Ph.D. Dissertation in Economics

Noncompetitive Labour Markets, Severance Payments and Unemployment

by

Giulio Fella

The London School of Economics and Political Science

Abstract

The dissertation analyses the possible welfare-enhancing role of severance payments when labour markets are non-competitive.
Chapter 1 introduces the material in the thesis.
Chapter 2 presents a short survey of the results of the existing literature on dismissal costs.
Chapter 3 uses a strategic bargaining model to show that, once dismissal costs are correctly modelled as a payment which takes place only in case firms sever the relationship, firing costs cannot affect the separation rate in models featuring voluntary severance in the absence of restrictions. Firms will always find it profitable to induce workers to quit whenever separation is efficient. Only if some other source of inefficiency prevents firms and workers to split the rents from continuation can firing costs result in a reduced number of separations. In this case they may be efficient.
Chapter 4 analyses non-contractible firms' investment in general training in the presence of frictional unemployment. It argues that consensual layoff measures and other institutions that oblige firms to share the total separation payoff result in higher training. Since general training is vested in the worker on separation, in the absence of such measures, the firm would not capture any return to training in case of separation.
Chapter 5 shows that in a dynamic efficiency wage model the time-inconsistency of firing decisions implies that severance payments increase aggregate employment and are second-best Pareto optimal as they induce firms to internalise the negative externality, in the form of foregone rents, that they impose on workers on severance.
Chapter 6 concludes.

Contents

List of Figures

To Claudia

Acknowledgments

My greatest debt of gratitude is to Charlie Bean. He has been a constant reference, providing constructive criticism, continuous support and keeping me on track. I have learned a lot from him. My thinking on the issue of dismissal costs has been profoundly influenced by Giuseppe Bertola. I owe him stimulating conversations, a crucial modelling suggestion and much encouragement.

Various people have provided useful comments to previous drafts of some of the chapters. I thank them individually at the beginning of the relevant chapters. Needless to say, I accept full responsibility for all remaining errors.

Other people have contributed with their friendship and encouragement. I would like to mention: Roberto Artoni, Paolo Buonanno, Eliane Catilina, Rita Compostella, Marcello D'Amato, Guido Cervigni, Ozlem Oz, Matteo Piazza, Alessandro Rebucci. But the chief support has been from my wife to whom this thesis is dedicated.

Chapter 1

Introduction

The issue of insufficient labour market flexibility has become prominent in the policy debate of the past ten-fifteen years, both inside and outside the economic profession. The large and persistent increase in unemployment in continental Europe over the past twenty-five years has been taken as a sign that some form of rigidity must be preventing labour markets from achieving equilibrium at unemployment rates in line with those prevailing before the first oil shock[1]. This thesis deals with one particular aspect of the labour market flexibility debate: the effect of economic dismissal costs[2] on quantity flexibility and their consequences for labour turnover, employment and aggregate efficiency.

Recent empirical studies on job and worker turnover have revealed that labour

[1]See Nickell (1997) and Siebert (1997) for two alternative views on this issue. Bean (1994) provides a critical appraisal of the competing (and possibly complementary) views of the causes of European unemployment.

[2]Although in what follows we will use the terms dismissal costs, firing costs and job security generically, the thesis deals mainly with the costs associated with labour force adjustments associated with shocks which are orthogonal to workers' individual characteristics. Chapter 4 discusses the possible costs stemming from the inability of third parties to observe the true cause for dismissal.

markets are characterized by a process of intense reallocation. Each year up to fifteen percent of the total number of existing jobs are destroyed and approximately the same amount are created in OECD countries[3]. Interestingly enough, the data display a substantial uniformity across industrialized countries on both sides of the Atlantic. Even more surprising is that similar figures apply to several semiindustrialized countries, as documented by Roberts (1996) for Colombia, Chile and Morocco. This process of reallocation clearly entails substantial costs, both in the form of unemployment spells and other mobility costs for workers and of set up, hiring and training costs for firms.

In an ideal first-best world, endogenous turnover costs are offset by improvements in match productivity which result from labour reallocation, as shown by Lucas and Prescott (1974). As private and social opportunity costs coincide, all separations are efficient: the individual quest for the highest return ensures that the net social marginal benefit from employment is equalized across productive units. In such a world, the introduction of dismissal costs cannot improve on the decentralized equilibrium. Furthermore, if (endogenous) growth in the economy is driven by a Schumpeterian process of creative destruction by which new, more productive firms displace ageing, less efficient ones, job turnover is the counterpart of the improvement in standards of living. Any obstacle to the relocation of labour from old and declining sectors to new and dynamic ones would depress the growth rate by lowering the return on new productive units.

Whether such an idealized world is a close enough approximation to real-world

[3]See Davis and Haltiwanger (1992) for the US and Garibaldi et al. (1997) for the other OECD countries.

labour markets is a debatable issue. Widespread informational asymmetries and market incompleteness may well prevent mutually-beneficial trade opportunities from being exploited. Empirical and survey evidence lend substantial support to the view that firms may find it profitable to pay above-market-clearing wages[4]. A now sizeable body of literature suggests that realistic labour market frictions and market incompleteness may result in inefficient human capital investment[5]. Dismissal regulations may well be an endogenous second-best response to the costs associated with these kinds of market failures, rather than an exogenous impediment to the efficient working of the labour market.

I will begin by connecting the contents of the thesis with the existing literature on firing costs and with the debate on job security in general.

1.1 Dismissal costs in models of voluntary separation

Concern about the effects of dismissal costs such as mandatory severance payments, advance notice and other administrative and legal costs first arose in the mid-Eighties. By that time, the equilibrium unemployment rate seemed to be reverting towards its pre-oil-shock level in the US, while it was showing no sign of stabilizing in Europe. The quest for an explanation for this asymmetric and unprecedented pattern induced researchers and policy-makers to look for Transat-

[4]Kaufman (1987), Raff and Summers (1987), Krueger and Summers (1988) and Blinder and Choi (1992) provide evidence on the issue.

[5]See, among others, Felli and Harris (1996), Chang and Wang (1996), Acemoglu (1996, 1997).

lantic differences which could account for it. One of these asymmetries is the different degree of employment protection which is high in Europe, but low, on average, in the US.

It was argued (e.g. Blanchard et al. [1985]) that the costs of job security, which had not been perceived in the high growth years before the first oil shock, when most firms were expanding and high attrition rates ensured that dismissals were hardly necessary, had been fully exposed by the slowdown in growth. The restrictive dismissal practices which unions had imposed in Europe in the Sixties and Seventies were now backfiring in the form of high and persistent unemployment. This explains why the literature concentrated on the introduction of an exogenous firing cost in models of voluntary separation.

An accurate statement of the view that dismissal regulations have *caused* a significant fraction of European unemployment could be the following. Legislated employment protection significantly depresses aggregate labour demand, or shifts up the aggregate wage curve, or both, at a given level of employment, after taking into account all general equilibrium effects.

The possibility that dismissal costs can have a significant partial equilibrium impact on aggregate labour demand is easily dismissed. In partial equilibrium, their net effect on average labour demand is small and ambiguous, as argued in Bertola (1990, 1992). On the one hand, dismissal costs intuitively reduce firing by downsizing firms. On the other hand, they also depress hiring by expanding ones, as forward looking firms discount the cost of future workforce reductions. These two effects largely offset each other.

The fact that firing costs do not have significant impact effects on aggregate labour demand does not mean that they are innocuous. First, if they decrease profits they may result in lower rates of creation of new firms and lower rates of investment and growth, as argued in Hopenhayn and Rogerson (1993) and Bertola (1994). Secondly, if they depress turnover they reduce static efficiency, in a first best world, by distorting the cross-sectional allocation of labour.

The important issue is, then, whether dismissal costs do reduce profits and distort labour allocation in general equilibrium[6]. Dismissal costs, it is argued, depress profits through two effects. First, they directly increase total labour costs at unchanged wages. Second, they result in higher insiders' wages by improving the bargaining power of incumbent workers. Yet, if markets are perfect and complete and workers and firms have the same risk-adjusted discount rate, a pure severance payment cannot affect either labour allocation or the present discounted value of firms' profits in models of voluntary separation, as argued by Lazear (1990). Being a pure transfer, redundancy pay does not alter the total surplus from a match and cannot affect the decision to separate or continue the relationship. On the other hand, outsiders should be willing to accept a fall in entrance wages equal to the present value of the increase in future insiders' wages and severance pay, thus leaving the present value of the unit labour cost unchanged.

