int_{R(a)}^{1} J(a,x) dG(x) \tag{4.3}$$

In the context of *ex ante* unobservable heterogeneity and following Chéron et al (2011),

[7]Since Shimer (005a) and Hall and Milgrom (2008), the Nash-bargaining of wages is somewhat a disputed assumption, at least from an empirical perspective. Hall and Milgrom (2008) point out that the rigidity of wages helps explain the observed volatility of unemployment over the business cycle. Nevertheless, Pissarides (2009) rehabilitates the Nash-bargaining showing that the failure of the Mortensen-Pissarides' framework rather relies on the size of labor turnover costs which are typically understated.

[8]Depending on productivity draws, wage earnings may increase or decrease at each period. In France, approximately 40% of workers experience a fall in their real wages from one year to another.

we show that the way search costs enter into the wage equation depends on the *ex post* value of the worker relative to the *ex ante* expected average value of job creation, defined over the whole pool of unemployed workers[9]. More precisely, making use of equation (4.2) and defining $\tau(a) = \frac{\int_{R(a)}^{1} J(a,x)dG(x)}{\int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right) \int_{R(a)}^{1} J(a,x)dG(x)da}$, the wage expression can be rewritten as:

$$w(a, \varepsilon) = \gamma \left[a + \varepsilon + c\theta\tau(a)\right] + (1 - \gamma)z \tag{4.4}$$

High-ability workers are characterized by $\tau(a) > 1$, which implies that they are rewarded for more than the saving of the average search costs ($c\theta$). Both productivity a and search costs $c\theta\tau(a)$ then push up wages. Conversely, workers with low abilities earn low wages not only due to a lower productivity value (for a given ε) but also because of a lower imputed value of search costs. Therefore, the density of low-wages should be high if the density of low-ability workers is high. On the other hand, a low density of high-ability workers should translate into low densities of high-wages.

Our model also implies, however, that heterogenous workers may earn the same wage: a low-ability worker who would have been hit by a good shock may earn the same wage as a high ability one but who would have been hit by a bad shock. This introduces a new mechanism with regard to the wage distribution that is related to the endogenous productivity threshold $R(a)$, which tends to decrease the density of low wages but to increase the one of medium wages. Typically, the reservation productivity is high (low) for low(high)-ability workers (see property 2 below). Accordingly, the set of wages, defined over $w(a, 1) - w(a, R(a))$, is narrower for low-ability workers while it could be very large for high-ability workers, as firms may have interest to keep them on working in case of very bad shock.[10] All else being equal, both high productivity thresholds for low-ability workers and low productivity thresholds for high-ability workers should

[9] In Chéron et al (2011), the role search costs play into the wage equation is emphasized considering age-differentiated workers.

[10] Considering heterogenous workers in terms of education instead of ability, Dolado, Jansen and Jimeno (2009) argue that the set of wages could also be very large for high-educated workers. Their mechanism rests on on-the-job search in a context of heterogenous firms (creating simple and complex jobs) and exogenous job destruction. While high-skill workers optimally accept simple rather than complex job, firms accept these matches at a lower wage than with appropriately matched workers (low-skill workers in simple jobs) as they anticipate that over-educated candidates may quit as soon as a better job becomes available. This gives rise to strong wage inequalities among those workers, depending on the job they work in. In our model, wage inequalities are also strong among high-abilities workers as they depend on idiosyncratic productivity shocks that can be bad or good.

therefore contribute to increase the density of medium wages.

4.2.4 Equilibrium definition

The conditions that simultaneously determine the labor market tightness θ, the set of productivity thresholds $R(a)$ and unemployment levels by ability $u(a)$, $\forall a \in [\underline{a}, \overline{a}]$ can be now defined:

Proposition 5. *The labor market equilibrium is defined by the following set of equations:*

$$\frac{c}{q(\theta)} = \left(\frac{1-\gamma}{r+s+\lambda}\right)\int_{\underline{a}}^{\overline{a}}\left(\frac{u(a)}{u}\right)\int_{R(a)}^{1}[1-G(\varepsilon)]\,d\varepsilon da$$

$$R(a) = -a+z-\left(\frac{\lambda-\gamma\theta q(\theta)}{r+s+\lambda}\right)\int_{R(a)}^{1}[1-G(\varepsilon)]\,d\varepsilon$$

$$u(a) = f(a)\frac{s+\lambda G(R(a))}{\theta q(\theta)+s+\lambda G(R(a))} \quad ; \quad u=\int_{\underline{a}}^{\bar{a}}u(a)da$$

Proof. First, from $J(a,R(a))=0$ and the wage expression (5.3) and using the fact that $(r+s+\lambda)J(a,\varepsilon)-(r+s+\lambda)J(a,R(a))=\varepsilon-R(a)-w(a,\varepsilon)+w(a,R(a))$, it follows that $(r+s+\lambda)J(a,\varepsilon)=(1-\gamma)[\varepsilon-R(a)]$. Second, integrating by parts leads to $\int_{R(a)}^{1}[\varepsilon-R(a)]\,dG(\varepsilon)=\int_{R(a)}^{1}[1-G(\varepsilon)]\,d\varepsilon$. □

Property 2. *The labor market equilibrium is characterized by $R'(a)<0$.*

Proof. The proof is straightforward by noticing that $dR=-da+\left(\frac{\lambda-\gamma\theta q(\theta)}{r+s+\lambda}\right)[1-G(R)]\,dR$, which implies that $\frac{dR}{da}\equiv R'(a)=-\frac{1}{1-\left(\frac{\lambda-\gamma\theta q(\theta)}{r+s+\lambda}\right)[1-G(R)]}<0$, $\forall\lambda,\gamma$. □

According to this property, the higher the worker's ability, the lower the productivity threshold below which the job is not maintained. This suggests, therefore, that high-ability workers may keep their jobs even though bad productivity shocks hit them.

4.3 Data

To examine the correlation between the employment to unemployment transition rates and wages, we used French Labor Force survey ("*Enquête Emploi*" provided by INSEE) in 1992, 1997, 2002 and 2007. For each year, we defined our sample in the following way. We focused on the population of respondents who were working at the beginning of the first quarter and considered their situation on the labor market at the end of the quarter. We chose to select the subsample of workers aged from 18 and 60, working full-time or part-time jobs and employed by the private sector. We exclude farmers and self-employed. We also deleted the few observations with missing values, mainly because of missing wages. Lastly, workers were sorted according to their socioeconomic status at the beginning of the quarter. In particular, we defined four groups of workers according their skill level: high-skilled workers (executives and managers), medium-skilled workers (technical supervisors and technicians), low-skilled workers (skilled manual workers) and unskilled workers (unskilled workers and employees).

We focused on the two following variables of interest. The first one was about transitions from employment to unemployment. Therefore, we defined a dummy variable which was equal to one when the worker has experienced a transition from employment to unemployment between the beginning and the end of the first quarter of each year under consideration. The second outcome was the monthly wage level, expressed in euros. Wages were divided into ten intervals computed from nine wage deciles. For each skill level, workers were then sorted according to the wage interval they belong to at the beginning of the quarter. We computed then quarterly employment-unemployment transition rates both by skill level and wage interval.

Firstly, figure 4.1 below shows the log-normal-like shape of wage distributions within skill groups for 2007, which suggests that France can truly represent what is observed in most of OECD countries. Figure 4.2 draws the same picture for 1992-2007. Ratios by skill of wage deciles to the median wage are also provided in table 4.1.[11]

Secondly, figure 4.3 shows that France is also characterized by a negative correlation between wage deciles and transition rates from employment to unemployment within each wage interval, both at the aggregate level and by skills. Figure 4.4 emphasizes also that such a decreasing relationship holds over the 1992-2007 period. We do not report statistics related to high-skilled workers as managers experience only very few employment to unemployment transitions in the French LFS (at least for high wage deciles). Furthermore, this negative slope is all the more important that wages below the fifth decile are considered.

[11] Quantitative properties of our model will be compared to those statistics.

Figure 4.1: Wage distributions in 2007 French LFS, by skill

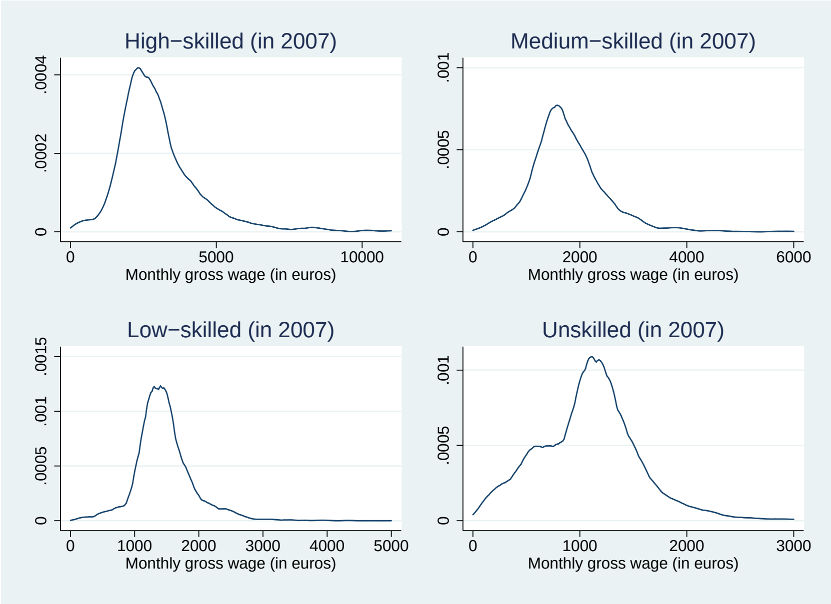


Figure 4.2: Wage densities by skill in 1992, 1997, 2002 and 2007 (in France)

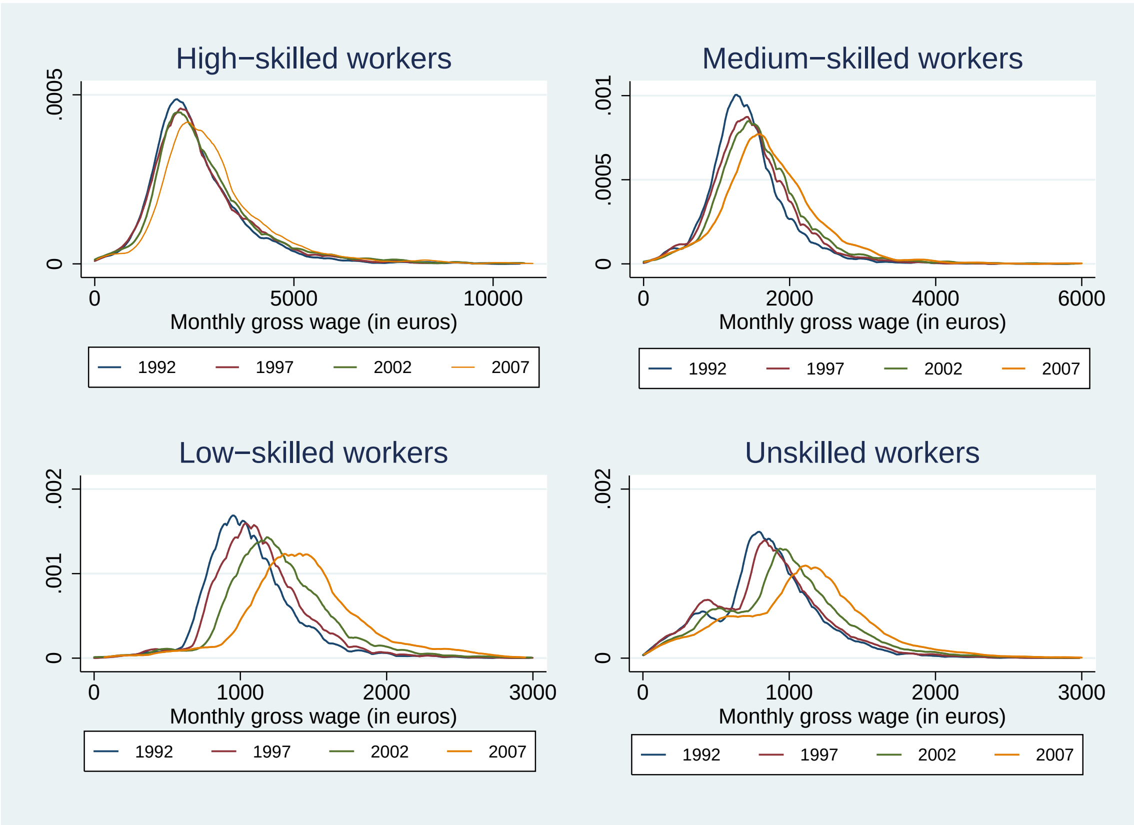


Table 4.1: Ratios by skill of wage deciles to the median wage (in France)*

Workers	D5/ average wage	D1/ D5	D2/ D5	D3/ D5	D4/ D5	D5/ D5	D6/ D5	D7/ D5	D8/ D5	D9/ D5
High-skilled										
In 1992	0.89	0.6	0.73	0.83	0.92	1	1.11	1.25	1.43	1.78
In 1997	0.89	0.59	0.72	0.83	0.91	1	1.10	1.27	1.46	1.80
In 2002	0.87	0.62	0.73	0.82	0.92	1	1.12	1.25	1.47	1.85
In 2007	0.90	0.62	0.74	0.83	0.91	1	1.10	0.94	1.42	1.73
Medium-skilled										
In 1992	0.94	0.64	0.78	0.85	0.93	1	1.08	1.17	1.30	1.53
In 1997	0.95	0.61	0.74	0.84	0.91	1	1.08	1.18	1.31	1.53
In 2002	0.92	0.63	0.76	0.85	0.93	1	1.08	1.20	1.31	1.57
In 2007	0.96	0.61	0.75	0.85	0.92	1	1.08	1.10	1.31	1.53
Low-skilled										
In 1992	0.94	0.74	0.82	0.89	0.94	1	1.06	1.13	1.24	1.42
In 1997	0.96	0.72	0.83	0.89	0.95	1	1.06	1.13	1.24	1.39
In 2002	0.95	0.73	0.81	0.88	0.95	1	1.07	1.15	1.25	1.41
In 2007	0.96	0.73	0.84	0.89	0.94	1	1.07	1.13	1.25	1.40
Unskilled										
In 1992	0.97	0.44	0.68	0.83	0.91	1	1.09	1.18	1.32	1.55
In 1997	0.97	0.44	0.61	0.80	0.93	1	1.09	1.20	1.35	1.59
In 2002	0.97	0.44	0.62	0.80	0.93	1	1.09	1.18	1.33	1.55
In 2007	1	0.43	0.61	0.79	0.91	1	1.09	1.18	1.30	1.49

* The first column "D5/average wage" gives the ratio of the fifth wage decile to the average wage.