The only way in which dismissal costs can have real effects is if their impact on wages does not fully offset their partial equilibrium effect. This is clearly the case

[6]To be precise, the narrow notion of general equilibrium that we and most of the literature adopt is one in which wages are endogenous, but other variables such as the interest rate or the determinants of demand are not explicitly modelled.

if either firms and workers have different discount rates or if binding wage rigidities prevent the necessary fall in entrance wages. While the existence of these market imperfections is not unrealistic, it does not imply that dismissal costs necessarily reduce efficiency. If risk-averse workers have a higher risk-adjusted discount rate than firms because they can only imperfectly insure against costly job loss due to asset market imperfections, redundancy payments may well improve on the decentralized equilibrium by inducing efficient risk-sharing, as argued in Bertola (1996a). A similar case, can be made if limited (real) wage flexibility implies that mobility is involuntarily costly for workers, as we argue in chapter 3.

A second possible channel by which dismissal costs can have real effects even in the absence of market imperfections was identified in the wedge between the cost born by the firm and the payment received by workers. In practice, dismissal regulations involve not only a pure redundancy payment from firms to workers, but also red tape administrative costs and other forms of third party payments, such as notice requirements and paperwork, which introduce a wedge between what firms pay and what workers receive. These deadweight loss components of firing costs, it is argued, result in higher relation specific quasi-rents thus decreasing separation rates. Also, given that workers receive less than the actual firing cost paid by the firm, hiring and investment are depressed even in the absence of market imperfections, as entrance wages do not fall by the amount necessary to undo the effect of dismissal costs. Since the frictionless decentralized equilibrium features no distortion apart from the exogenously imposed firing costs, the latter have obvious welfare costs. These and all the main conclusions of the existing literature on the issue are surveyed in chapter 2.

The relevant questions concerning the effect of dismissal regulations are: 1) do dismissal costs affect job destruction and *ex post* profits; 2) do firing costs reduce *ex ante* profits, hence job creation. The second point is relevant only if firing restrictions do reduce *ex post* profits and hinges on the flexibility of entrance wages, as argued above. The first question has received a positive, unconditional answer by the literature. Dismissal costs, it is argued, increase the cost of job termination and result in higher wages by increasing workers' bargaining power. Yet, in models of voluntary and efficient separation it is unclear why it should always be so. If the wage that a worker can receive in her present job is strictly lower than the alternative wage outside, for example, one would expect the worker to quit voluntarily and no firing cost to be paid. The existing literature *assumes* that dismissal costs always result in higher insider wages, by treating them as a tax on separation whoever initiates it. This is clearly not the case in reality.

Chapter 3 takes issue with some of the findings of the existing literature. It analyses the effect of dismissal costs on *ex post* actions and payoffs in a strategic bargaining model. The framework it uses encompasses all existing models of voluntary separation under risk-neutrality and symmetric information. Once the process of bargaining is modelled along the lines of the strategic bargaining literature pioneered by Rubinstein (1982), one can meaningfully distinguish which party gains (*ex ante*) from the end of the match. Firing costs alter the firm's outside option, but not the worker's. Workers do not get them if they stop bargaining and quit to trade with a third party. As argued in Binmore, Rubinstein and Wolinsky (1986), an outside option affects the bargained payoffs only if the threat to take it is credible. So they provide workers with a credible threat only in those cases in which the firm's outside option would be binding in the frictionless equilibrium.

We show that even though, under some circumstances, dismissal costs can have redistributive effects and reduce *ex post* profits, they can never induce inefficient termination in models of voluntary separation. Firing is never an equilibrium strategy in this class of models. So, firing costs alter the equilibrium payoffs only if some exogenous event may force the firm to fire workers despite it being suboptimal, or if separation would allow the firm to free assets which have a positive market value. Intuitively, firing costs increase the specificity of the firm's capital if it decides to trade with a third party. If its assets are already fully specific - i.e. if its outside option gross of the firing cost is zero - under no circumstance will the firm find it profitable to pay the firing cost in order to trade with a third party. In this case dismissal costs cannot affect equilibrium payoffs unless involuntary firing can take place with strictly positive probability.

Even more important is that firing costs do not generate any additional joint quasi-rent from continuation of the match. They have no effect on the separation rate in models in which severance is voluntary in the absence of firing restrictions. Even when they do generate quasi-rents for workers, these rents and, hence, workers' mobility cost are always lower than the statutory firing cost. So, a firm will never dismiss a worker. When the joint surplus from the match is nonpositive, it will rather induce a quit by paying the worker a *voluntary* severance payment equal to her foregone rent. Legislated firing costs, whether received by workers or not, cannot affect the separation decision which is always socially efficient. Hence they cannot have any effect on tenure in models of efficient separation. This is at odds with the empirical evidence of a positive relationship between tenure and dismissal costs.

There is little reason, though, to expect separation decisions to be jointly optimal in the real world, or, equivalently, the quasi-rents from the match to have been completely dissipated by the time severance takes place. Exogenous or endogenous constraints on wage flexibility may well result in involuntary costly mobility for workers. Binding minimum wage constraints, efficiency wage considerations and wage-setting by right-to-manage collective agreements, all entail inefficient separation.

Real wage rigidity implies that, in the absence of dismissal costs, severance takes place whenever the firm's payoff is non-positive whether or not the joint return within the match is strictly lower than the joint return from separation. Dismissal costs, whether fully received by workers or not, reduce the firm's return from separation and result in labour hoarding and longer tenure. Provided they are high enough, they induce the firm to terminate the relationship efficiently, when the joint rent from the match is zero. We show that in a simplified version of the exogenous mobility cost model analysed in Bertola and Ichino (1995b), a pure severance payment unambiguously increases aggregate welfare if exogenous wage compression results in involuntarily costly workers' mobility.

The existing empirical evidence supports the relevance of this interaction between real wage rigidity and dismissal regulations. Nickell (1998) finds a strong positive correlation in a cross-country regression between an index of job security and workers' tenure. Our result implies that tenure and, consequently, human capital investment are not efficient under "employment-at-will".

1.2 Dismissal costs and labour market imperfections

In a first-best world, dismissal costs, if they have any effect at all, can only redistribute total returns from firms to incumbent workers. Even when dismissal costs improve on a decentralized allocation which features excessive turnover, they do reduce firms' profits in the stripped-down set up analysed in chapter 3. This is at odds with the observation that some firms do offer forms of job security over and above statutory minima. Apart from Japanese firms, a significant number of US firms such as DEC, IBM, Kodak, Eli Lily are committed to a no-firing policy. In the UK, a sizeable proportion of firms contractually commit to severance payments significantly in excess of legislated minima, as documented in Millward et al. (1992). Since a great number of these firms, especially in the US, are non-unionized, this is difficult to explain with Booth's (1997) otherwise appealing idea that bargaining over dismissal costs is a substitute for efficient bargaining over wages and employment in situations in which wage-setting by right-to-manage collective agreements results in inefficient separation *ex post*.

The industrial relations literature has often emphasized the role that job security can play in improving workers' cooperation and boosting productivity or reducing costs in union and non-union firms alike. In the US there has been a long standing debate about whether excessive labour market flexibility could be responsible for the poor growth performance of the Eighties and early Nineties.

The mechanism through which job security may improve firms' profitability that have been suggested are: a reduction in wage costs if more risk-averse workers accept wage reductions in exchange for job security, higher investment in human capital, higher effort and lower worker-initiated turnover. The first aspect has already been emphasized by the implicit contract literature and has recently been restated in a fully dynamic context by Bertola (1996a).

Chapter 4 builds on the results in chapter 3 and explores the effect of job security measures on firms' investment in general training in the presence of frictional unemployment. The model takes as a stylized fact firms' investment in training. Search frictions imply that, in the absence of contracts, the level of investment is inefficiently low due to hold up.

The literature on incomplete contracts has explored various possible solutions to hold up. Differently from other types of investment discussed in this literature, employer-provided general training is productive outside the current relationship, but vested in the non-investing party on separation. So, no return accrues to the investing firm when the match is destroyed. Search frictions imply that the return is shared between the worker and the *future* employer. Acemoglu (1997) has demonstrated that this spillover onto future employers cannot be internalized in the *laissez-faire* equilibrium, since the identity of future employers is unknown at the time of investment. We show that an appropriate redistribution of property rights can reduce the other type of spillover - that onto the worker - thus improving firms' incentives to train. Consensual layoff measures achieve the required reallocation of property rights. By preventing employers from unilaterally terminating the relationship, consensual layoff measures oblige the investing firm to bargain over the total payoff from separation, that is over the transfer which induces workers to accept severance. This sharing of the joint separation payoff implies that firms capture part of the *marginal* return to training that accrues to the worker, thus improving ex ante incentives to invest.