Figure 4.3: Employment to unemployment transition (E− >U) rates in French LFS

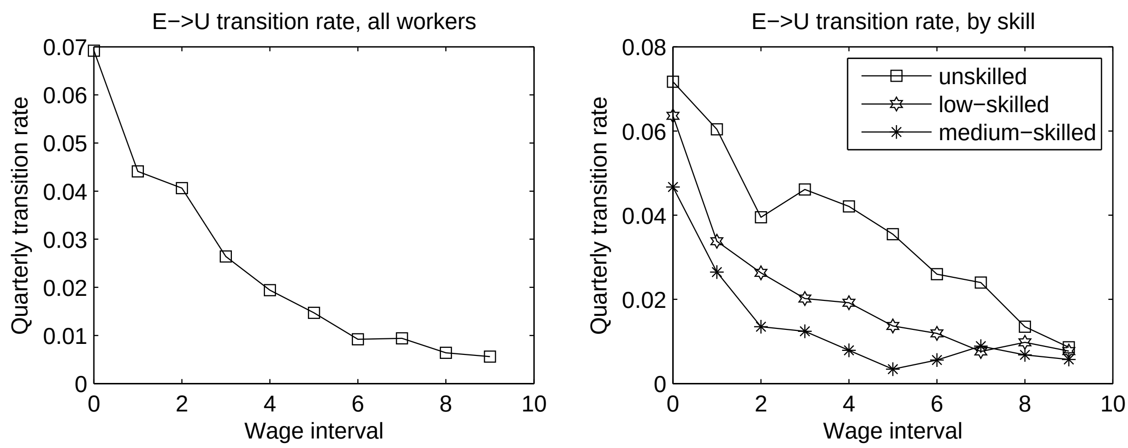
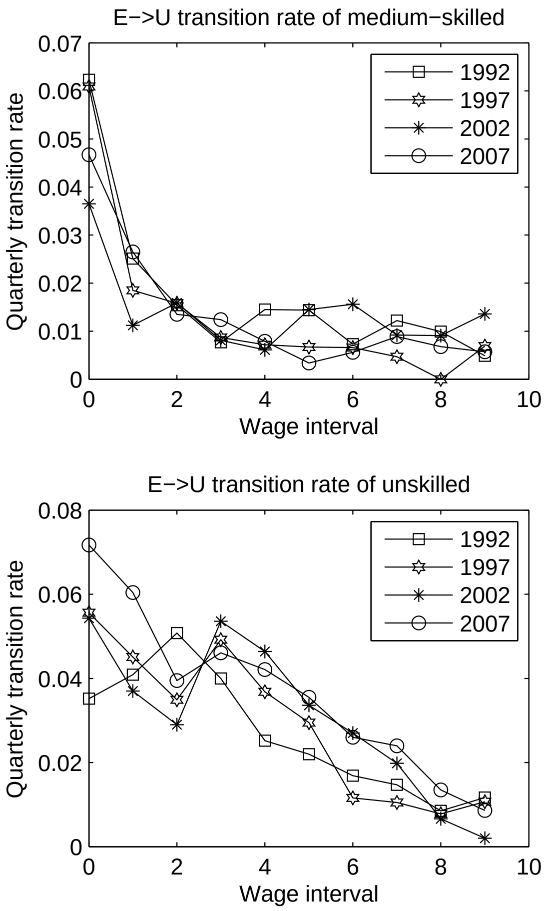


Figure 4.4: Employment to unemployment (E− >U) transition rates by skill in 1992, 1997, 2002 and 2007 (in France)



The next section aims at enlightening the potential empirical relevance of combining productivity shocks with *ex ante* unobservable heterogeneity of workers as a way to explain the distribution of wages. Obviously, our approach leaves aside several dimensions of the labor market that have turned out to be important determinants of the wage distribution, such as on-the-job search (see *e.g.* Bontemps, Robin and van der Berg (2000), Postel-Vinay and Robin (2002)). However, our objective is to focus on a particular context that has not been yet examined -heterogenous unobservable abilities of workers interacting with idiosyncratic productivity shocks- and to show that firms' decision rule about productivity can help explain the shape of the wage distribution.

4.4 Simulations of the equilibrium wage distribution

Our overall strategy consists in showing how sensitive is the wage distribution to endogenous productivity thresholds, by letting the lower bound of productivity shocks $\underline{\varepsilon}$ vary. An important implication of our model is that the distribution of wages depends on endogenous productivity thresholds. Accordingly, the shape of wage dispersion can be dissimilar to the distribution we typically assume for workers' abilities. This is clearly so in case of exogenous job destruction (i.e. $\underline{\varepsilon} = 1$ and $s > 0$).

We briefly present the model calibration, discuss the computation of wage distributions and finally examine some numerical experiments.

4.4.1 Calibration

We consider a quarterly calibration of the model. A first set of parameters is based on external information. A second one aims at replicating some stylized facts that characterize the French low-skilled workers ("ouvriers qualifiés") data set over the 2006-2007 period. We consider a homogenous period of the business cycle (before the current economic crisis) as our model does not allow for macroeconomic shocks and fluctuations.[12]. As detailed below, we consider two targets: the unemployment rate and the employment to unemployment exit rate over this period.

[12]Postel-Vinay and Robin (2002) and Jolivet, Postel-Vinay and Robin (2006) (among many others) follow a similar strategy to estimate their steady-state search model.

The empirical investigation of the quantitative properties of the model then consists in examining *(i)* the shape of the wage distribution, and *(ii)* statistics summarizing wage deciles and average separation rates within each wage decile interval. We consider three specifications of productivity shocks that allows us to analyze the sensitivity of the quantitative properties of the model.

As a preliminary step, specifications of functional forms for the matching function and the distributions of idiosyncratic shocks and abilities are required. We choose the simplest functions based on existing assumptions in the literature. In particular, as in Mortensen and Pissarides (1994), we consider a uniform distribution of shocks $G(x) = x \ \forall x \in [\underline{\varepsilon}, 1]$ and a Cobb-Douglas matching function $M(v,u) = v^{\psi}u^{1-\psi}$. We also follow Mortensen (2003) by assuming that the dispersion of abilities is defined by a Pareto distribution. More specifically, we assume that $F(a) = 2\left(1 - \frac{1}{a}\right) \ \forall a \in [1, 2]$.[13]

These specifications imply that the equilibrium conditions collapse to:

$$c\theta^{1-\psi} = \left(\frac{1-\gamma}{r+s+\lambda}\right)\int_1^2 \left(\frac{u(a)}{u}\right)\frac{1}{2(1-\underline{\varepsilon})}\left[1-R(a)\right]^2 da \qquad (4.5)$$

$$R(a) = -a + z - \left(\frac{\lambda - \gamma\theta^{\psi}}{r+s+\lambda}\right)\frac{1}{2(1-\underline{\varepsilon})}\left[1-R(a)\right]^2 \qquad (4.6)$$

$$u(a) = \left(\frac{2}{a^2}\right)\left(\frac{s+\lambda G(R(a))}{\theta^{\psi}+s+\lambda G(R(a))}\right) \quad ; \quad u = \int_1^2 u(a)da \qquad (4.7)$$

The first set of parameters is consistent with conventional values assumed in the literature: $r = 0.01$ and $\psi = \gamma = 0.5$. The second set of parameters then includes $\{z, c, s, \lambda, \underline{\varepsilon}\}$. The role of productivity shocks and endogenous productivity threshold crucially depends on the value of $\underline{\varepsilon}$. Therefore, we examine three model specifications to underline mechanisms at work:

- "Model 1" is the benchmark model with endogenous job destruction but without any exogenous job destruction. Hence, we set $s = 0$. Following Ljungqvist and Sargent (2008), the Poisson arrival rate of productivity shocks is consistent with an expected shock every two years, i.e. $\lambda = 1/8$. The support of productivity shocks is assumed to be continuously distributed over the range $[-1, 1]$ by setting $\underline{\varepsilon} = -1$. Accordingly, the lowest produc-

[13] Alternative calibrations of $\underline{a}$ and $\underline{a}$ would lead to similar quantitative conclusions. As we are interested in the shape of wage distributions, wage levels do not matter. Hence, we will report simulation results by dividing wages by the lowest wage in the economy.

tivity $(a+\varepsilon)$ is zero in this economy and the highest one is 3 as $a \in [1,2]$. We choose $z = 0.52$ to get an average job destruction rate of 2.1% per quarter, consistent with the observed employment to unemployment exit rate for the skilled manual workers over the period 2006-2007 in France. The labor market tightness θ should be consistent with an unemployment rate of 8.2% (from equation (4.7)), which gives the value of c as a solution of equation (4.5).[14]. This implies that the contact rate is $\theta^{\psi} = 35\%$, which means that it takes on average 8.5 months to get a job offer. The simulated average duration of an unemployment spell is then approximately 11 months.

To test the sensitivity of the distributional implications of the model, we then consider two other calibrations:

- "Model 2" refers to the case of exogenous job destruction where it is assumed that $\underline{\varepsilon} = 1$. There are no productivity shocks in this economy. Instead of equations (4.5)-(4.7), the equilibrium is characterized by:[15]

$$c\theta^{1-\psi} = \int_1^2 \left(\frac{u(a)}{u}\right)\left(\frac{(1-\gamma)(a+1-z)-\gamma c\theta}{r+s}\right) da$$

$$u(a) = \left(\frac{2}{a^2}\right)\left(\frac{s}{\theta^{\psi}+s}\right) \quad ; \quad u = \int_1^2 u(a)da$$

  As we aim at comparing Model 2 with Model 1, home production is still set to $z = 0.52$, and we assume $s = 2.1\%$ to match the average separation rate of low-skilled workers. θ (hence c) is also set to match the unemployment rate (8.2%), which is now consistent with an unemployment spell of about 13 months for each ability.

- "Model 3" is a mixed version of Model 1 and Model 2. It allows both for exogenous and endogenous productivity shocks but a smaller rate of exogenous job destruction is chosen: $s = 1\%$. Still, $\lambda = 1/8$ (as in Model 1) and we assume $\underline{\varepsilon} = 0$. Both remaining parameters, z and θ (hence c), are set to fit the average exit rate from employment to unemployment

[14] Actually, only the steady-state value of θ^{ψ} matters in our simulation procedure since equations (4.6)-(4.7) for the productivity threshold and the unemployment, as well as upcoming equation (4.8) for the wage, are expressed as functions of θ^{ψ}. c then satisfies equation (4.7) for a given θ.

[15] With only exogenous job destruction, $c\theta^{1-\psi} = J(a,1)$ at equilibrium, where $J(a,1) = \frac{a+1-w(a)}{r+s}$ and $w(a) = \gamma(1+a+c\theta)+(1-\gamma)z$.

and the unemployment rate. This leads to $z = 1.29$, and the average duration of an unemployment spell is approximately 11 months.

4.4.2 Computation of wage distribution

Without any productivity shock (Model 2), the distribution of wages can be straightforwardly derived from that of abilities, as wages only depend on the ability a. More precisely, assuming (as in the calibration) $\varepsilon = 1$, we get from equation (5.3):[16]

$$\begin{aligned} w(a) &= (1-\gamma)z + \gamma(a+1) + \gamma\theta q(\theta)J(a,1) \\ &= \left(\frac{1-\gamma}{1+\gamma\Psi(\theta)}\right) z + \gamma\left(\frac{1+\Psi(\theta)}{1+\gamma\Psi(\theta)}\right)(a+1) \end{aligned}$$

with $\Psi(\theta) = \frac{\theta q(\theta)}{r+s}$ and making use of the fact that $J(a,1) = \frac{a+1-w(a)}{r+s}$. The corresponding density function of wages, denoted $\phi(w)$, is then given by:

$$\begin{aligned} \phi(w) = \phi(w(a)) &= f(a) - u(a) \\ &= f(a)\left(\frac{\theta^{\psi}}{s+\theta^{\psi}}\right) \end{aligned}$$

From the Pareto distribution, $f(a) \equiv F'(a) = \frac{2}{a^2}$, which leads to $f'(a) < 0$. Therefore, $\phi'(w) < 0$ unambiguously. But such a strictly decreasing shape of the wage density function is clearly at odds with the well-documented hump-shaped wage distributions.