Interestingly, there exist real world institutions that resemble the kind of optimal arrangements highlighted. Social plans legislation in Germany and other continental European countries prescribes that firms cannot initiate mass redundancies unless they have agreed with workers' representatives on the details of the procedure and compensation packages. As mentioned above, a number of firms contractually commit to a zero-firing policy that effectively prevents them from laying off workers unless by mutual consent. The institution of lifetime employment in Japan has the same effects. We also show that large enough severance payments can achieve the same result. The result is not restricted to firm-provided training. It applies to all forms of general investment which is vested in the non-investing party. So, by the same token consensual layoff measures also increase workers' investment in activities that increase firms' goodwill, such as effort to improve product quality or to develop products that remain the intellectual property of firms.

The model also highlights a kind of inefficiency which is not present in search models with homogeneous workers. It is well known that in those models the decentralized equilibrium is efficient, provided the share parameter satisfies the Hosios (1990) condition, thus balancing the congestion and thick market externalities. In our model, both trained and untrained workers coexist and the share parameter alone cannot ensure efficiency on both the job creation and job destruction margins. Such a minor and realistic deviation from the usual set up highlights the fragility of the efficiency result in search models.

The other mechanism through which dismissal costs can increase firms' profitability is by reducing turnover and improving workers' effort. The turnover and effort arguments are usually associated with efficiency wage theories. Efficiency wage considerations are the most widely accepted candidate for the existence of real wage rigidity. Apart from the empirical study of Krueger and Summers (1988), the relevance of efficiency wage consideration is supported by the case-study and survey evidence in Raff and Summers (1987), Kaufman (1987) and Blinder and Choi (1992).

Chapter 5 analyses the equilibrium effect of dismissal costs in a dynamic version of Shapiro and Stiglitz (1984) shirking model. We endogenize the separation rate by assuming that firms' productivity is subject to idiosyncratic uncertainty. Since mobility has to be costly for unemployment to be a credible punishment, redundant workers bear an involuntary mobility cost. This raises incentive compatible wages at expanding firms, as workers have to be compensated for the expected mobility cost. Firms control the firing probability and can reduce their wage bill by committing to fire less in downturns. Yet, at the time of firing, their pledge is not time-consistent as the benefit from lower wages is sunk. Dismissal costs not only unambiguously increase aggregate employment, but allow firms to credibly commit to fire less in downturns. Most importantly, a pure severance payment unambiguously increases both the value of expanding firms and aggregate welfare.

In the absence of dismissal costs or other commitment or reputational devices, the externality, in the form of foregone rents, that firms impose on workers cannot be traded. Workers' moral hazard prevents it to be traded *ex post,* while the firm's moral hazard rules out *ex ante* trading. Severance payments effectively work as a Pigovian tax on firing and induce firms to internalize the externality. Welfare may still increase even if only part of the cost born by the firm accrues to workers.

Chapter 2

The existing literature

2.1 Introduction

This chapter is a survey of the main results of the existing literature on dismissal costs. This line of research has been mainly concerned with the employment and welfare implications of the introduction of an exogenous firing cost in models of voluntary separation. While, on the whole, it has little to say on the possibility of a welfare-enhancing role for job security[1], the existing literature has explored at length the mechanisms through which dismissal regulations may affect actions and payoffs. This survey does not attempt to be exhaustive, but rather to highlight these mechanisms in an intuitive way. For this purpose, we construct a simple dynamic competitive model and analyse the effects of a separation tax on the equilibrium allocation.

[1]The only exceptions are Bertola (1996a), which explores the possibility of efficient risk-sharing and Booth (1997). Otherwise, in so far as separation decisions are assumed to be jointly-efficient for the firm-worker pair in the absence of dismissal regulations, any positive effect of job security on aggregate welfare is apriori ruled out. In the absence of externalities, private and social opportunity costs coincide and the separation decision is socially optimal.

The model reproduces all the main findings on the existing literature on firing costs. A separation tax has small and ambiguous effects on aggregate labour demand, as it depresses both job creation and job destruction by reducing the forward-looking shadow cost of labour at contracting firms and increasing it at expanding ones. On the other hand, the part of the tax which accrues to workers on separation pushes up the wage of the workers who are entitled to it, by increasing their outside option, and reduces the wage of newly hired workers who accept a lower entrance wage in exchange for a higher pay in the future. These two effects increase the shadow cost of labour at firing firms and reduce it at hiring ones thus offsetting the partial equilibrium impact. If the separation tax is a pure transfer from firms to workers and markets are complete and perfect, then the tax has no effect on the shadow cost of employment and labour allocation. It just redistributes workers' and firms' payoffs across time.

There is no reason to expect this idealized world to be an accurate description of reality: asset market imperfections may imply that in equilibrium marginal rates of substitution differ across agents, or exogenous wage rigidities may prevent competition among workers from eliminating rents to job seeking, or the tax may involve a deadweight loss and the cost to the firm of severing the employment relationship may exceed the payment which accrues to the worker. In all these instances a separation tax does have real effects. While in the case of incomplete insurance markets it may actually increase aggregate welfare by providing insurance against income uncertainty, in the other two cases it unambiguously reduces efficiency as it distorts the cross-sectional labour allocation, depresses firms' profitability and investment.

One should be somewhat suspicious that a tax on separation, whoever initiates it, has exactly the same effects that the existing literature attributes to dismissal costs. This implies that either firing costs work in the same way as our tax on separation or that they have been treated as if they do. A moment's reflection suggests that the latter has to be the case. In reality, firing costs are not due if workers voluntary quit. So workers cannot credibly threat to quit if their pay does not reflect the dismissal cost. Secondly, firing costs are not a tax in so far as no third party derives any revenue from them. Feldstein (1976) has shown how workers and firms have an incentive to exploit third-party payments when these take the form of an unemployment subsidy through an imperfectly experienced-rate unemployment insurance. Symmetrically, in the case of firing costs, the parties have an incentive to find an agreement to label separations as quits and economize on deadweight losses associated with job security. We address these issues in the following chapter.

The present chapter is structured as follows. Section 2.2 describes the economic environment. Section 2.3 discusses the partial equilibrium effects of a tax on separation. Section 2.4 analyses its impact on wages under perfect competition. Section 2.5 is concerned with the conditions under which a tax on separation has real effects in equilibrium. Section 2.6 concludes.

2.2 Economic environment

The following stripped down economic environment provides a minimum setup within which to review the results of the existing literature on firing costs under risk-neutrality and symmetric information.

We consider a labour market with a homogeneous and infinitely divisible labour force whose size is normalized to one. Time is discrete. Workers are infinitely lived and endowed with one indivisible unit of labour. They can either choose to be unemployed and earn a non-market return equal to z units of the unique good produced within the economy[2] or work at the going wage. Mobility both across jobs and in and out of unemployment involves no cost for workers. Both product and labour markets are perfectly competitive. On the demand side, there is a continuum of firms indexed by n, where $n \in [0,1]$. They produce a homogeneous good whose price is normalized to one. Though *ex ante* identical, firms are *ex post* heterogeneous, as they are subject to idiosyncratic exogenous shocks to their production function

$$R(l_t^n; \alpha_t^n),$$

where l_t^n is the employment level at firm n at time t and α_t^n is a, publicly observable, *portmanteau* index for the state of firm-level productivity. Higher values of α_t^n are associated with higher productivity. The production function is assumed to be strictly concave in l_t^n, so that the marginal product of labour $R'(l_t^n; \alpha_t^n)$ is

[2] z could also be interpreted as the flow utility of leisure.

decreasing in employment.

As a useful benchmark, we assume away any capital market imperfection. Agents can freely borrow and lend at a common interest rate, subject to solvency. Firms, then, choose employment to maximize the present discounted value of profits net of turnover costs at any instant. In the absence of turnover costs, labour demand would be given by the textbook static condition that the marginal product of labour equals the current (real) wage, i.e. $R'(l_t^n; \alpha_t^n) = w_t^n$.

We assume that a tax is levied on any separation which involves a worker with at least one period of tenure. For any such worker that leaves her current job for whatever reason, the firm has to pay a time-independent tax equal to F units of output. Part or all of the tax can be transferred to the worker. Q is the part of the cost F born by the firm which accrues to the worker. The excess $F - Q$, if positive, is a pure deadweight loss. As we will argue in chapter 3, we believe this way of modelling dismissal costs is misleading. In the real world firing costs differ from the tax on separation just described in so far as they are due only in case a firm makes workers redundant, but not if workers voluntary quit their employer. Yet, our purpose here is to show how some crucial conclusions of the existing literature on the effects of dismissal costs rely upon the assumption that they effectively work as a tax on mobility. The assumption that the firing tax is paid only on separations that involve workers with at least one year of tenure is necessary (and sufficient) to ensure the existence of a well-defined competitive equilibrium. Since workers bear no mobility cost, if however small a part of the tax on separation accrued to them ($Q > 0$), independently of tenure, they could increase their lifetime wealth unboundedly by reshuffling themselves across jobs at an infinite rate and no firm would enter the market.