On the other hand, in the benchmark economy, the wage earning of a worker is not only related to her ability but also to idiosyncratic shocks that hit the job. Formally, this results in $w = w(a, \varepsilon)$ defined by (5.3), which can also be rewritten as:

$$w(a,\varepsilon) = (1-\gamma)z + \gamma(a+\varepsilon) + \left(\frac{\gamma\theta^{\psi}}{r+\lambda+s}\right)\left(\frac{1-\gamma}{2(1-\underline{\varepsilon})}\right)[1-R(a)]^2 \qquad (4.8)$$

To compute wages densities, we then need to account for the endogenous job destruction

[16] This equation is the same as $w(a) = \gamma(a+1+c\theta) + (1-\gamma)z$.

decision as well. For instance, some high-ability workers may earn low wages at a certain point because their productivity threshold is low. Beyond this intuitive statement, the density of wages can be derived as follows:

$$\phi(w) \quad = \quad \int_1^2 \Psi(w,a)da \text{ with } \begin{cases} \Psi(w,a) = 0 \;\; \forall w < w(a, R(a)) \\ \Psi(w,a) = f(a) - u(a) \;\; \forall w \geq w(a, R(a)) \end{cases}$$

where we make use of the fact that the uniform distribution of shocks implies that the density of each productivity draw is unchanged all across the support of shocks.

4.4.3 Numerical experiments

This quantitative analysis first aims at showing how the shape of the wage distribution depends on endogenous reservation productivity. We then examine whether our model is able to produce realistic properties concerning both wage deciles and separation rates by wage interval.

Figures 4.5 and 4.6 below show the properties of the model. Figure 4.7 compares some statistics calculated from the simulated data of the model with the same statistics based on empirical data for the French low-skilled workers segment.

Figure 4.5 focuses on job destruction rates. Benchmark Model 1 implies a decreasing relationship between the job destruction rate and worker ability: it starts with a quarterly rate of approximately 4% (per quarter) and falls to zero for workers whose ability is 1.9 times higher than the lowest one. Given the equilibrium distribution of abilities, around 5% of workers do not experience any employment to unemployment transitions. In Model 3, the lower bound of idiosyncratic shock is 0 instead of -1 in Model 1. Workers whose abilities are above 1.4 then no longer experience endogenous firing, but face the exogenous employment exit rate of 1%. This is due to the fact that their reservation productivity is limited by the lower bound 0. Lastly, the endogenous job destruction rate in Model 2 clearly does not depend at all on abilities.

As a combination of the exogenous distributions of abilities and productivity shocks, the wage

Figure 4.5: Model properties (I): reservation productivity and job destruction rate by ability

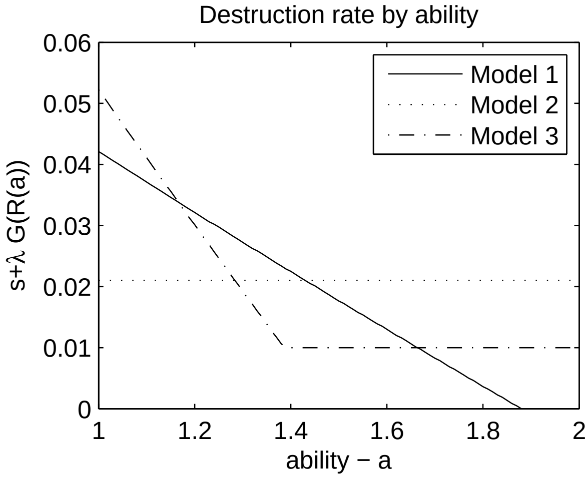
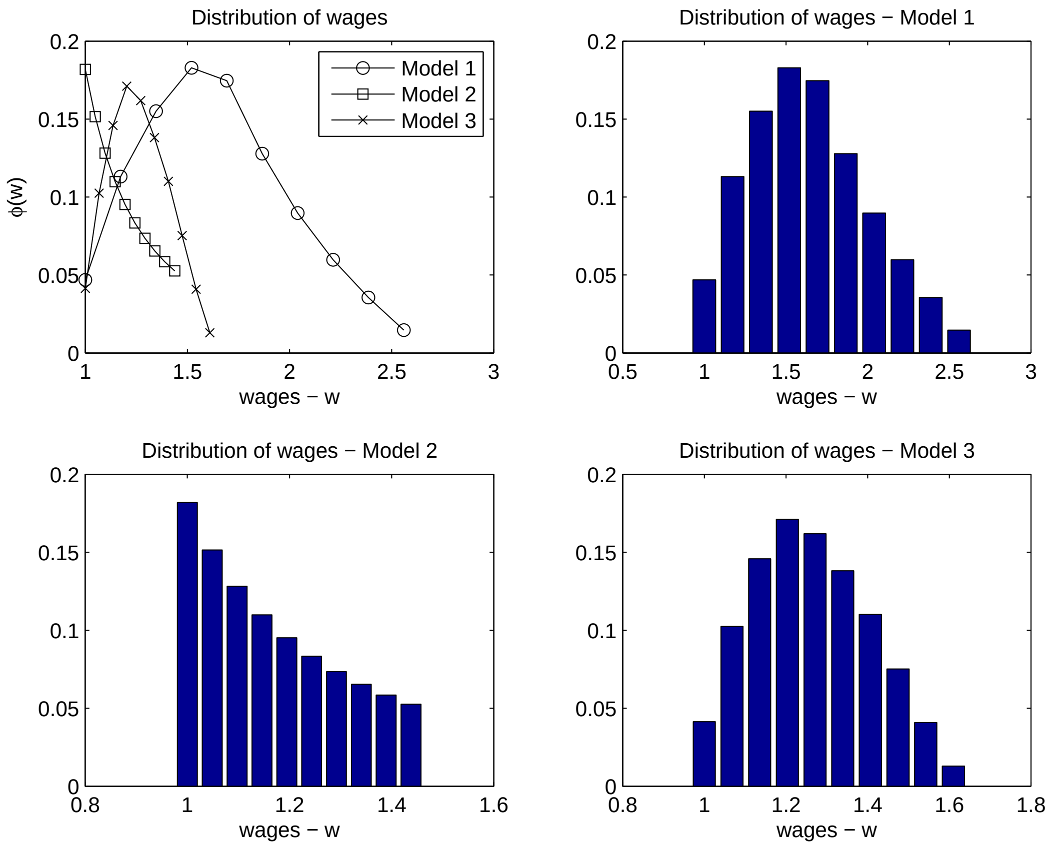


Figure 4.6: Model properties (II): the distribution of wages



distribution is determined by firms' decision rule about reservation productivity. In figure 4.6[17], the most striking feature is that introducing idiosyncratic productivity shocks generates hump-shaped wage distributions (models 1 and 3). This first relies on the fact that the reservation productivity of low-ability workers is high, which implies that only low-ability workers who draw a good productivity are in a position to keep their job. All else being equal and compared with the case of exogenous separations, this raises the average wage of low-ability workers. Secondly, the reservation productivity of high-ability workers is low. Again, all else being equal, high-ability workers who have been hit by a bad productivity shock earn lower wages, which move them to the left in the wage distribution.

The magnitude of these mechanisms is all the more important when the gap between the lower bound of productivity shocks and the upper bound is large. This means that the potential role of firms' decisions is larger in Model 1. In Model 3, this gap is 1 (since $\underline{\varepsilon} = 0$) while it is 2 in Model 1 (since $\underline{\varepsilon} = -1$)[18]. Thus, reservation productivities for high abilities workers turn out to be smaller in Model 1 than the lower bound of productivity shocks in Model 3. This implies that the shift to the left of high ability workers in the wage distribution is much stronger in Model 1 than in Model 3.

Figure 4.7: Model assessment

Finally, we compare the models' implications to some statistics computed from the group of low-skilled (manual skilled) workers in France. Figure 4.7 reports two kinds of statistics. The

[17]For each model calibration, we divide wages by the lowest wage.

[18]Keep in mind that we assume an upper bound equal to 1.

panel on the left gives the value of each wage decile with respect to the median wage and the panel on the right gives the average job destruction rate according to the wage interval workers belong to. While other factors such as on-the-job search should improve our understanding of wage distributions, Model 1 performs surprisingly well. Ratios of wage deciles over the median wage are well-replicated, in particular deciles 1 to 4. In the panel on the right, Model 2, by definition, cannot account for the decreasing shape of employment exit rates by wage interval. Both Model 1 and 3 match the data relatively well and generate similar patterns. Therefore, there is no basic difference between Model 1 and 3 in terms of worker flows by wage interval. But Model 1 clearly does a better job in replicating ratios of wages deciles than Model 3 thanks to the difference between both models in the lower bound of of productivity shocks. As the latter is smaller in Model 1 (compared to Model 3), the mix of abilities actually rises within the wage distribution. This implies for instance that some-high ability workers could have much smaller wages in Model 1 than in Model 3, since their reservation productivity thresholds turn out to be lower than zero (the lower bound of shocks in Model 3).

Overall, although the performance of the model is not perfect, we think that these numerical experiments highlight the potential role of firms' decisions about reservation productivity in wage dispersion analysis.

4.5 Conclusion

The goal of this paper is to highlight the role of firms' decisions about reservation productivity (hence determining whether a job may be closed down) in the wage dispersion analysis. This has been neglected until now since existing models put the emphasis on on-the-job-search, and are usually characterized by exogenous firings. We have developed a matching model with endogenous job destructions (which implies endogenous reservation productivity) in combination with heterogenous workers. By letting the range of productivity shocks vary from zero, we showed that the model can generate a hump-shaped wage distribution.

We do not want to disregard or diminish the role the contribution of on-the-job search in analyzing wage dispersion. Rather, on-the-job search and endogenous firing decisions should be considered together in order to provide a good description of wage inequality. This gives a research agenda, as well as providing a decomposition of cross-employee wage variance in line with Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay and Robin (2006) who use matched employer-employee data.

Chapter 5

Training, job destruction and wage distribution[1]

This paper accounts for distributions of wages, job destruction rates and training investments at the same time. To that end, we consider a matching model with idiosyncratic productivity shocks, endogenous firm-specific investments in training and workers' heterogeneity according to *ex ante* unobservable abilities. Two sources of inefficiency arise in such a theoretical framework: a holdup problem and a composition externality in the search process. We examine the quantitative model properties, calibrated to French low-skilled workers, and then characterize the optimal labor market policy that leads to reach the efficient allocation.

JEL Classification: J31, J38, J41

Keywords: Training, job destruction, wage distribution

[1] I am really grateful to Arnaud Chéron for his suggestions that improved this work.

5.1 Introduction

Extended job search environments from Burdett and Mortensen (1998) with firm-specific training investments have been used to explore wage dispersion (Mortensen (2000), Rosholm and Svarer (2004) and Quercioli (2005) for instance)[2]. In particular, the equilibrium wage distribution obtained in this way can be hump-shaped, as in real data. As an extension of Burdett-Mortensen framework, these papers combine on-the-job search, exogenous firing decisions and homogenous agents. Endogenous firing decisions are put aside. Nevertheless, Chéron and Rouland (2011a) have recently highlighted the potential role of idiosyncratic productivity shocks and endogenous firing decisions in the wage dispersion analysis. Chéron and Rouland (2011b) also show that job destruction and training investment decisions are highly complementary. Firms have strong incentives to invest in training to protect matches from idiosyncratic productivity shocks. Expected productivity gains due to training investments rise the job tenure, which in turn encourages firms to invest more.

This paper extends these two contributions in order to deal both with positive and normative issues. In particular, we address heterogeneity issues about wages, employment to unemployment transitions and training, and wonder if there is room for labor market policy. To that end, we consider a matching model with endogenous job destructions, heterogenous workers due to *ex ante* unobservable abilities and where firms decide at job entry how much to invest in match-specific skills. This framework generates a wage distribution, transition rates from employment to unemployment and average training amounts per worker by wage interval. From numerical experiments calibrated and confronted to real French data, we show that this model is a credible framework. In particular, the wage distribution has a log-normal-like shape and job destruction rates and abilities are negatively related. The model also generates a strictly increasing correlation of the amount firms are willing to invest in specific-training with abilities. Further, our framework produces significant disparities between training amounts and between job destruction rates according to the wage interval a worker belongs to, as in real data. But, differences are not high enough. So, the model performance is far from perfect but could be improved considering on-the-job training and firms' heterogeneity as well.

[2] Fu (2011) considers firms' decisions on general human capital in a Burdett-Mortensen framework.

Considering firm-specific training investments in a labor market with frictions and unobservable heterogeneity across workers' abilities raises inefficiency issues. The Hosios condition no longer achieves efficiency as in the equilibrium unemployment benchmark. Therefore, there is room for labor market policies. First, the introduction of workers heterogeneity in a matching model with non-directed search -due to *ex ante* unobservable heterogeneity- implies a composition externality in the search process since the composition of the group of the unemployed has an effect on the average expected value of a contact. Such a theoretical issue has been examined by Shimer and Smith (2001), Albrecht and Vroman (2002), Blažquez and Jansen (2008) and Chéron, Hairault and Langot (2011) for instance. In our framework, the more unemployed workers with high abilities there are, the higher the probability to contact a high-ability worker, and hence the higher the expected return on a vacancy. Second, the other source of inefficiency in our framework comes from training investments that firms made before the wage bargaining[3], which comes to introduce a fixed job creation cost, fully lost if the wage negotiation fails. So, the ex-post wage bargaining process increases workers' bargaining power who takes advantage of the time to get a higher wage. Therefore, firms pay all the costs but cannot get all the returns on their investment, leading finally to a holdup problem[4]. Since both inefficiencies are not related, implications for policy are also different. Hence, as a second contribution of this paper, we characterize the optimal labor market policy by removing inefficiencies one after the other. Considering a first-best policy (in the sense that policy intruments could depend on ability levels), we show that abilities and optimal firing taxes and training subsidies that must be implemented are ambiguously related.

The remaining of the paper is organized as follows. The next section presents the theoretical framework. The quantitative analysis of the model is carried out then in the third section. The fourth section deals with inefficiencies issues. Finally, the last section concludes.

[3]We consider holdup issues to be in line with Chéron and Rouland (2011b).

[4]Jansen (2010) shows that competition among rival applicants may prevent hold-ups in markets with frictions when firms need to invest in capital before posting a vacancy.

5.2 Model

5.2.1 Model environment and labor market flows

We extend the framework used in Chéron and Rouland (2011a) to account for firms' decisions to engage in training. In particular, we consider a continuous-time matching model in steady state with endogenous job creations and destructions. Workers are heterogenous due to unobservable ability a along the interval $[\underline{a}, \overline{a}]$, with $F(a)$ the exogenous cumulative distribution function of abilities. Opening a job vacancy, the firm knows the distribution function of abilities but does not *ex ante* observe the ability of the contacted worker. This ability is revealed once the worker has been hired. In this way, as in Dolado, Jansen and Jimeno (2009) and Chéron, Hairault and Langot (2011), we consider heterogenous workers in the context of a non-directed search process. More precisely, we assume that firms cannot *ex ante* direct their search toward (unobservable) workers' ability. An aggregate matching function $M(v, u)$ then determines the number of hirings, where v and u denote the number of vacancies and unemployed workers, respectively. The matching function is increasing and concave in both arguments. Accordingly, the contact rate for each worker is given by $\theta q(\theta) \equiv \frac{M(v,u)}{u}$, where $q(\theta) \equiv \frac{M(v,u)}{v}$, and $\theta \equiv \frac{v}{u}$ is the labor market tightness.

A productive unit is the association of one worker and one firm. The productivity of the job/firm is the sum of a random component ε and a deterministic one $y(a, k(a))$, derived from a training investment in firm-specific skills that depends on the worker's ability a[5]. We assume that the human capital level of a worker is determined at the entry into the job and is constant for all the job tenure.

The time of events and of decisions is as follows. First, at the time of match formation, firms decide on the investment in firm-specific skills $k(a)$ that determines the human capital of the worker $y(a, k(a))$ as long as the job lasts[6]. Training investments increase the output of the worker only if she stays with the training firm. In this way, training is assumed to be specific in Becker's (1962) sense as it is fully lost on separation[7]. Second, an idiosyncratic productivity

[5]The additive form of the output of the match we assumed between an endogenous component ($y(k(a))$) and another exogenous one (ε) clearly simplifies calculations but also fits the usual definition of training. Usually, training is considered as a way to improve workers' skills. Without training, workers are still able to produce but at lower productivity levels. To mention only a few, Lechthaler (2009) and Belot, Boone and van Ours (2007) consider an additive form of the output of the match as well, within the framework of endogenous human capital and productivity shocks.

[6]The function $y(a, k(a))$ is supposed strictly increasing and concave, with $y(0) = 0$.

[7]Our choice of training modelling is supported by data. For instance, O'Connell (1999) reports that most of

shock ε is realized according to a Poisson rate λ where $G(\varepsilon)$ is the cdf, $\forall \;\; \varepsilon \in [0,1]$. Firms then decide to close down any jobs which productivity is below an (endogenous) productivity threshold denoted $R(k(a))$ that obviously depends on the amount the firm invested in training. Jobs can also be destroyed exogenously at rate s in the form of voluntary quits of workers[8]. Accordingly, $\lambda G(R(k(a)) + s$ gives the overall job destruction rate. The firm and its worker then bargain over the strating wage. Lastly, whenever an idiosyncratic shock arrives, the firm either accounts for this new value of ε in a new wage negotiation or destroy the job for a zero return.

The transition rate from unemployment to employment for a worker with ability a is given by $\theta q(\theta)[1 - G(R(k(a)))]$. Therefore, denoting $u(a)$ the number of unemployed workers with ability a and defining $f(a) \equiv F'(a)$, equilibrium labor market flows in steady state imply:

$$u(a)\theta q(\theta) = [s + \lambda G(R(k(a)))]\,[f(a) - u(a)] \quad \forall a \in [\underline{a}, \overline{a}]$$

The overall unemployment rate is written: $u = \int_{\underline{a}}^{\bar{a}} u(a)\mathrm{d}a$.

5.2.2 Firms' decisions

Hiring and firing decisions The value of a filled job is assumed to be defined by:

$$rJ(k(a),\varepsilon) = y(a,k(a)) + \varepsilon - w(k(a),\varepsilon) + \lambda \int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x) - (\lambda + s)J(k(a),\varepsilon)$$

with r the interest rate and where $w(k(a),\varepsilon)$ stands for the wage.

It is not in the best interest of firms to keep workers on working if the job value is negative, i.e. $J(k(a),\varepsilon) \leq 0$.[9] Therefore, the threshold value for productivity $R(k(a))$ satisfies $J(k(a), R(k(a))) = 0$ and is positively related to wages but negatively to the labor hoarding job value since a new shock may hit the job:

the training sessions are enrolled while employed and are not only firms-financed but also job-related, making them apparently more specific.

[8]Introducing an exogenous job destruction rate allows for a positive unemployment rate.

[9]Actually, due to Nash bargaining of wages, this separation rule is also optimal from workers' point of view.

$$R(k(a)) = -y(a,k(a)) + w\left(k(a), R(k(a))\right) - \lambda \int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x) \tag{5.1}$$

The recruiting policy is determined by the expected average value of the job once filled. But, as firms cannot *ex ante* target hirings among heterogenous workers and as a particular productivity shock is drawn once a worker is contacted, the vacancy decision depends both on abilities' and on productivity shocks' distributions. Besides, we assume that the training cost $C(k(a))$[10] is fully born by the firm at the time of match formation (before the wage bargaining). The value of a vacancy is therefore defined as follows:

$$rV = -c + q(\theta)\int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right)\left\{\int_{R(k(a))}^{1}\left[J(k(a),\varepsilon) - V\right]\mathrm{d}G(\varepsilon) - C(k(a))\right\}\mathrm{d}a$$

with $c \geq 0$ the flow cost of recruiting a worker.

A standard free entry condition (such that the value of vacancies vanishes in equilibrium) then determines the labor market tightness θ and implies that the expected recruitment cost equalizes *ex ante* the expected value of job creation:

$$\frac{c}{q(\theta)} = \int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right)\left\{\max_{k(a)} \int_{R(k(a))}^{1} J(k(a),\varepsilon)\mathrm{d}G(\varepsilon) - C(k(a))\right\}\mathrm{d}a \tag{5.2}$$

The average expected value of a contact especially depends on the ability distribution of unemployed workers. As we will show in Section 5.4, the heterogeneity across abilities in hiring values imply the existence of a composition externality in the search process[11]: the more unemployed workers with low abilities there are, the higher the probability to contact a low-ability worker, and hence the lower the expected return on a vacancy.

[10] with $C'(0) = 0$, $C'(k(a)) > 0$ and $C''(k(a)) = 0$.

[11] This composition externality in the search process due to the heterogeneity across abilities is similar to the intergenerational externality first highlighted by Chéron, Hairault and Langot (2011).

Training investment decision At the time of match formation, before the shock occurs, the firm chooses how much specific training they invest in, in order to maximize the net expected value of a filled job. It follows that the investment decision is stated as:

$$\max_{k(a)\geq 0} \int_{R(k(a))}^{1} J(k(a),\varepsilon)\mathrm{d}G(\varepsilon) - C(k(a)) \quad \Longrightarrow \quad C'(k(a)) = \int_{R(k(a))}^{1} J_1(k(a),\varepsilon)\mathrm{d}G(\varepsilon)$$

In this way, firms decide on the sum they invest in specific training so that the expected marginal return on investment is equaled to its marginal cost. Beyond the ability level itself, the marginal return particulary depends on the relation between the bargained wage and the investment level, underlying a potential holdup problem. Effects of training investments on wages and job destructions are highly dependent on the wage setting game.

5.2.3 Wage setting

Wages are determined by a Nash bargaining. The firm and the worker share the global surplus generated by a job according to their bargaining power. But, when firms support hiring costs (such as a training cost), a natural holdup problem may arise. Indeed, the initial wage of the two-tier wage structure (that should optimally arise in such a case)[12] may not be credible. The reason is that new workers have an incentive to renegotiate immediately after been hired as training investments require continuing relationships to be efficient. Therefore, second-tier wages apply initially as well as subsequent to any shock to match productivity. The ex-post bargaining process increases employees' threat point. Demanding a higher wage, workers capture some of the rents created by the training cost without paying for, leading finally to a holdup problem. Therefore, we consider another source of inefficiency (holdup) in addition to the composition externality arising out of the heterogeneity across abilities. In section 5.4, we will deal with these two sources of inefficiency.

As proposed by Mortensen and Pissarides(1999), we consider an insider wage structure that applies initially and subsequently to any shock to match productivity. Since firms have to

[12]Following Mortensen and Pissarides (1999), the wage structure that arises when firms are liable for hiring costs (a training cost here) is a two-tier one. On one hand, the initial wage reflects the fact that workers share in the initial hiring (training) cost by accepting a lower wage. On the other hand, renegotiated wages subsequent to match productivity shocks ("insider wage") no longer include training costs since they are already sunk.

pay the training cost in both cases of success and failure of the wage bargaining, the global surplus generated by a job that is shared between the firm and the worker, writes: $S(k(a),\varepsilon) = J(k(a),\varepsilon) + W(k(a),\varepsilon) - U(k(a))$, where workers' values of unemployment and employment are respectively given by:

$$rU(a) = z + \theta q(\theta) \int_{R(k(a))}^{1} \big[W(k(a),x) - U(a)\big] \mathrm{d}G(x)$$
$$rW(k(a),\varepsilon) = w(k(a),\varepsilon) + \lambda \int_{R(k(a))}^{1} \left[W(k(a),x) - W(k(a),\varepsilon)\right] \mathrm{d}G(x)$$
$$+ \left[s + \lambda G(R(k(a)))\right] \left[U(a) - W(k(a),\varepsilon)\right]$$

The standard Nash-sharing rule is:

$$W(k(a),\varepsilon) - U(a) = \gamma S(k(a),\varepsilon)$$

where γ stands for the bargaining power of workers. The following expression for the wage can be then derived:

$$w(k(a),\varepsilon) = (1-\gamma)z + \gamma\left[y(a,k(a)) + \varepsilon\right] + \gamma\theta q(\theta) \int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x) \tag{5.3}$$

In the context of *ex ante* unobservable heterogeneity and following Chéron et al (2011), the way search costs enter into the wage equation depends on the *ex post* value of the worker relative to the *ex ante* expected average value of job creation, defined over the whole pool of unemployed workers[13]. More precisely, defining $\tau(k(a)) = \frac{\int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x) - C(k(a))}{\int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right)\left\{\int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x) - C(k(a))\right\}\mathrm{d}a}$ and using equation (5.2), the wage expression can be rewritten as:

$$w(k(a),\varepsilon) = (1-\gamma)z + \gamma\left[y(a,k(a)) + \varepsilon + c\theta\tau(k(a))\right] + \gamma\theta q(\theta)C(k(a)) \tag{5.4}$$

Wages are ability-specific and depend of course on worker's reservation wage (z) and on her

[13] Chéron et al. (2011) emphasize the role search costs into the wage equation considering age-differentiated workers.

productivity $(y(a,k(a))+\varepsilon)$. Further, the last term in the right-hand side refers to the holdup problem. It rises the bargained wage. If the negotiation fails, the firm will have to pay another training cost $C(k(a))$ when a new worker will be hired. This event takes place at rate $\theta q(\theta)$. So, staying in the match, the worker enables the firm to save the expected cost $\theta q(\theta)C(k(a))$ and wages increase by a fraction γ of that saving by the Nash assumptions. And the higher the training investment, the stronger the holdup problem. Finally, wages depend on search costs the firm saves when the wage bargaining does not fail. High-ability workers are characterized by $\tau(k(a)) > 1$, which means that they receive more than the average search costs ($c\theta$). Both productivity $y(k(a))$ and search costs $c\theta\tau(k(a))$ then push up wages. Conversely, workers with low abilities get low wages not only due to a lower productivity value (for a given ε) but also because of a lower imputed value of search costs ($\tau(k(a)) < 1$).

5.2.4 Equilibrium definition

The conditions that simultaneously determine the labor market tightness θ, sets of training investments $k(a)$ and productivity thresholds $R(k(a))$ and unemployment levels by ability $u(a)$, $\forall a \in [\underline{a}, \overline{a}]$ can be now defined.