Allowing for the cost born by the firm to exceed the payment which accrues to workers is meant to capture the fact that dismissal regulations often involve third party payments in the form of red tape costs such as advance notice requirements, paperwork and other administrative costs.

A constant unit tax on separations is equivalent to a linear adjustment cost. This is desirable for two reasons. First, in many countries dismissal costs take the form of a notice period and payment for each worker made redundant. Secondly, linear adjustment costs realistically imply discrete rather than continuous labour force adjustment which is consistent with the empirical evidence[3].

In a competitive labour market firms have nothing to gain from committing on employment levels[4], so they adjust their labour force optimally after observing the current shock. Optimality requires the shadow value of employment J_t^n to satisfy the Bellman equation

$$J_t^n\left(l_t^n, w_t^n, \alpha_t^n, r_t, t, ...\right) = R'(l_t^n; \alpha_t^n) - w_t^n + \frac{1}{1+r_t} E_t\left[J_{t+1}^n\right] \tag{2.1}$$

at any instant t. The shadow value of the marginal worker equals the instantaneous marginal profit plus the expected future value discounted at the market cost of funds to the firm. In general, the shadow value of labour depends not only on the

[3]Evidence in support of lumpy labour force adjustment is provided by Hamermesh (1989) and Holtz-Eakin and Rosen (1991).

[4]If firms are wage-setters they may reduce their wage bill by committing to fire less in downturns, as in Saint-Paul (1995a).

whole path of current and future realizations of state and control variables, but also on past realizations, such as the inherited labour force, insofar as these affect future profits and the conditional expectation $E_t[.]$ in equation (4.36).

It is possible to restrict the state space in a way which ensures that the conditional probability distribution of future realizations is fully characterized by their current value. The simplified stochastic structure makes the problem much more tractable to the benefit of economic intuition at little cost in terms of rigour. Following Bertola (1990) we assume a time-invariant discount rate r and a Markov process in levels for the shocks. The process is a two-state Markov chain with symmetric transition probabilities[5] given by

$$\alpha_{t+1}^{n} = \begin{cases} \alpha_g \,\text{with prob. } p \text{ if} \alpha_t^n = \alpha_b, \text{ with prob. } (1-p) \text{ if} \alpha_t^n = \alpha_g \\ \alpha_b \,\text{with prob. } p \text{ if} \alpha_t^n = \alpha_g, \text{ with prob. } (1-p) \text{ if} \alpha_t^n = \alpha_b. \end{cases}$$

At time t, firms are either in the good state and enjoying high productivity - $\alpha_t^n = \alpha_g$ - or in the bad state characterized by low productivity - $\alpha_t^n = \alpha_b$ - where $\alpha_g > \alpha_b$. Within each group, all firms are identical, but, over time, they switch between states according to the above Markov process. Since shocks are independent across firms and the number of firms is infinite, the cross-sectional conditional and unconditional distributions coincide with the conditional and unconditional distributions of shocks for the individual firm. The proportion of firms in each state which are hit by a shock coincides with the conditional probability

[5] Allowing for asymmetric transition probabilities would be straightforward, but would just complicate the algebra with little gain in economic insight.

p that a firm currently in the bad (good) state will experience a positive (negative) productivity shock in the next period. Also, since the ergodic unconditional probability that a given firm is enjoying good or bad business conditions is 0.5, this is also the steady state proportion of firms in each state. Given that the total number of firms has been normalized to one, this means that in every time interval $p/2$ firms move from the bad to the good state and an equal number faces the opposite transition. p is also a measure of shock persistence. Shocks are permanent if $p = 0$, while they are serially uncorrelated when $p = 0.5$.

In the presence of turnover costs, firms should increase their employment as long as the expected present value of profits associated with the marginal worker exceeds the hiring cost. Conversely, they should reduce their labour force whenever the shadow value of labour falls below (minus) the firing cost, as the associated increase in the firm's value exceeds the cost of reducing the labour force at the margin. If turnover costs are linear, labour demand is unaffected by small changes in the shadow value of labour, as for small enough shocks the strictly positive marginal cost of adjustment exceeds the opportunity cost of inaction. We assume that the change in business conditions is large enough to generate positive turnover. In our two state world, this implies that firms are always indifferent (at the margin) between keeping employment constant and hiring a new worker, if in the good state, and between retaining and firing employees, in the bad state.

2.3 Labour demand

In the present model shocks to labour productivity are the only source of uncertainty. So workers have no incentive to quit[6] a firm which pays the market wage and firms adjust employment only if they undergo a state transition. With shocks to productivity oscillating between two states, the shadow value of labour can take at most three values. It is J_g^j, if the firm has just transited into the good state and increased its labour force. In this case the marginal worker has less than one period of tenure (is "junior"). Alternatively, a firm has not hired new workers in the current period and employs only workers with at least one period of tenure ("senior")[7]. The shadow value of employment is either J_g^s and J_b^s depending on whether the firm is in the good or the bad state. Distinguishing between firms in the same state according to the tenure of their marginal worker is necessary if the wage-tenure profile is not flat.

The shadow value of labour changes from J_g^j or J_g^s to J_b^s with probability p, the conditional probability that a firm moves from the good to the bad state and reduces its labour force. With the same probability p the contribution of the *currently* marginal worker to the firm's value transit from J_b^s to J_g^s in response to a positive shock[8]. With the complementary probability $(1-p)$ the firm's business conditions do not change and either the shadow value of employment stays the

[6]The effect of the endogeneity of quits is analysed in Saint-Paul (1995b). He shows how the procyclicality of quits may result in multiple, Pareto rankable, equilibria.

[7]We terms "junior" and "senior" workers are used rather than the common "outsider" and "insider" to emphasize that, differently from the literature pioneered by Lindbeck and Snower (1988), workers have no market power in the present setup.

[8]The transition is to J_g^s rather than J_g^j as the currently marginal worker becomes intramarginal when productivity improves and the firm hires.

same, if the firm was already employing only senior workers, or it moves from J_g^j to J_g^s for those firms that expanded their labour force in the previous period and whose junior workers achieve tenure. So, the optimality condition (4.36) can be rewritten as

$$J_g^j = R'(l_g; \alpha_g) - w_g^j + \frac{1}{1+r}\left[(1-p)J_g^s + pJ_b^s\right], \tag{2.2}$$

$$J_g^s = R'(l_g; \alpha_g) - w_g^s + \frac{1}{1+r}\left[(1-p)J_g^s + pJ_b^s\right] \tag{2.3}$$

and

$$J_b^s = R'(l_b; \alpha_b) - w_b^s + \frac{1}{1+r}\left[(1-p)J_b^s + pJ_g^s\right] \tag{2.4}$$

where w_g^j is the wage of junior workers at hiring firms and w_g^s and w_b^s are the wages of senior workers at firms in the good and the bad state respectively.

Since, hiring is costless it must be $J_g^j = 0$: firms increase their labour force as long as the contribution of an additional worker to the firm's value is positive. Conversely, firms have an incentive to cut their labour force in response to a negative shock until the saving made by firing one more worker exceeds the dismissal cost. So, it must be $J_b^s = -F$.

One will have noted that we have let the employment levels in equations (2.2)-(2.4) take only two values - l_g and l_b - thus assuming that they depend only on the state of productivity and not on the wage of the marginal worker. For an arbitrary wage-tenure profile this may not necessarily be the case. Optimality may require firms in the good state to adjust employment after one period, even at unchanged business conditions, if the wage differential between junior and senior workers is large enough to drive the shadow value of labour above zero or below $-F$. As we show in the next section, our assumption turns out to be verified in equilibrium: the equilibrium wage-tenure profile never justifies adjustment in the absence of shocks. This implies that equations (2.2) and (2.4) determine labour demand respectively in the good and the bad state, while (2.3) determines the shadow value of senior workers in the good state J_g^s. This equals $J_g^j - (w_g^s - w_g^j)$, as can be seen by subtracting equation (2.2) from (2.3). Remembering that $J_g^j = 0$, it is $J_g^s = -(w_g^s - w_g^j)$. If the marginal product of labour is unchanged, changes in the shadow value of employment can only stem from changes in wages.