Definition 4. *The labor market equilibrium is defined by the following set of equations*[14]*:*

$$\frac{c}{q(\theta)} = \int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right) \left\{ \left(\frac{1-\gamma}{r+s+\lambda}\right) \int_{R(k(a))}^{1} [1-G(\varepsilon)]\,\mathrm{d}\varepsilon - C(k(a)) \right\} \mathrm{d}a \tag{5.5a}$$

$$R(k(a)) = -y(a,k(a)) + z - \left(\frac{\lambda - \gamma\theta q(\theta)}{r+s+\lambda}\right) \int_{R(k(a))}^{1} [1-G(\varepsilon)]\,\mathrm{d}\varepsilon \tag{5.5b}$$

$$C'(k(a)) = \left\{ \frac{(1-\gamma)\,[1-G(R(k(a)))]}{r+s+\lambda G(R(k(a))) + \gamma\theta q(\theta)\,[1-G(R(k(a)))]} \right\} y'(k(a)) \tag{5.5c}$$

$$u(a) = f(a)\frac{s+\lambda G(R(k(a)))}{s+\theta q(\theta)+\lambda G(R(k(a)))} \quad ; \quad u = \int_{\underline{a}}^{\bar{a}} u(a)\mathrm{d}a \tag{5.5d}$$

Proof. First, from $J(k(a), R(k(a))) = 0$ and the wage expression (5.3) and using the fact that $(r+s+\lambda)J(k(a),\varepsilon) - (r+s+\lambda)J(k(a),R(k(a))) = \varepsilon - R(k(a)) - w(k(a),\varepsilon) + w(k(a),R(k(a)))$, it follows that $(r+s+\lambda)J(k(a),\varepsilon) = (1-\gamma)\,[\varepsilon - R(k(a))]$. Second, integrating by parts

[14] In appendix 5.6.1, we remind equilibrium conditions in case of observable heterogeneity of workers, as in Chéron and Rouland (2011b).

$\int_{R(k(a))}^{1} J(k(a),x)\mathrm{d}G(x)$ leads to $\left(\frac{1-\gamma}{r+s+\lambda}\right)\int_{R(k(a))}^{1}[1-G(x)]\,\mathrm{d}x$.

□

The ability level of a worker has an ambiguous effect on the amount the firm is willing to invest in training (equation (5.5c)). On one hand, the higher the ability, the higher the training incentives because of a smaller endogenous job destruction probability ($\lambda G(R(k(a)))$ in the denominator). Indeed, training investments rise the worker productivity, which in turn increases both present and future expected average job value. But, on the other hand, the higher the ability, the more significant the holdup problem. The term $\gamma\theta q(\theta)$ refers to holdup. It cuts down on firms' training incentives because of the wage surplus workers get by threatening firms. Nevertheless, the smaller the bargaining power of workers, the lower the holdup, hence the higher training incentives.

5.3 Model properties: Quantitative analysis

Our strategy is to derive model properties and to confront them to real data. We aim at getting a credible enough model that reproduces some stylized facts about French workers.

We first present the data we use. Then, we describe the model calibration and finally, we examine some numerical experiments.

5.3.1 Data

We use French Labor Force surveys ("*Enquête Emploi*" provided by INSEE) in 2006 and 2007. For each year, we define our sample in the following way. We focus on the population of respondents who were working at the beginning of the first quarter and consider their situation on the labor market at the end of the quarter. We choose to select the subsample of low-skilled workers (skilled manual workers) aged from 18 and 60, working full-time jobs and employed by the private sector. We exclude farmers and self-employed. We also delete the few observations with missing values, mainly because of missing wages. Altogether that comes to a sample of 2981 individuals for 2006 and 3099 individuals for 2007.

We focus on the three following variables of interest. The first one is the monthly wage level, expressed in euros. Wages are divided into four intervals computed from three wage quartiles. Workers are then sorted according to the wage interval they belong to at the beginning of the quarter. The second outcome is about transitions from employment to unemployment. Therefore, we define a dummy variable which is equal to one when the worker has experienced a transition from employment to unemployment between the beginning and the end of the the quarter of each year under consideration. Then, for each wage quartile, we construct a quarterly employment-unemployment transition rate over the whole period by dividing the sum of workers who experienced a transition from employment to unemployment during a quarter (in 2006 and 2007), by the total number of workers in the quartile. The third outcome is the training amount a worker receive in average during a quarter of each year. We only consider firm-provided training related to jobs as we are interested in firm-specific skills. For both years and for each wage interval, we add all the training hours firms provided to workers during a quarter and we divide this sum by the total number of workers in the quartile to get the average amount of training per worker during a quarter (in hours).

Figure 5.1 shows the log-normal-like shape of low-skilled workers wage distribution in 2007. Figure 5.2 gives quarterly transition rates from employment to unemployment and average training amounts, both by wage quartile. First, French low-skilled workers are characterized by a negative correlation between wage quartiles and transition rates from employment to unemployment. Workers who belong to the last wage quartile are much less likely to experience such a transition than workers of the first wage quartile (about four times fewer). Second, French low-skilled workers are also characterized by a positive correlation between wages quartiles and training levels. The average quarterly training amount per worker is about twice higher for workers who belong to the fourth quartile. More precisely, all wage quartiles taken together, individuals in the sample have a 4.26% quarterly probability of benefiting from a firm-specific training session on average (*i.e.* about 17% a year). And the average training amount per worker who has been trained is about 25.87 hours a quarter[15].

[15]Keep in mind that workers in the model are trained at the time of match formation once and for all while these figures correspond to quarterly data. Hence, training amounts the model will generate will not be comparable to these training amounts found in data. Therefore, in the quantitative analysis of the model, we will focus on differences in training as a percentage of the average training amount of workers in the first wage quartile, rather than talking about pure differences in training amounts.

Figure 5.1: Low-skilled wage distribution (in 2007 French LFS)

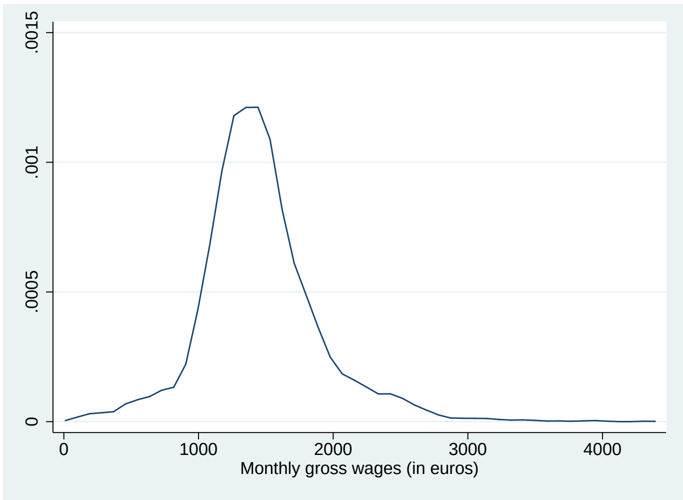
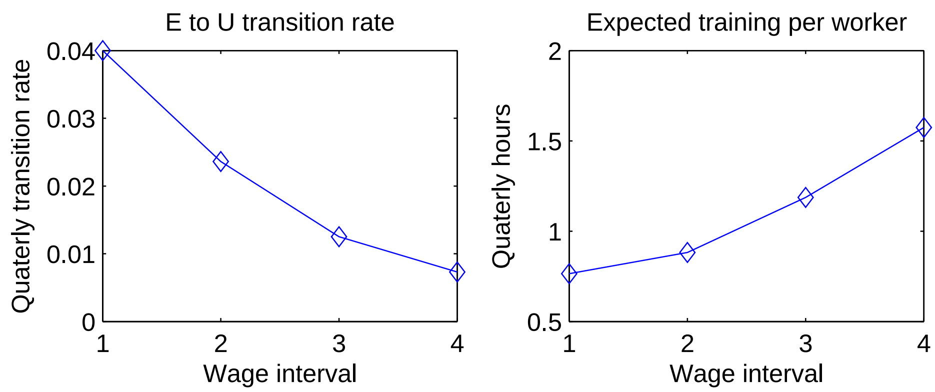


Figure 5.2: Quarterly E→U transition rates and expected quarterly training amounts of low-skilled workers, by wage interval (in 2006-2007 French LFS)



5.3.2 Calibration

We consider a quarterly calibration of the model. A first set of parameters is based on external information. A second one aims at replicating some stylized facts that characterize the French low-skilled workers ("ouvriers qualifiés") data set over the 2006-2007 period. We consider a homogenous period of the business cycle (before the current economic crisis) as our model does not allow for macroeconomic shocks and fluctuations[16]. We consider three targets: the unemployment rate, the average employment to unemployment exit rate and the median-to-mean wage ratio over this period. The empirical investigation of the quantitative properties of the model then consists in examining statistics summarizing wage quartiles, average separation rates and training amounts within each wage quartile interval.

As a preliminary step, specifications of functional forms for the matching function and the distributions of idiosyncratic shocks and abilities are required. We choose the simplest functions based on existing assumptions in the literature. In particular, as in Mortensen and Pissarides (1994), we consider a uniform distribution of shocks $G(x) = x \; \forall x \in [0,1]$ and a Cobb-Douglas matching function $M(v,u) = v^{\psi}u^{1-\psi}$. We also follow Mortensen (2003) by assuming that the dispersion of abilities is defined by a Pareto distribution. More specifically, $F(a) = 1 - \frac{1}{a} \; \forall a \in [1,2]$. We also have to define specifications of functional forms for the endogenous productivity $y(a,k(a))$ and the training cost $C(k(a))$. We assume a quadratic training cost $C(k) = \frac{1}{2}k^2$ and the productivity function is supposed strictly decreasing and concave $y(a,k) = ak^{\alpha}$. These specifications imply that the equilibrium conditions collapse to:

$$c\theta^{1-\psi} = \int_1^2 \left(\frac{u(a)}{u}\right)\left\{\left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\frac{1}{2(1-\underline{\varepsilon})}\right][1-R(k(a))]^2 - \frac{1}{2}k^{\alpha}(a)\right\}\mathrm{d}a \tag{5.6}$$

$$R(k(a)) = -ak^{\alpha}(a) + z - \left(\frac{\lambda-\gamma\theta^{\psi}}{r+s+\lambda}\right)\left[\frac{1}{2(1-\underline{\varepsilon})}\right][1-R(k(a))]^2 \tag{5.7}$$

$$k(a) = \left[\frac{a\alpha(1-\gamma)[1-R(k(a))]}{r+s+\lambda R(k(a)) + \gamma\theta^{\psi}[1-R(k(a))]}\right]^{\frac{1}{2-\alpha}} \tag{5.8}$$

$$u(a) = \left(\frac{1}{a^2}\right)\left(\frac{s+\lambda R(k(a))}{\theta^{\psi}+s+\lambda R(k(a))}\right) \quad ; \quad u = \int_1^2 u(a)\mathrm{d}a \tag{5.9}$$

[16] Postel-Vinay and Robin (2002) and Jolivet, Postel-Vinay and Robin (2006) (among many others) follow a similar strategy to estimate their steady-state search model.

Further, the wage earning of a worker is not only related to her ability but also to idiosyncratic shocks that hit the job. Formally, this results in $w = w(k(a), \varepsilon)$ defined by (5.3). From above specifications, wages can be rewritten as:

$$w(k(a,\varepsilon) = (1-\gamma)z + \gamma a k(a)^{\alpha} + \gamma\varepsilon + \left(\frac{\gamma\theta^{\psi}}{r+\lambda+s}\right)\left[\frac{1-\gamma}{2(1-\underline{\varepsilon})}\right][1-R(k(a))]^2 \tag{5.10}$$

To compute wages densities, we need to account for the endogenous job destruction decision as well. The density of wages denoted $\phi(w)$ can be then derived as follows:

$$\phi(w) \quad = \quad \int_1^2 \Psi(w,a)da \ \text{ with } \begin{cases} \Psi(w,a) = 0 \ \ \forall w < w(a, R(a)) \\ \Psi(w,a) = f(a) - u(a) \ \ \forall w \geq w(a, R(a)) \end{cases}$$

where we make use of the fact that the uniform distribution of shocks implies that the density of each productivity draw is unchanged all across the support of shocks.

The first set of parameters is consistent with conventional values assumed in the literature: $r = 0.01$ and $\psi = \gamma = 0.5$. The second set of parameters then includes $\{z, c, s, \lambda, \alpha\}$.

The model allows both for exogenous and endogenous productivity shocks. The exogenous job destruction rate is set to $s = 1.3\%$ to be consistent with the average quitting rate over the period 2006-2007 in France. Following Ljungqvist and Sargent (2008), the Poisson arrival rate of productivity shocks is consistent with an expected shock every two years, i.e. $\lambda = 1/8$. We choose $z = 1.23$ to get an average job destruction rate of 2.1% a quarter, consistent with the observed employment to unemployment exit rate for skilled manual workers over the period 2006-2007 in France. The labor market tightness θ should be consistent with an unemployment rate of 7.5% (from equation (5.9)), which gives the value of c as a solution of equation (5.6)[17]. This implies that the contact rate for a worker is $\theta^{\psi} = 31\%$ a quarter. The simulated average duration of an unemployment spell is then approximately 11 months. Lastly, we choose $\alpha = 0.29$ so that the ratio of median wage to mean wage is 0.95, as observed for the skilled manual workers

[17] Actually, only the steady-state value of θ^{ψ} matters in our simulation procedure since equations (5.7)-(5.9) for productivity thresholds, training amounts and unemployment, as well as equation (5.10) for wages, are expressed as functions of θ^{ψ}. c then satisfies equation (5.9) for a given θ.

over the period 2006-2007 in France.

5.3.3 Numerical experiments

This quantitative analysis first aims at looking at model properties according to the level ability. We then examine how good the model is at matching data.

Figures 5.3 and 5.4 show model properties. Figure 5.3 focuses on job destruction rates and training amounts. First, the model implies a decreasing relationship between job destruction rates and abilities for workers whose ability is 1.4 times higher than the lowest one: it starts with a quarterly rate of about 5% and falls to 1.3%. Above, workers no longer experience endogenous firing, but face the exogenous employment exit rate of 1.3%.
Second, the model implies a strictly increasing relationship between the amount firms are willing to invest in training and abilities. The higher the level ability, the higher the training volume.