Replacing the values of J_g^j, J_g^s and J_b^s in equations (2.2) and (2.4) results in the two hiring and firing conditions

$$R'\left(l_g; \alpha_g\right) = w_g^j + \frac{p}{1+r}F + \frac{(1-p)}{1+r}\left(w_g^s - w_g^j\right), \tag{2.5}$$

$$R'\left(l_b; \alpha_b\right) = w_b^s - \frac{(p+r)}{1+r}F + \frac{p}{1+r}\left(w_g^s - w_g^j\right). \tag{2.6}$$

Let us abstract for a moment from the wage differential. Equations (2.5) and (2.6) show that firing costs introduce a wedge between the wage and the marginal product of labour. In hiring firms the wage is lower than marginal productivity while the opposite is true in firing firms. For a given wage, firing costs reduce employment in the former and increase it in the latter through the well known option-value effect. The value of the option is higher the lower is the degree of shock persistence (the higher is p). The more likely a shock is to be reversed in the next period, the higher is the expected saving from delaying adjustment. Furthermore, dismissal costs have a more obvious depressing effect on firing, the term $rF/1+r$ in equation (2.6), as layoffs involve a direct cost in the present. The last terms in equation (2.5) and (2.6) imply that, as continuous adjustment is not optimal in the presence of turnover costs, firms must take into account the steepness of the wage-tenure profile. Hiring is depressed, if the latter is increasing, and boosted if it is decreasing. Conversely, job destruction is increased if entrance wages are lower than senior workers' pay, as firms take into account that firing one more worker implies the possibility of replacing it with a "junior" worker and save on wage costs in case business conditions improve in the future.

A tax on separation increases labour demand at downsizing firms, but reduces it at hiring ones. So it reduces labour turnover. Yet, less intense labour reallocation does not have to be necessarily associated with a higher number of unemployed workers. As the long run equilibrium number of firms in each state is 0.5, aggregate labour demand equals

$$L = \frac{l_b + l_g}{2}. \tag{2.7}$$

So the net partial equilibrium effect of dismissal costs depends on whether they reduce employment at hiring firms more than they increase it at firing ones or viceversa. Bertola (1990, 1992) has forcefully argued that whichever the direction of the net effect it is likely to be small. It is easy to see why. Suppose that firms pay a state independent wage $\bar{w}$. The labour demand equations (2.5) and (2.6) become

$$R'(l_g; \alpha_g) = \bar{w} + \frac{p}{1+r} F \tag{2.8}$$

and

$$R'(l_b; \alpha_b) = \bar{w} - \frac{(p+r)}{1+r} F. \tag{2.9}$$

Apart from signs, the only source of asymmetry between equation (2.8) and (2.9) is the discount rate r in the numerator of the second addendum on the right hand side of (2.9) and the steepness of the marginal product curve. If the marginal productivity curve has a constant slope (e.g. it is linear and subject to additive shocks), asymmetric employment effects can only stem from different changes in the shadow cost of labour, the right-hand-side of (2.8) and (2.9), with respect to the frictionless equilibrium. If $r > 0$, dismissal costs reduce the shadow cost of labour at firms in the bad state by more than they increase it at firms in the good one and aggregate demand increases. This is illustrated in figure 2.1, where l_b and l_g stand for employment in the two states in the absence of frictions, when the marginal product of labour coincides with the wage. Aggregate labour demand is given by L, the midpoint between l_b and l_g. Dismissal costs drive a wedge between the wage and marginal productivity and the new employment levels are given by l_b' and l_g'. Average labour demand increases to L'. This effect, though, is small for realistically small values of the discount rate. On the other hand, a given change in the shadow cost of labour will result in a bigger change in labour demand the flatter the marginal product curve is. So for firing costs to reduce aggregate labour demand, the marginal product curve must be flatter in good states. In

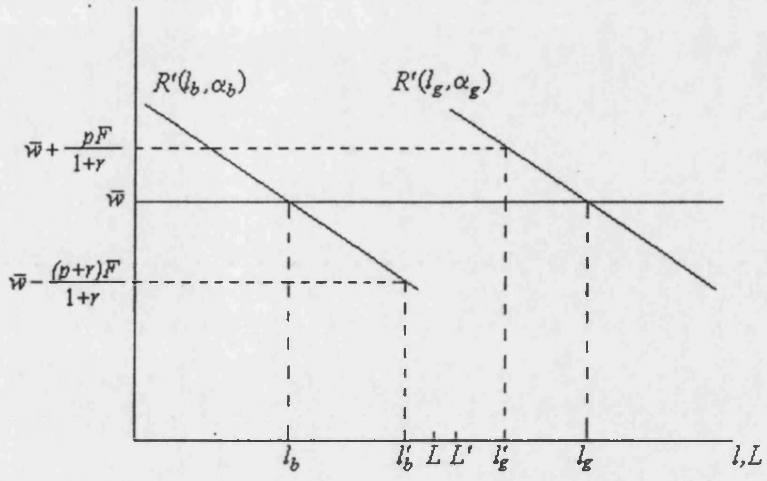

Figure 2-1: Effect of dismissal costs on average labour demand with constant slope marginal product and $r > 0$.

the case of a Cobb-Douglas production function subject to multiplicative shocks, for example, firing costs result in lower average labour demand if the discount rate equals zero. This is illustrated in figure 2.2, where aggregate demand falls from L to L'. Bertola (1992) proves that, even in the extreme case in which firing costs completely prevent labour force adjustment, this effect has the same order of magnitude as Jensen inequality effects and is, hence, second order with respect to the change in firm-level employment. Allowing for a positive discount rate further dampens this negative effect and may even reverse it. Bentolila and Bertola (1990) calibrate a partial equilibrium model with non-stationary shocks and Cobb-Douglas production function and find that firing costs have a small positive impact on average labour demand[9].

[9] Bentolila and Saint-Paul (1994) allow for firms' exit and find that the partial equilibrium effect of firing costs, though limited, may be non-monotonic. Small costs may depress average labour demand, but large ones increase it.

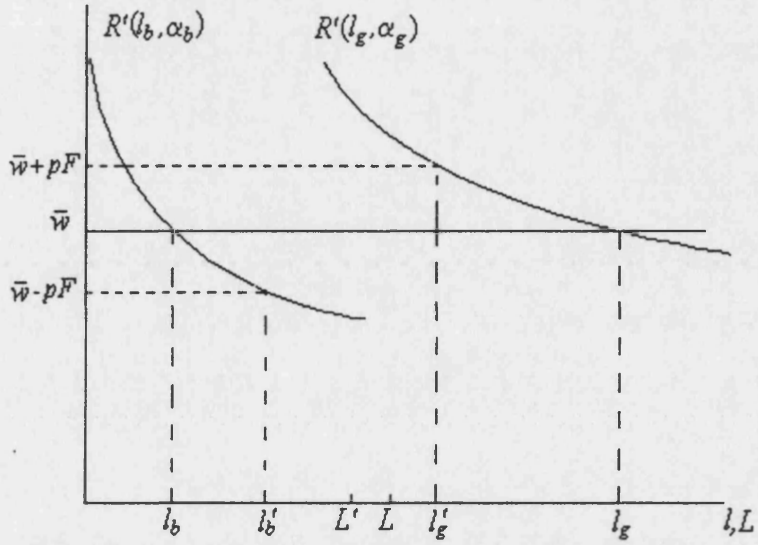

Figure 2-2: Effect of dismissal costs on average labour demand with Cobb-Douglas technology, multiplicative shocks and $r = 0$.

There is clearly a second channel through which a dismissal costs can affect employment. By reducing the market value of firms firing costs should reduce firms' entry. Yet, the effect of dismissal costs on the value of firms cannot be properly assessed without taking into account how wages respond and it is to this that we now turn.

2.4 Wage determination

Workers' labour supply decisions maximize expected lifetime utility subject to the intertemporal budget constraint. Under our maintained assumption of risk-neutrality, lifetime utility coincides with the infinite stream (discounted at the subjective discount rate ρ) of labour and non-labour income. In fact, if utility is linear (affine), workers have no incentive to reallocate consumption over time, as long as their subjective discount rate ρ is no larger (smaller) that the rate of interest at which they can borrow (lend)[10]. Indexing workers by i, their maximized lifetime labour[11] income W_t^i needs to satisfy the Bellman equation

$$W_t^i\left(w_t^i, \rho, \ldots\right) = w_t^i + \frac{1}{1+\rho} E_t[W_{t+1}^i] \tag{2.10}$$

where w_t^i is the flow return the worker is earning in period t.

In the frictionless competitive equilibrium, firms must pay identical wages to identical workers, as workers can move across jobs at no cost and would arbitrage away any pay differential. So, in the absence of turnover costs, the equilibrium wage is the higher between the market clearing wage and the return to non-market activity z. In the second case, the market clearing wage lies below what workers could earn as self-employed and part of the labour force is voluntary unemployed, as jobs are rationed at wages at which workers are willing to work. It is convenient, though not at all necessary, to assume that the return to non-market activity z is high enough to ensure the existence of a positive stock of (voluntarily) unemployed workers, so that wages are fully determined by the supply side of the market.

In the present setup all unemployment is voluntary and, with no labour force growth, the stock of employed workers is constant in steady state or, equivalently, job creation exactly equals job destruction. It is, then, convenient to assume that unemployed workers never accept employment, while all job creation and

[10]If risk-neutral workers could borrow (lend) at a market rate lower (higher) than their subjective discount rate, they would concentrate consumption of maximized resources at the end (beginning) of their lifetimes. As noted in Bertola (1996b), equilibrium considerations rule out such degenerate cases.

[11]Since workers are indifferent about the timing of consumption under our assumptions, optimal labour supply decisions are independent from non-human-capital endowments.

destruction is accounted for by workers' churning. So, employed workers transit across three states: employed with no tenure at a firm which is enjoying good business conditions, employed with at least one-year tenure at a firm in the same state, employed with at least one-year tenure at a firm in the bad state. Let us denote by W_u the expected lifetime human capital of an unemployed worker, while W_g^j, W_g^s and W_b^s represent lifetime labour income for respectively a junior and senior worker at a firm in the good state and a (senior) worker at a firm in the bad state. The Bellman equations corresponding to expected wealth in each state are

$$W_g^j = w_g^j + \frac{1}{1+\rho}\left[(1-p)W_g^s + p\frac{l_b}{l_g}W_b^s + p\left(1-\frac{l_b}{l_g}\right)\left(W_g^j + Q\right)\right], \qquad (2.11)$$

$$W_g^s = w_g^s + \frac{1}{1+\rho}\left[(1-p)W_g^s + p\frac{l_b}{l_g}W_b^s + p\left(1-\frac{l_b}{l_g}\right)\left(W_g^j + Q\right)\right], \qquad (2.12)$$

$$W_b^s = w_b^s + \frac{1}{1+\rho}\left[(1-p)W_b^s + pW_g^s\right], \qquad (2.13)$$

where the terms in square brackets represent the current expectation of next-period wealth. With conditional probability $(1-p)$ the firm's business conditions are unchanged from the following period. In this case only junior workers's lifetime income changes, since they will have now become senior workers and their wealth will now be W_g^s. With probability p the firm changes state. If it transits from the bad to the good state (equation 2.13), it keeps all its workers and their expected wealth becomes W_g^s. If the firm undergoes the opposite transition (equations 2.11 and 2.12), it will fire a proportion l_b/l_g of its workers. Assuming random dismissals, workers face a probability pl_b/l_g of seeing their wealth change to W_b^s and a probability $p(1 - l_b/l_g)$ of being dismissed and finding employment at expanding firms. As tenure is lost on separation, their lifetime income becomes W_g^j, but they receive a severance payment Q.

The expected wealth of an unemployed worker coincides with the present discounted value of lifetime income from self-employment, i.e.

$$W_u = \frac{(1+\rho)}{\rho} z. \tag{2.14}$$

Since the only firms that hire new workers in this model are firms which have been hit by a favourable shock, voluntary unemployment implies that the expected utility of a junior worker is given by

$$W_g^j = W_u. \tag{2.15}$$

If the permanent income of newly employed workers were higher then the permanent return from unemployment, workers would be leaving unemployment at an infinite rate.

The lifetime income of a senior worker, whether employed at a firm in the good or the bad state, is given by

$$W_g^s = W_b^s = W_u + Q. \tag{2.16}$$

Since the tax is due independently from the reason for separation, by quitting their firm, senior workers could get the payment Q and enjoy a level of wealth equal to W_u. So firms need to pay a wage which ensures at least the same lifetime return.

Equilibrium wages must satisfy the optimality equations (2.11)-(2.13) and the no-arbitrage conditions (2.15) and (2.16). Replacing the latter into the former results in

$$w_g^j = z - \frac{1}{1+\rho} Q, \tag{2.17}$$

$$w_g^s = w_b^s = z + \frac{\rho}{1+\rho} Q. \tag{2.18}$$

If workers were not entitled to any payment in case of separation ($Q = 0$), wages would be equalized across states - as mobility is costless for workers - and would coincide with the opportunity cost of labour z. A positive separation payment, instead, implies a wage differential $w_i^s - w_g^j = Q$ (where $i = b,\ g$) between senior and junior workers. Effectively, with a positive severance payment Q senior workers' mobility is subsidized and the wage differential must compensate them for foregoing the subsidy. So their wage w_i^s is increased by the flow equivalent of the severance payment they could obtain by quitting. Conversely, the wage of junior workers w_g^j falls by the present value of the rent they get after one period. The rent accrues to them unconditionally in the form of either a higher wage or a severance payment in case of dismissal. The only effect of the severance payment Q is to tilt the wage-tenure profile upwards, while leaving the lifetime income of junior workers unchanged[12].

[12]Equivalently, upon its introduction the firing tax results in a windfall gain for employed workers, but has no effect on the lifetime utility of unemployed and newly hired workers.

It is apparent from equation (2.17) that, for large enough values of the payment Q, the entrance wage may be negative. In the presence of exogenous wage rigidities such as minimum wage legislation, the wage may be prevented from falling by the full amount of the future rent even if workers would be willing to accept the wage cut in exchange for the higher future return. On the other hand, the size of the fall in the entrance wage depends on the extreme assumption that workers are entitled to a payment after only one period of employment at the same firm. In most countries it takes more time for workers to qualify for severance payment at all and the payment itself is a function of seniority. So the necessary fall in the entrance wage is likely to be significantly smaller.

It should be noted that the fact that the wage differential $w_g^s - w_g^j = Q$ implies $J_g^s = -Q$ and, since $Q \leq F$, $-F \leq J_g^s \leq 0$. As argued in the previous section, the firm does not change its labour force in the absence of shocks.

2.5 Equilibrium

The market equilibrium conditions can be recovered by replacing the equilibrium wages (2.17)-(2.18) in the labour demand equations (2.5)-(2.6). With some manipulation they can be written as

$$R'\left(l_g; \alpha_g\right) = z + \frac{p}{1+r}\left(F - Q\right) + \left(\frac{1}{1+r} - \frac{1}{1+\rho}\right)Q \qquad (2.19)$$

and

$$R'(l_b;\alpha_b) = z - \frac{(p+r)}{1+r}(F-Q) - \left(\frac{r}{1+r} - \frac{\rho}{1+\rho}\right)Q. \qquad (2.20)$$

Equations (2.19) and (2.20) implicitly give the equilibrium employment levels at firms in the good and the bad state. It is apparent that the equilibrium allocation in the presence of a firing tax can differ from the frictionless one in which the marginal product of labour coincides with its opportunity cost z for only two reasons. Either the firing tax does not fully accrue to workers $(F > Q)$[13] or workers and firms discount the future at different rates $(\rho \neq r)$. As argued in Lazear (1990), if capital markets are perfect and agents have the same, risk-adjusted, discount rate, the only effect of a pure severance payment $(F = Q)$ is to redistribute firms' costs and workers' return across time with no real effect. The equilibrium allocation coincides with the Pareto optimal allocation that would prevail in a frictionless market.

This conclusion does not rely in any way on the assumption that both parties lack market power. Market power can only affect the size of the rent which accrues to senior workers. The intuition behind the result is that, provided workers are not entitled to any severance payment at the moment a match is formed, the higher future wage bill is fully offset by the fall in the entrance wage, as unemployed workers bid wages down by the full present value of the future rent. Bertola (1990) obtains the same result in an atomistic insider-outsider model. Burda

[13] Theoretically, all one needs is that F differs from Q. This may be the case, for example, if unemployment benefits are imperfectly experienced rated, as first noted by Feldstein (1976). A careful treatment of this case, though, would require modelling how the excess of Q over F is financed and lies outside the scope of this survey.

(1992) and Cabrales and Hopenhayn (1996) have shown that the same outcome obtains in a version of Mortensen and Pissarides (1994) matching model in which wages are determined by bilateral bargaining. As long as the tax on separation is realistically zero when a match is created, it cannot affect the parties' shares of the surplus, as these depend on *ex ante* bargaining power. Our tax on separation does increase (reduce) *ex post* workers' (firms') bargaining power and result in an increasing wage-tenure profile for workers, but if both firms and workers have rational expectations about the process for wages, the entrance wage would ensure a division of the surplus according to *ex ante* bargaining power.