Figure 5.4 shows the hump-shaped wage distribution the model generates. As a combination of the exogenous distributions of abilities and productivity shocks, the wage distribution is determined by firms' decision rule about reservation productivity. As explained in Chéron and Rouland (2011a), this first relies on the fact that the reservation productivity of low-ability workers is high, which implies that only low-ability workers who draw a good productivity are in a position to keep their job. All else being equal, this raises the average wage of low-ability workers. Secondly, the reservation productivity of high-ability workers is low. Again, all else being equal, high-ability workers whose job has been hit by a bad productivity shock earn lower wages, which move them to the left in the wage distribution.

Finally, we compare model implications to statistics computed from the group of low-skilled (manual skilled) workers in France. Figure 5.5 reports three statistics. The panel up left gives the value of each wage quartile with respect to the median wage. The panel up right gives the average job destruction rate in each wage interval over the job destruction rate of workers who belong to the first wage quartile. And the panel down left shows the average quarterly training amount per worker in each wage interval over the average training amount per worker in the first wage quartile. All in all, the model performs pretty well. Ratios of wage quartiles over the median wage are quite well replicated, so are job destructions rate by wage interval. However,

Figure 5.3: Model properties (I): Quarterly job destruction rates and training levels, by ability

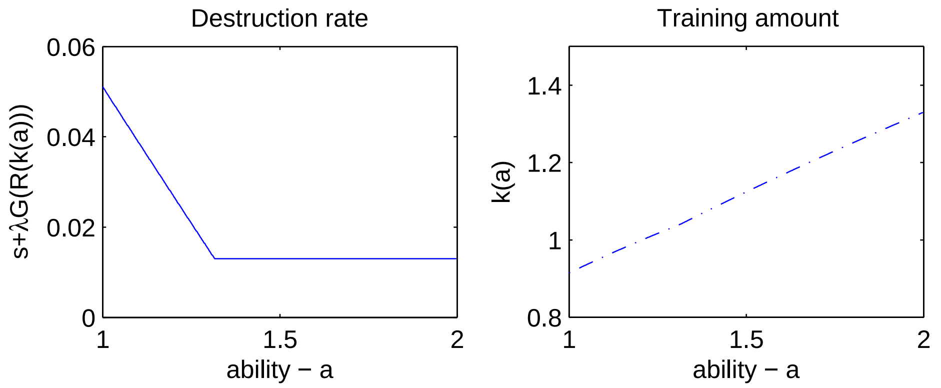
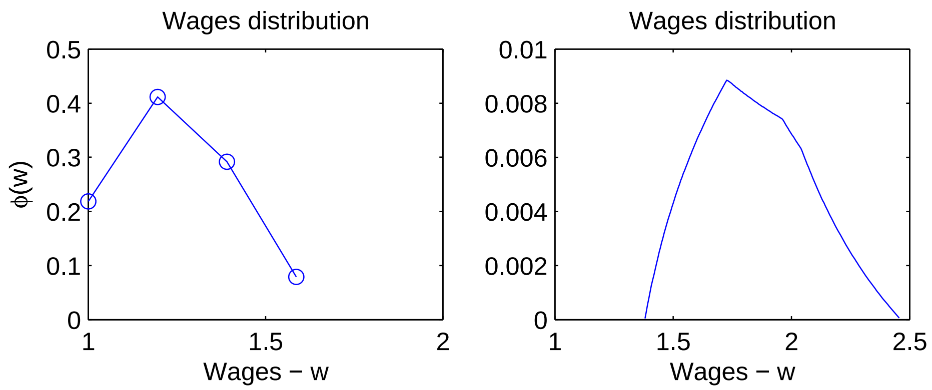
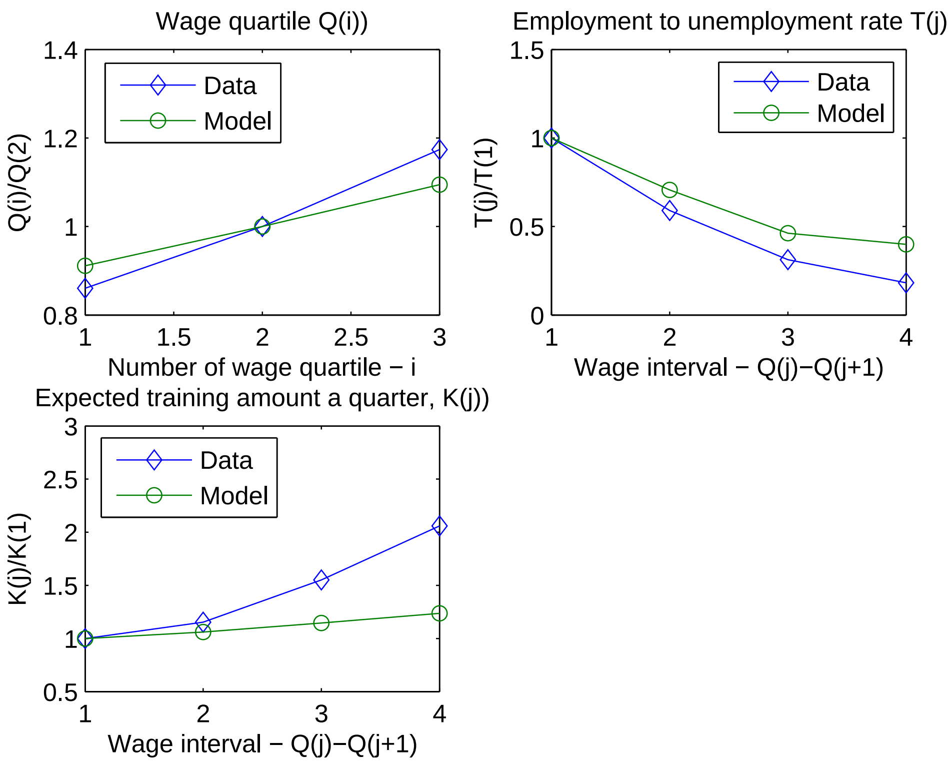


Figure 5.4: Model properties (II): the distribution of wages



the model does not generates high enough differences between the first and the last wage quartile, especially for the training amount. According to data, the average training volume of workers who belong to the fourth wage quartile is supposed to be twice higher than the average volume of workers who belong to the first wage quartile but it is only about one 25% higher in the model. The heterogeneity across abilities implies not high enough differences in training but they are qualitatively correct. The model performance could be improved considering on-the-job training and firms' heterogeneity as well.

Figure 5.5: Model assessment



Overall, although its performance is far from perfect, these numerical experiments show that the model is credible enough to analyze the design of optimal labor market policy. Next section deals with this issue.

5.4 Optimal labor market policy

Since unobservable heterogeneity of workers brings inefficiency *per se*, we should correct it even though the whole heterogeneity will not be removed. Besides, the holdup problem is another source of inefficiency, also depending on unobservable heterogeneity of workers. Altogether, the Hosios condition (the elasticity relative to vacancies in the matching function is equal to the bargaining power of workers) cannot lead to reach social optimality, as in the equilibrium unemployment benchmark.

We first characterize the efficient allocation and we present then the optimal labor market policy that reaches it.

5.4.1 The efficient allocation

We derive the optimal allocation by maximizing steady-state output with respect to the labor market tightness $\theta^\star$, the reservation productivity $R^\star(k(a))$ and the training investment $k^\star(a)$. Steady-state output is made up of the total production of both unemployed and employed workers, net of search and training costs. The problem of the planner is stated as follows:

$$\underset{\{\theta, R(k(a)), k(a)\}}{max} \int_0^\infty e^{-rt} \left\{ \int_{\underline{a}}^{\overline{a}} \left[\overline{y}(a, k(a)) + u(a)z - \frac{c\theta u}{[\overline{a} - \underline{a}]} - \theta q(\theta) u(a) C(k(a)) \right] \mathrm{d}a \right\} \mathrm{d}t.$$

where $u = \int_{\underline{a}}^{\overline{a}} u(a)da$, and subject to the evolution of $u(a)$ and of the average output of employed workers with specific ability a denoted by $\overline{y}(k(a))$:

$$\dot{u}(a) = [f(a) - u(a)] [\lambda G(R(a)) + s] - u(a)\theta q(\theta) \tag{5.11}$$

and

$$\begin{aligned} \dot{\overline{y}}(a, k(a)) = u(a)\theta q(\theta) & \int_{R(k)}^1 [\varepsilon + y(a, k(a))] \,\mathrm{d}G(\varepsilon) \\ & + \lambda [f(a) - u(a)] \int_{R(k)}^1 [\varepsilon + y(a, k(a))] \,\mathrm{d}G(\varepsilon) \\ & - (\lambda + s)\overline{y}(a, k(a)) \end{aligned} \tag{5.12}$$

Definition 5. *Defining* $\eta(\theta) = \frac{-\theta q'(\theta)}{q(\theta)}$, *the efficient labor market allocation is then characterized by a triplet* $\{\theta^\star, R^\star(k(a)), k^\star(a)\}$ *solving:*

$$\frac{c}{q(\theta^\star)} = \int_{\underline{a}}^{\bar{a}} \left(\frac{u(a)}{u}\right) \left\{ \left(\frac{1-\eta(\theta^\star)}{r+\lambda+s}\right) \int_{R^\star(k(a))}^{1} [1-G(\varepsilon)]\,\mathrm{d}\varepsilon - (1-\eta(\theta^\star))C(k^\star(a)) \right\} \mathrm{d}a \tag{5.13a}$$

$$R^\star(k(a)) = -y(a,k^\star(a)) + z - c\theta^\star - \theta^\star q(\theta^\star)C(k^\star(a)) - \left(\frac{\lambda - \theta^\star q(\theta^\star)}{r+s+\lambda}\right) \int_{R^\star(k(a))}^{1} [1-G(\varepsilon)]\,\mathrm{d}\varepsilon \tag{5.13b}$$

$$C'(k^\star(a)) = \left(\frac{1-G(R^\star(k(a)))}{r+\lambda G(R^\star(k(a)))+s}\right) y'(k^\star(a)) \tag{5.13c}$$

$$u(a) = f(a)\frac{s+\lambda G(R^\star(k(a)))}{s+\theta^\star q(\theta^\star)+\lambda G(R^\star(k(a)))} \quad ; \quad u = \int_{\underline{a}}^{\bar{a}} u(a)\mathrm{d}a$$

5.4.2 Removing inefficiencies

This last section investigates the way to restore the optimality of equilibrium choices. Distortions come from two sources: holdup and unobservable heterogeneity of workers across abilities. But both inefficiencies are not related. Therefore, implications for policy should be different. Hence, we characterize the optimal labor market policy by removing inefficiencies one after the other. In a first step, we address the holdup issue by considering observable heterogeneity of workers. Then, we characterize the optimal labor market policy when workers' heterogeneity is no longer observable and training investments lead to a holdup problem.

Removing inefficiencies in case of observable heterogeneity

When workers' heterogeneity is observable, there is no longer a composition externality in the search process. Firms can direct their search according to the ability level. Hence, the expected average hiring value only depends on distribution of shocks.

Job destruction decisions and training investment decisions are strongly complementary: a fraction of the expected training cost that the firm saves when the worker stays in the match is captured by the worker through the wage bargaining (holdup). This rises the productivity threshold, leading finally to an excess of job destructions. And, the higher the amount firms are willing to invest in training, the higher the distortion. Therefore, firing taxes F can be implemented to reach the efficient level of job destructions. Training subsidies T that firms would get at the time of match formation should also be implemented in order to lower the training cost.

Proposition 6. *Assuming $\gamma = \eta(\theta)$, the optimal labor market policy with training subsidies and firing taxes $\{T, F\}$ when workers' heterogeneity is observable, solves*[18]*:*

$$\begin{cases} T &= \gamma C(k^\star) + \left[1 - G(R^\star(k))\right] F \\ \\ F &= \left(\frac{1}{r+s}\right) \gamma \theta^\star q(\theta^\star) C(k^\star) \end{cases}$$

where $k^\star$, $R^\star(k)$ and $\theta^\star$ solve the optimal allocation.

Proof. See appendix 5.6.1. □

Firstly, $1/[r+s]$ defines the discount factor that depends both on the interest and on the likelihood of voluntary quits: with probability s, firms will not have to pay firing taxes. Therefore, the higher the probability the worker quits voluntary the firm, the lower the distortions, hence firing taxes. $\gamma\theta q(\theta)C(k(a))$ defines the instantaneous value of the distortion due to holdup. It is all the more important that the training investment is high. Therefore, firing taxes should be strictly positively related to training amounts.

Secondly, training subsidies should be implemented because workers do not share in the training cost while they benefit from the investment (through the productivity). On the contrary, firms have to pay all the training cost without benefiting from the whole return. Therefore, the optimal training subsidy should integrate this distortion plus the negative incidence of the firing tax on job creations. Since firing taxes are not due if the worker is not hired actually, training subsidies also depend on the probability of endogenous job destruction after the first shock draw. Again, since the optimal profile of firing taxes is strictly with invested amounts, training subsidies should also be.

[18]This section is very close to Chéron and Rouland (2011b) who deal with holdup inefficiencies in case of directed search. Therefore, we directly present the optimal policy. Details can be found in Appendix 5.6.1.

Removing inefficiencies in case of unobservable heterogeneity

In case of unobservable heterogeneity, job destructions are inefficient not only because of the holdup problem but also because of the excess (lack) of search costs -compared to the average value of search costs- high-ability (low-ability) workers get (suffer from) when the wage bargaining does not fail. For now, we assume that policy can be ability-targeted while obviously, it cannot be in real world. Therefore, this analysis needs to be extended considering a proxy of this first-best policy.

Again, firing taxes $F(a)$ and training subsidies $T(a)$ can be implemented to reach the optimal allocation. With training subsidies, the value of a vacancy is now such as:

$$rV = -c + q(\theta)\int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right)\left\{\int_{R(k(a))}^{1} J(k(a),\varepsilon)\mathrm{d}G(\varepsilon) - C(k(a)) + T - V\right\}\mathrm{d}a$$

The free entry condition implies $V = 0$ and then:

$$\frac{c}{q(\theta)} = \int_{\underline{a}}^{\overline{a}} \left(\frac{u(a)}{u}\right)\left\{\int_{R(k(a))}^{1} J(k(a),\varepsilon)\mathrm{d}G(\varepsilon) - C(k(a)) + T\right\}\mathrm{d}a \tag{5.14}$$

The reservation productivity $R(k(a))$ is defined by $J(k(a), R(k(a))) = -F$. In the context of an insider wage structure, the surplus sharing rule is now such that $W(k(a),\varepsilon) - U(a) = \gamma S(k(a),\varepsilon)$ and $J(k(a),\varepsilon)+F = (1-\gamma)S(k(a),\varepsilon)$ where $S(k(a),\varepsilon) = J(k(a),\varepsilon)+F+W(k(a),\varepsilon)-U(a)$. We therefore derive the following wage expression:

$$\begin{aligned} w^P(k(a),\varepsilon) = (1-\gamma)z &+ \gamma\left[y(a,k(a)) + \varepsilon + c\theta\tau(k(a))\right] \\ &+ \gamma\left\{r + s + \theta q(\theta)\left[1 - G(R(k(a)))\right]\right\}F \\ &+ \gamma\theta q(\theta)\left[C(k(a)) - T\right] \end{aligned} \tag{5.15}$$

On the one hand, the training subsidy reduces the training cost (last term of the right-hand side). But, on the other hand, workers are now in a position to capture also a fraction γ of the firing taxes the firm saves if the job is not destroyed and if the negotiation does not fail (second term of the right-hand side).

Proposition 7. *Assuming $\gamma = \eta(\theta)$, the optimal labor market policy with training subsidies and firing taxes $\{T(a), F(a)\}$ when workers' heterogeneity is unobservable, solves:*

$$\begin{cases} T(a) & = \gamma C(k^\star(a)) + [1 - G(R^\star(k(a)))]\, F(a) \\ \\ F(a) & = \left(\frac{1}{r+s}\right) \{\gamma\theta^\star q(\theta^\star) C(k^\star(a)) + c\theta^\star [1 - \tau(k^\star(a))]\} \end{cases}$$

where $k(a)^\star$, $R^\star(k(a))$ and $\theta^\star$ solve the optimal allocation.

Proof. See appendix 5.6.2. □

Compared with the previous case where the heterogeneity of workers was observable, there is now another source of inefficiency. $c\theta\,[1 - \tau(k(a))]$ stands for the distortion that comes from search costs the firm saves when the wage bargaining does not fail. Since high-ability workers are characterized by $\tau(k(a)) > 1$ (which means that they receive more than the average search costs $c\theta$), this distortion from search costs reduces the value of their firing taxes[19]. Conversely, $\tau(k(a)) < 1$ for low-ability workers, which rises the part of their firing taxes related to the distortion from search costs. But, compared to high-ability workers, the inefficiency arising from holdup is low. So is the corresponding part of firing taxes. In the end, the optimal profile of firing taxes should clearly depend on worker's ability but both distortions are oppositely related to ability. A low-ability level implies fewer holdup (since the training investment is low), which reduces the value of firing taxes. But, the negative composition externality low-ability workers imply in the search-process rises firing taxes of workers with the lowest abilities. Therefore, we cannot determine analytically the ability-profile of firing taxes.

As before, training subsidies reflect both the negative incidence of firing taxes on job creations and the fact that workers do not share in the training cost. Again, since the optimal profile of

[19]With a two-tier wage structure, there would be no holdup, hence no related distortion. Firing taxes of high-ability workers would be then strictly negative and negatively correlated to worker's level ability. Firms should have incentives to get rid of workers with a high ability level, who impose an additional wage as they have a better ability than the average one. Conversely, firing taxes of low-ability workers would be strictly positive to encourage firms to keep them in employment.

firing taxes is not clearly stated, we cannot determine analytically the optimal ability-profile of training subsidies.

The optimal design of firing taxes and training subsidies totally depends on workers heterogeneity. But the level ability is precisely unobservable, both for firms and the planner. Therefore, the optimal policy we defined cannot be implemented in this condition. We are aware of this shortcoming. But at least it highlights how the optimal labor market policy should respond to workers' unobservable heterogeneity, compared to the case of observable heterogeneity (Chéron and Rouland (2011b)). One way to implement such a policy would be to rely on an observable characteristic of worker's heterogeneity instead of abilities. Wages could be useful. But, wages and optimal firing taxes and training subsidies would be still ambiguously related.

5.5 Conclusion

The goal of this paper is to address heterogeneity issues about wages, job destruction rates and training investments. To that end, we have considered a matching model with idiosyncratic productivity shocks, endogenous firm-specific investments in training and workers' heterogeneity according to *ex ante* unobservable abilities. From numerical experiments confronted to real data, we have shown that this framework is credible enough to deal with the two sources of inefficiency that arise in such a theoretical framework -a holdup problem and a composition externality in the search process. The optimal labor market policy then consists in implementing both training subsidies firing taxes. In a first-best approach, both instruments should depend on the ability level.

5.6 Appendix

5.6.1 Equilibrium equations with observable heterogeneity

In case of homogeneous workers, the expected average value of a job once filled only depends on the productivity shocks' distribution (no longer on the ability distribution). All workers receive the same training investment (still depending on the ability level).

Proposition 8. *For a particular ability level a, the labor market equilibrium is defined by the following set of equations:*

$$
\begin{aligned}
\frac{c}{q(\theta)} &= \left(\frac{1-\gamma}{r+s+\lambda}\right)\int_{R(k)}^{1}[1-G(x)]\,\mathrm{d}x - C(k)\\
R(k) &= -y(a,k)+z+\left(\frac{\gamma}{1-\gamma}\right)c\theta+\left(\frac{\gamma}{1-\gamma}\right)\theta q(\theta)C(k)-\frac{\lambda}{r+s+\lambda}\int_{R(k)}^{1}[1-G(x)]\,\mathrm{d}x\\
C'(k) &= \left\{\frac{(1-\gamma)[1-G(R(k))]}{r+s+\lambda G(R(k))+\gamma\theta q(\theta)\left[1-G(R(k))\right]}\right\}y'(k)
\end{aligned}
$$

Proof. $\tau(a)=1$ in equations (5.5a) to (5.5c). □

This implies $\theta(a)$, $R(a)$ and $k(a)$.

5.6.2 Equilibrium equations with labor market policy

Wage setting First, the surplus from a match, $S(k(a),\varepsilon)=J(k(a),\varepsilon)+F(a)+W(k(a),\varepsilon)-U(a)$ shared such as $W(k(a),\varepsilon)-U(a)=\frac{\gamma}{1-\gamma}J(k(a),\varepsilon)$, implies $w(k(a),\varepsilon)=(1-\gamma)z+\gamma\left[y(k(a))+\varepsilon\right]+\gamma\theta q(\theta)[1-G(R(k(a)))]F(a)+\gamma(r+s)F(a)+\gamma\theta q(\theta)\int_{R(k(a))}^{1}J(k(a),x)\mathrm{d}G(x)$.

Then, defining $\tau(k(a))=\frac{\int_{R(k(a))}^{1}[J(k(a),x)\mathrm{d}G(x)-C(k(a))+T(a)}{\int_{\underline{a}}^{\overline{a}}\left(\frac{u(a)}{u}\right)\left\{\int_{R(k(a))}^{1}J(k(a),x)\mathrm{d}G(x)-C(k(a))+T(a)\right\}\mathrm{d}a}$, wages finally write:

$$
\begin{aligned}
w^P(k(a),\varepsilon) = {} & (1-\gamma)z+\gamma\left[y(k(a))+\varepsilon+c\theta\tau(k(a))\right]+\gamma\theta q(\theta)\left[C(k(a))-T(a)\right]\\
& +\gamma\left\{r+s+\theta q(\theta)\left[1-G(R(k(a)))\right]\right\}F(a)
\end{aligned}
$$

Job creations Endogenous job destruction rule now such as $J(k(a),\varepsilon)<-F(a)$ leads to a reservation productivity $R(k(a))$ defined by $J(k(a),R(k(a)))=-F(a)$. Therefore, it follows $(r+s+\lambda)\left[J(k(a),\varepsilon)-J(k(a),R(k(a)))\right]=\varepsilon-R(k(a))-w^P(k(a),\varepsilon)+w^P(k(a),R(k(a)))$. Using the wage expression $w^P(k(a),\varepsilon)$ it comes that:

$$\begin{aligned}(r+s+\lambda)\left[J(k(a),\varepsilon)-J(k(a),R(k(a)))\right] &= (1-\gamma)\left[\varepsilon-R(k(a))\right]\\ \Leftrightarrow\quad J(k(a),\varepsilon) &= \left(\tfrac{1-\gamma}{r+\lambda+s}\right)\left[\varepsilon-R(k(a))\right]-F(a)\end{aligned}$$

Job creations equation is then defined by:

$$\frac{c}{q(\theta^P)}=\int_{\underline{a}}^{\overline{a}}\left(\frac{u(a)}{u}\right)\Bigg\{\left(\frac{1-\gamma}{r+s+\lambda}\right)\int_{R^P(k(a))}^{1}\left[1-G(\varepsilon)\right]\mathrm{d}\varepsilon-C(k^P(a))+T(a)$$
$$-\left[1-G(R^P(k(a)))\right]F(a)\Bigg\}\,\mathrm{d}a$$

Making job creations optimal when Hosios condition holds implies:

$$\frac{c}{q(\theta^P)}=\frac{c}{q(\theta^\star)}\quad\Leftrightarrow\quad T(a)=\gamma C(k^\star(a))+\left[1-G(R(k(a)))\right]F(a)$$

where $k^\star(a)$ solves the optimal allocation.

Job destructions The reservation productivity $R(k(a))$ defined by $J(k(a),R(k(a)))=-F(a)$ is such as:

$$R(k(a))=-y(k(a))+w^P\left(k(a),R(k(a))\right)-(r+s)F(a)-\lambda\int_{R(k(a))}^{1}\left[J(k(a),x)+F(a)\right]\mathrm{d}G(x)$$

As mentioned before, $J(k(a),R(k(a)))=-F(a)$ implies $J(k(a),\varepsilon)=\left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\varepsilon-R(k(a))\right]-F(a)$. It then comes that $\int_{R(k(a))}^{1}\left[J(k(a),\varepsilon)+F(a)\right]\mathrm{d}G(x)=\left(\frac{1-\gamma}{r+\lambda+s}\right)\int_{R(k(a))}^{1}\left[1-G(\varepsilon)\right]\mathrm{d}\varepsilon$ by integrating by parts. Replacing the wage expression and using the definition of $\tau(k(a))$ finally leads to the job destruction equation:

$$R^P(k(a))=z-y(k^P(a))-(r+s)F(a)-\left[\frac{\lambda-\gamma\theta^P q(\theta^P)}{r+s+\lambda}\right]\int_{R^P(k(a))}^{1}\left[1-G(x)\right]\mathrm{d}x$$

Given that $T(a)=\gamma C(k(a))+\left[1-G(R(k(a)))\right]F(a)$, making job destructions optimal when

Hosios condition holds implies:

$$R^P(k) = R^\star(k) \quad \Leftrightarrow \quad F(a) = \left(\frac{1}{r+s}\right)\{\gamma\theta^\star q(\theta^\star)C(k^\star(a)) + c\theta^\star\left[1-\tau(k^\star(a))\right]\}$$

where $k^\star(a)$ and $\theta^\star$ solve the optimal allocation.

Training investment level Since $T(a) = \gamma C(K(a)) + \left[1 - G(R(k(a)))\right] F(a)$ and since $J(k(a), \varepsilon) = \left(\frac{1-\gamma}{r+\lambda+s}\right)\left[\varepsilon - R(k(a))\right] - F(a)$ at equilibrium, the training investment level is determined by:

$$\begin{aligned} & \underset{k(a)}{max} \int_{R(k(a))}^{1} J(k(a),\varepsilon)\mathrm{d}G(\varepsilon) - C(k(a)) + T(a) \\ \Leftrightarrow \quad & \underset{k(a)}{max} \left(\frac{1-\gamma}{\lambda+s}\right)\int_{R(k(a))}^{1}\left[\bar{\varepsilon} - R(k)\right]\mathrm{d}G(\varepsilon) - (1-\gamma)C(k(a)) \end{aligned}$$

where the first order condition implies $C'(k(a)) = -\left(\frac{1-G(R(k(a)))}{r+\lambda+s}\right)\frac{\partial R(k(a))}{\partial k(a)}$ with $\frac{\partial R(k(a))}{\partial k(a)} = -\left\{\frac{r+\lambda+s}{r+\lambda G(R(k(a)))+s+\theta q(\theta)[1-G(R(k(a)))]}\right\}\left[y'(k(a)) + \theta q(\theta)C'(k(a))\right]$.