This result makes clear that a pure severance payment can affect labour allocation only in two cases. The first, which we have already discussed in the previous section, when constraints on downward wage flexibility, such as minimum wage legislation or insiders' setting of outsiders' pay, prevent entrance wages from falling by the necessary amount. In this case, a firing tax does not affect employment at firing firms, but results in involuntary unemployment. Unemployed workers are prevented from accepting employment at a wage at which they would be willing to and hiring firms create less jobs, as the tax increases the present value of the marginal worker's wage bill. Alternatively, a pure separation-contingent transfer may have real effects if the assumptions of perfect and complete markets fail to be satisfied. Workers' effective discount rate can differ from the firms' market cost of funds either because of capital market imperfections or because, with incomplete asset markets, risk-averse workers may not be able to fully diversify labour income uncertainty and would not behave *as if* risk-neutral, thus discounting high wages more heavily than low ones[14]. With respect to the case in which non-neutrality was driven by downward wage rigidity both hiring and firing would be affected as can be seen from equations (2.19) and (2.20).

While asset market incompleteness may have little bearing for the risk-neutral agents in the model's economy, in general it implies that there is no presumption that the frictionless decentralized allocation is efficient to start with. Yet, even allowing for risk-aversion, it is easy to see why a pure severance payment could only reduce aggregate welfare in the present set up where mobility is costless for workers. The frictionless equilibrium features no fluctuation in labour income. Redundancy payments, by resulting in positive wage differentials, are inefficient, as they introduce income uncertainty which, in the absence of complete contingency markets, can only be imperfectly diversified. This conclusion, though, rests on the assumption that dismissal costs are the only source of income uncertainty.

In reality, job switching does involve income fluctuations in the form of unemployment spells, relocation expenses and so on. Topel (1991) documents that displaced workers suffer significant wage losses. As argued in Merz (1996), to support a first best allocation in an environment characterized by costly mobility, labour contracts need to be made contingent on all possible idiosyncratic state transitions. Efficient contracts should specify not only state-contingent wages, but also mobility-contingent payments[15]. Bertola (1996a) analyses the welfare implications of severance payments in a model of costly mobility featuring both

[14]One additional reason for which workers' and firms' can have different rates of discount is if firms are exempted from paying severance payments in case of bankruptcy or takeover.

[15]Symmetrically, in our model legislated severance payments result in negative mobility costs. The unenforceability of separation-contingent payments from workers to firms implies that redundancy payments are welfare reducing.

incomplete insurance markets and capital market imperfections. He shows that, if workers cannot insure against labour income uncertainty and borrowing constrains imply that risk-averse workers can only imperfectly self-insure through accumulation and decumulation of assets, then the decentralized frictionless equilibrium features too little rather than too much mobility, as workers have to finance part of mobility costs out of reduced consumption. In such a set up severance payments increase workers' welfare by providing a substitute for incomplete contingency markets. If firms are less risk-averse than workers, this form of risk sharing is welfare improving and results in increased turnover and more efficient labour allocation[16].

To sum up, a pure separation-contingent transfer can have real effects only if either some rigidity prevents market prices to fully reflect its value to one of the parties involved or if the parties' intertemporal valuations of it differ. In this second instance, it is not necessarily the case that the value of the transfer is higher for the firm than for the worker and an inefficient trade is imposed on the parties. Dismissal costs can well be welfare improving if mobility is, realistically, costly for risk-averse workers with imperfect access to asset markets.

Booth (1997) makes a similar point in a model in which mobility is endogenously costly for workers, as right-to-manage wage-setting results in inefficient separation. Again, she finds that dismissal costs are welfare improving for the risk-averse workers in her model. There exists an optimal size for the dismissal

[16]With its emphasis on efficient risk-sharing Bertola's model clearly belongs to the implicit contract literature pioneered by Baily (1974) and Azariadis (1975). Its novelty lies in its fully dynamic nature.

cost which ensures the same outcome that would prevail under efficient bargaining. The assumption of risk-aversion in Booth's model actually obscures the higher generality of her conclusion. The crucial issue is that right-to-manage wage setting results in excessive firing even under workers' risk-neutrality, if the marginal product of labour is decreasing in employment[17]. As we argue in chapter 3, as long as separation is involuntary, dismissal costs reduce firing and induce firms to internalize the mobility cost that separation imposes on dismissed workers.

With the above two exceptions, the existing literature has focused its attention on the possible welfare reducing consequences of dismissal regulations. For this reason, it highlighted the other possible source of distortion in the labour allocation: the case in which the dismissal costs involve a deadweight loss in the form of a wedge ($F - Q > 0$) between the cost to the firm and the payment to the worker. After all, in the real world a significant part of dismissal costs consists of red-tape administrative costs such as consultation and notice requirements and paperwork which do not involve direct pecuniary benefit for workers.

The effect of dismissal costs which do not fully accrue to workers is evident from equations (2.19) and (2.20) and is basically equivalent to the partial equilibrium effect of a firing cost equal to $F - Q$. Dismissal costs reduce both job creation at expanding firms and job destruction by contracting ones. While the net effect on aggregate employment is ambiguous, the welfare implications are clearly negative as labour reallocation from low to high productivity firms is slowed down. Also,

[17]Under workers' risk-neutrality *and* constant returns to labour right-to-manage and efficient bargaining produce the same outcome, as both the marginal utility of income and the marginal product of labour are constant.

by increasing total labour costs redundancy payments depress profits, hence firms' entry and investment, and, possibly, the economy's rate of growth, as argued in Hopenhayn and Rogerson[18] (1993) and Bertola (1994). Furthermore the third-party payment $F - Q$ constitutes a net loss for the economy as a whole.

This result that dismissal costs are a source of labour market sclerosis corresponds to common perception and has been widely accepted by the profession. Intuitively, a compulsory third-party payment $F - Q$ increases the quasi-rent from continuation of the match and its duration.

The same result has been used to argue that wasteful dismissal costs may have positive in addition to negative welfare implications. By increasing long-term attachment they may provide an incentive to invest in human capital. In general, there is no reason to assume that investment in human capital is efficient. Even ruling out contracting problems[19], in the presence of positive turnover and search costs part of the returns to the investment accrue to third parties (future employers) not involved in the current match, as argued by Acemoglu (1996, 1997). Any reduction in turnover reduces the extent of the spillover thus increasing marginal returns to investment. By reducing separation rates, firing costs involving third-party payments would have two effects. On the one hand, they would depress vacancy posting, by reducing the share of total surplus which accrues to firms. On the other, they would increase investment in human capital by increasing *mar-*

[18]Surprisingly, Hopenhayn and Rogerson obtain their result in the case of a pure severance payment. The result stems from assuming, as in Hansen (1985), that workers choose over employment lotteries, receive the same income, whether employed or not, and do not compete for the future rent. Dismissal costs reduce labour supply through a pure income effect.

[19]MacLeod and Malcomson (1993) provide possible contractual solutions to the problem of inefficient investment under incomplete contracting.

ginal returns to investment. Their net welfare effect (gross of the deadweight loss) may then be ambiguous, as argued by Jansen (1997), though, once the deadweight loss is correctly taken into account, it is almost surely negative.

2.6 Conclusion

To sum up, the existing literature on dismissal costs argues that they can be decomposed into two components: a pure transfer (the payment which directly accrues to the worker) and a deadweight loss (the waste of resources associated with the necessary paperwork, etc.). The first one has no real effects if utility is transferable or, equivalently, if there are no wage rigidities or asset market imperfections, but it results in a net redistribution from firms to workers and has real effects otherwise. The second component drives a wedge between the marginal labour product and the wage and distorts labour allocation, reduces profits and investment and the aggregate rate of growth.

We have shown how all these results follow if firing costs are modelled as a tax on separation whoever initiates it. In fact, this is what all the literature on dismissal costs implicitly assumes.

The next chapter shows how this assumption is far from innocuous and how most of the above results stand and fall with it.

Chapter 3

When do firing costs matter * ?

3.1 Introduction

The aim of this chapter is to reassess the effect of firing costs in models of symmetric information and voluntary separation.