Training equation is finally defined by $C'(k^P(a)) = \left[\frac{1-G(R^P(k(a)))}{r+\lambda G(R^P(k(a)))+s}\right] y'(k^P(a))$.

Bibliography

Acemoglu D. 1997. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies* 64 (3): 445–64.

Acemoglu D., and Pischke J.-S. 1998. "Why do firms train? Theory and evidence." *Quarterly Journal of Economics*, no. 113:79–119.

———. 1999a. "Beyond Becker: Training in imperfect labour markets." *Economic Journal* 453 (109): F112–F142.

———. 1999b. "The structure of wages and investment in general training." *Journal of Political Economy* 3 (107): 539–572.

Acemoglu D., and Shimer R. 1999. "Holdups and efficiency with search frictions." *International Economic Review* 109 (40): 827–849.

Aeberhardt R., Givord P., and Marbot C. 2011. "Minimum wage and wage inequality in France: An unconditional quantile regresion approach." *Forthcoming in Documents de Travail de la DESE-Working Papers of the DESE.*

Albrecht J., and Vroman S. 2002. "A matching model with endogenous skill requirements." *International Economic Review* 43 (1): 283–305.

Andolfatto D. 1996. "Business cycles and labor market search." *American Economic Review* 86 (1): 112–132.

Bassanini A., Booth A., Brunello G., De Paola M., and Leuven E. 2007. In *Workplace training in Europe*, edited by Wasmer E. (eds) Education in Brunello G., Garibaldi P. and Training in Europe. Oxford University Press.

Becker G. 1962. “Investment in human capital: a theoretical analysis.” *Journal of Political Economy*, no. 70:9–49.

———. 1964. *Human Capital: A theorical and empirical analysis with special reference to education.* University of Chicago Press.

Behaghel L., Crépon B., and Sédillot B. 2008. “The perverse effects of partial employement protection reform: The case of French older workers.” *Journal of Public Economics* 92 (3-4): 696–721.

Belot M., Boone J., and van Ours J. 2007. “Welfare-improving employment protection.” *Economica*, no. 295.

Bishop J. H. 1991. Chapter On-the-job training of new hires of *Market failure in training? New economic analysis and evidence on training of adult employees*, edited by in Stern D. and Ritzen J., 61–98. Berlin: Springer Verlag.

Blàquez M., and Jansen M. 2008. “Search, mismatch and unemployment.” *European Economic Review* 52 (3): 498–526.

Blasco S., Crépon B., and Kamionka T. 2008, January. “Evaluation of the impact of training on individual labor market transitions.”

Blundell R., and Costa Dias M. 2000. “Evaluation methods for non-experimental data.” *Fiscal Studies* 21 (4): 427–468.

Blundell R., Costa Dias M., Meghir C., and Van Reenen J. 2004. “Evaluating the Employment Impact of a Mandatory Job Search Program.” *Journal of the European Economic Association* 2 (4): 569–606.

Bontemps C., Robin J.M., and van den Berg G. 1999. “An empirical equilibrium job search model with search on the job and heterogeneous workers and firms.” *International Economic Review* 40 (4): 1039–1074.

———. 2000. “Equilibrium search with continuous productivity dispersion: Theory and non-parametric estimation.” *International Economic Review* 41 (2): 305–358.

Budria S., and Telhado Pereira P. 2004. “On the returns to training in Portugal.” IZA Discussion Paper No. 1429.

———. 2007. “The wage effects of taining in Portugal: Differences across skill groups, genders, sectors, and training types.” *Applied Economics*, no. 39:787–807.

Burdett K., and Mortensen D. 1998. “Wage differentials, employer size, and unemployment,.” *International Economic Review,* 2 (39): 257–273.

Cahuc P., Postel-Vinay F., and Robin J.M. 2006. “Wage Bargaining with On-the-Job Search: Theory and Evidence.” *Econometrica* 74 (2): 323–364.

Cahuc P., and Zylberberg A. 2006. “La formation professionnelle des adultes : un systéme á la dérive.” Technical Report, COEP, CCIP.

Chernozhukov V., and Hansen C. 2005. “An IV Model of Quantile Treatment Effects.” *Econometrica* 1 (73): 245–261.

Chéron A. 2005. “Efficient v.s. equilibrium unemployment with match-specific costs.” *Economics Letters*, no. 88:176–183.

Chéron A., Hairault J.-O., and Langot F. 2007, February. “Job creation and job destruction over the life cycle: The older workers in the spotlight.” Iza discussion papers 2597, Institute for the Study of Labor (IZA).

Chéron A., Hairault J.-O., and Langot F. 2011. “Age-dependent employment protection.” *Forthcoming in Economic Journal.*

Chéron A., and Langot F. 2004. “Labor market search and real business cycles: reconciling Nash-bargaining with real aage dynamics.” *Review of Economic Dynamics* 7 (2): 476–493.

Chéron A., and Rouland B. 2011a. “Endogenous Job Destructions and the Distribution of Wages.” *forthcoming in Labour Economics.*

———. 2011b. “Inefficient Job Destructions and Training with Holdup.” *Labour: Review of labour economics and industrial relations* 25 (4): 397–420.

Chéron A., Rouland B., and Wolff F.-C. 2008. “Returns to firm-provided training in France: Evidence on mobility and wages.” Working Paper.

Cole H., and Rogerson R. 1999. “Can the Mortensen-Pissarides matching model match the Business Cycle facts.” *International Economic Review* 40 (4): 933–959.

Dearden L., Machin S., Reed H., and Wilkinson D. 1997. "Labour Turnover and work-related training." Technical Report, The Institute for Fiscal Studies.

Dolado J., Jansen M., and Jimeno J. 2009. "On-the-Job Search in a Matching Model with Heterogeneous Jobs and Workers." *Economic Journal* 119 (534): 200–228.

Fella G. 2005. "Termination restrictions and investment in general training." *European Economic Review* 6 (49): 1479–1499.

Firpo S. 2007. "Efficient semiparametric estimation of quantile treatment effects." *Econometrica* 75 (1): 259–276.

Fu C. 2011. "Training, search and wage dispersion." *Review of Economic Dynamics* 14 (4): 650–666.

Goux D., and Maurin E. 2000. "Returns to firm-provided training: evidence from French worker-firm matched data." *Labour Economics*, no. 7:1–19.

Green F., Felstead A, Mayhew K., and Pack A. 2000. "The impact of training on labour mobility: individual and firm-level evidence from Britain." *British Journal of Industrial Relations* 2 (38): 261–275.

Greene W.H. 1998. "Gender economics courses in liberal arts colleges: further results." *Journal of Economic Education* 29 (4): 291–300.

Hairault J.-O., Langot F., and Sopraseuth T. 2010. "Distance to retirement and older workers' employment: The case for delaying the retirement age." *Journal of the European Economic Association* 8 (5): 1034–1076.

Hall R.E. 2005a. "Employment fluctuations with equilibrium wage stickiness." *American Economic Review* 95 (1): 50–65.

Hall R.E., and Milgrom P. 2008. "The limited influence of unemployment on the wage bargain." *American Economic Review* 98 (4): 1653–1674.

Hashimoto M. 1981. "Firm-specific human capital as a shared investment." *American Economic Review*, no. 71:475–482.

Heckman J., Ichimura H., and Todd P. 1998. "Matching as an econometric evaluation estimator." *Review of Economic Studies* 65 (2): 261–294.

Heckman J., Lalonde R., and Smith J. 1999. “The economics and econometrics of active labor market programs.” In *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, Volume 3, 1865–2097. Amsterdam: Elsevier Science.

Jansen M. 2010. “Job auctions and hold-ups.” *Labour Economics* 17 (3): 608–619.

Jolivet G., Postel-Vinay F., and Robin J.-M. 2006. “The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US.” *European Economic Review* 50 (4): 877–907.

Khaskhoussi T., and Langot F. 2008. “La proximité de l’âge de la retraite. Quels effets sur l’accumulation du capital humain et sur la recherche d’emploi?” *Revue Économique* 59 (3): 641–652.

Koenker R., and Bassett G. 1978. “Regression Quantiles.” *Econometrica* 46 (1): 33–50.

Kugler A., and Pica G. 2008. “Effects of employment protection on job and workers flows: Evidence from the 1990 Italian reform.” *Labour Economics* 15 (1): 78–95.

Lechthaler W. 2009. “The interaction of firing costs and firm training.” *Empirica* 36 (3): 331–350.

Leuven E. 2005. “The economics of private sector training: a survey of the literature.” *Journal of Economic Surveys* 19 (1): 91–111.

Leuven E., and Oosterbeek H. 2001. “Firm-specific human capital as a shared investment: Comment.” *American Economic Review*, no. 91:342–347.

———. 2008. “An alternative approach to estimate the wage returns to private-sector training.” *Journal of Applied Econometrics*, no. 23:423–434.

Ljungqvist L., and Sargent T. 2008. “Two Questions about European Unemployment.” *Econometrica* 76 (1): 1–29.

Loewenstein M.A., and Spletzer J.R. 1999. “The role of off-the-job training vs. On-the-job training for the mobility of women workers.” *Journal of Human Ressources*, no. 34:710–733.

Lynch L. 1991. “The role of off-the-job training vs. On-the-job training for the mobility of women workers.” *American Economic Review* 2 (81): 151–156.

Malcomson J. 1997. "Contracts, hold-up and labor markets." *Journal of Economic Literature*, no. 35:1916–1957.

Merz M. 1995. "Search in the labor market and the Real Business cycle." *Journal of Monetary Economics* 36 (2): 269–300.

Messe P.-J., and Rouland B. 2011. "Stricter employment protection on older workers and firms' incentives to train: Evidence from a natural experiment in France." Working Paper.

Millard S. 1996. "The cyclial effects of labour market policy." In *Advanced Business Cycle Theory*, edited by P.Y. Hénin. Springler Verlaag.

Mincer J. 1974. *Schooling, Experience, and Earnings.* New York: NBER Press.

Montizaan R., Cörvers F., and De Grip A. 2010. "The effects of pension rights and retirement age on training participation: Evidence from a natural experiment." *Labour Economics* 17 (1): 240–247.

Mortensen D. 2000. "Equilibrium unemployment with wage posting: Burdett-Mortensen meet Pissarides." In *Panel Data and Structural Labor Market Models*, edited by P. Jensen N.M. Kiefer in H. Bunzel, B.J. Christiansen and eds. D.T. Mortensen, 281–292. Amsterdan: Elsevier.

———. 2003. *Wage Dispersion: Why are Similar People Paid Differently.* The MIT Press.

Mortensen D., and Pissarides C. 1994. "Job creation and job destruction in the theory of unemployment." *Review of Economic Studies* 61 (3): 397–415.

———. 1999. "New developments in models of search in the labor market." In *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, Volume 3 of *Handbook of Labor Economics*, 2567–2627. Elsevier.

O'Connell P. 1999. "Adults in training: An international comparison of continuing education and training." Technical Report WD(99)1, OECD Center for Educational Research and Innovation.

Ok W., and Tergeist P. 2003. "Improving workers' skills: anatycal evidence and the role of the social partners." *OECD Social Employment and Migration Working Papers*, vol. No 10.

Parent D. "Wages and mobility: the impact of employer-provided training."

Picchio M., and van Ours J. C. 2011, Mar. “Retaining through Training: Even for Older Workers.” Iza discussion papers 5591.

Pierre G., and Scarpetta S. 2006. “Employment protection: do firms’ perceptions match with legislation?” *Economics Letters* 90 (3): 328–334.

Pigou A.C. 1912. *Wealth and welfare.* London: Macmillan.

Pissarides C. 2000. *Equilibrium unemployment theory.* 2nd. Volume 1. Edited by MIT Press Books. The MIT Press.

———. 2009. “The unemployment volatility puzzle: Is wage stickiness the answer?” *Econometrica* 77 (5): 1339–1369.

Postel-Vinay F., and Robin J.M. 2002. “Equilibrium wage dispersion with worker and employer heterogeneity.” *Econometrica* 70 (6): 2295–2350.

Quercioli E. 2005. “Training, turnover, and search.” *International Economic Review* 46 (1): 133–143.

Rosenbaum P., and Rubin D. 1983. “The central role of the propensity score in observational studies for causal effects.” *Biometrika* 70:41–55.

Rosholm M., and Svarer M. 2004. “Endogenous wage dispersion in a search-matching model.” *Labour Economics* 11 (5): 623–645.

Sato Y., and Sugiura H. 2003. “On labor policies affecting investment in general human capital.” *Labour* 17 (4): 599–622.

Schnalzenberger M., and Winter-Ebmer R. 2009. “Layoff tax and employment of the elderly.” *Labour Economics* 16 (6): 618–624.

Shimer R. 2005a. “The cyclial behavior of equilibrium unemployment and vacancies.” *The American economic Review* 95 (1): 25–49.

Shimer R., and Smith L. 2001. “Matching, search, and heterogeneity.” *The B.E. Journal of Macroeconomics)* 0, no. 1.

Smith J., and Todd P. 2005. “Does matching overcome LaLonde’s critique of nonexperimental estimators?” *Journal of Econometrics* 125 (1-2): 305–353.

Wilde J. 2000. "Identification of multiple equation probit models with endogenous dummy regressors." *Economic Letters* 69:309–312.

Young D. 2003. "Employment Protection Legislation: its economic impact and the case for reform." Economic paper 186, European Commission Directorate-General for Economic and Financial Affairs.