The existing consensus has been reviewed in chapter 2. Firing costs, it is maintained, have two effects. First, they result in higher wages for incumbent workers and reduce firms' *ex post* profits. Yet, the literature in question *assumes* that dismissal costs result in higher insider wages, but does not explicitly model the mechanism behind this result. As we have argued, firing costs are effectively treated as a tax on separation, whoever initiates it. This is clearly not the case in reality. Voluntary quits involve no statutory cost for firms. Second, it is contended that dismissal regulations result in higher relation-specific quasi-rents and an inefficiently lower separation rate, if only a part of the firing cost born

*I am grateful to Charlie Bean and Paola Manzini for helpful comments.

by the firm accrues to workers[1]. On the other hand, a pure severance payment should not affect job duration.

The intuition behind this second claim is the following. If part of the firing cost is a tax (a third party payment), it has real effects if both demand and supply for labour are elastic. In fact, even red tape firing costs are not a tax, but rather a redistribution of property rights on the job. *Ex post,* nobody has an interest that the cost is actually paid. Since all the above mentioned models assume transferable utility and complete and symmetric information, the Coase theorem would predict efficient separation independently from the distribution of the property rights. Something is clearly missing from this literature and, given the widespread acceptance of this paradigm and the range of implications which have been derived from it, it seems important to look inside the black box to understand the mechanism through which dismissal costs affect actions and payoffs.

It is easy to understand why the existing literature has treated firing costs as a tax on any separation. If separation is voluntary, it is impossible *ex post* to distinguish between dismissals and quits. Modelling bargaining explicitly, as it is done in this chapter, allows to meaningfully distinguish which party gains (*ex ante*) by the end of the match.

The framework we use encompasses all existing models of voluntary separation under risk-neutrality and symmetric information[2].

[1]These two predictions are common to all equilibrium models of voluntary separation and symmetric information. See Lazear (1990), Burda (1992), Hopenhayn and Rogerson (1993) and Millard and Mortensen (1997).

[2]The competitive outcome obtains as a limit case when workers' bargaining power converges

We model the bargaining process along the lines of the strategic bargaining literature pioneered by Rubinstein (1982) use the random proposer model of Binmore (1987) to avoid asymmetries in the order of offers.

Firing costs alter the firm's outside option, but not the worker's. Workers do not get them if they stop bargaining and quit to trade with a third party.

We show that firing is never an equilibrium strategy. So, for firing costs to have an effect on the equilibrium outcome it has to be the case that either some exogenous event may force the firm to fire workers despite it being suboptimal, or that separation allows the firm to free assets which have a positive market value. Intuitively, firing costs increase the specificity of the firm's capital. If its assets are already fully specific - i.e. if its outside option gross of the firing cost is zero - under no circumstance will the firm find it profitable to pay the firing cost in order to trade with a third party. So the latter cannot affect equilibrium payoffs unless involuntary firing can take place with strictly positive probability.

Finally, the main result of the chapter. Firing costs do not generate any additional joint quasi-rent and, hence, have no effect on the separation rate in models in which severance is voluntary in the absence of firing restrictions. Even when they do generate rents for workers, these rents and, hence, workers' mobility cost are always lower than the statutory firing cost. So, a firm will never dismiss a worker. When its surplus from the match is nonpositive, it will rather induce a quit by paying the worker a *voluntary* severance payment equal to her foregone rent. Under symmetric information legislated firing costs, whether received by workers

---

to zero.

or not, cannot affect the separation decision which is always socially efficient.

Dismissal regulations may result in labour hoarding only when separation would otherwise be inefficient, that is when some form of exogenous or endogenous real wage rigidity prevents the parties from agreeing on a wage reduction whenever mobility is costly for workers. Since utility is non-transferable, efficiency depends on the initial distribution of property rights on the job. In the absence of dismissal costs, severance takes place whenever the firm's payoff is non-positive whether or not the joint return within the match is strictly lower than the joint return from separation. Dismissal costs reduce the firm's return from separation, by setting a minimum price for the exercise of the right to fire. So, provided they are high enough to ensure that the minimum trade price of the property right equals the value of employment continuation to workers, they induce the firm to terminate the relationship efficiently, when the joint rent from the match is zero. We show that in a simplified version of the exogenous mobility cost model of Bertola and Ichino (1995b) a pure severance payment unambiguously increases aggregate welfare if exogenous wage rigidity results in involuntarily costly workers' mobility.

The existing empirical evidence provides support to the relevance of the mechanism we highlight in this chapter. Both the positive correlation between measures of job security and tenure and the impact on net employment creations of reforms increasing labour market flexibility are consistent with the involuntariness of separation and difficult to reconcile with the predictions of models of efficient separation

The chapter is structured as follows. Section 3.2 describes the economic environment. Sections 3.3.1 shows the irrelevance of dismissal costs when firms' investment is fully irreversible. Section 3.3.2 analyses the effect of dismissal costs on payoffs when either the firm's assets have a positive resale value or involuntary firing is possible. Section 3.4 shows that dismissal costs cannot affect separation rates and tenure in models of voluntary unemployment, but may induce efficient separation when real wage rigidity results in involuntary separation. Section 3.5 briefly discusses the empirical evidence on the relevance of involuntary separation. Section 3.6 concludes.

3.2 Economic environment

We consider a firm and a worker bargaining over the wage at which to exchange one unit of labour. This is sensible if the net return to labour is higher within the relationship due to specific investment or (non-statutory) turnover costs. As in most of the literature on search unemployment, bargaining is assumed to take place over a stock (i.e. the present value of expected surplus) rather than a flow[3]. The value of labour to the firm is a random variable y with nonnegative support and, without loss of generality, the utility of leisure is normalized to zero. So the net surplus from trade is y and the worker's utility, if trade takes place, coincides with the wage. Information is assumed to be symmetric. Both parties can decide to break the relationship irrevocably and trade outside. The return to doing so

[3] In reality, trade over labour services concerns a flow rather than a stock. In the absence of contracts this makes no difference, as the parties do not trade during bargaining. The effect of contracts is discussed in section 3.3.1. MacLeod and Malcomson (1995) survey the literature on bargaining over flows in the presence of contracts.

is w^o for the worker and π^o for the firm[4]. Also w^o is assumed to be a random variable with nonnegative support.

Since the scope for efficient risk-sharing is well understood[5], we assume both parties are risk-neutral.

Bargaining takes place over an infinite[6] time interval $[0, \infty)$ divided into discrete periods of length Δ. Each bargaining round is indexed by n. As in Binmore (1987) and MacLeod and Malcomson (1995), we assume that at the beginning of each bargaining round n, nature, $\mathfrak{N}$, chooses which player is entitled to make an offer, the worker being chosen with probability γ. γ is then a measure of the worker's bargaining power; the higher is γ, the higher is the cost to the firm of rejecting an offer, since there is a high probability that it will find itself in the same position in the next bargaining round.

The parties discount the future at the common instantaneous rate r. So $\delta^\Delta = e^{-r\Delta}$ is the discount factor for one bargaining round.

If the firm fires the worker and trades with a third party, it is bound to pay a firing cost F of which $Q \leq F$ accrues to the worker. The possibility that the cost to the firm exceeds the payment to the worker is meant to capture any deadweight loss, such as red-tape administrative costs, associated with firing. Contrary to all the existing literature, we realistically assume that quits involve no statutory cost to the firm and no payment to the worker.

The negotiation process has the following structure and is represented by the

[4]The outside payoffs are to be intended net of any non-statutory exogenous turnover cost born by either party.

[5]See, for example, Bertola (1996a) and the original implicit contract literature pioneered by Baily (1974) and Azariadis (1975) and surveyed in Hart (1983) and Stiglitz (1986).

[6]Most of the results carry over to the case of finite time horizon.

game tree in figure 3.1. In each bargaining round the agent entitled to make an offer proposes a wage w_n. The counterpart decides whether to accept (A), take its outside option (O) or reject (R).

1. If w_n is accepted, trade takes place and the game ends. The worker earns the wage w_n and the firm makes a profit $y - w_n$.
2. If the party faced with an offer unilaterally abandons the negotiations and enjoys its outside option by trading with a third party, the game ends. If the worker quits to take her outside option, the firm gets π^o and the worker earns the outside wage w^o. If it is the firm which fires the worker, it has to pay the firing cost F. In the latter case, the payoffs are respectively $\pi^o - F$ for the firm and $w^o + Q$ for the worker.
3. If w_n is rejected, the game moves onto a new bargaining round. The parties's flow payoffs in this circumstance are normalized to zero.

These moves are illustrated by the choices at stages n.1 and n.2 in figure 3.1.

A note is necessary regarding the effect of outside options on the bargaining outcome. The extensive form of the game makes it clear that we allow agents to take their outside option only after receiving an offer. Shaked (1987) has proved that if each party can take its outside option immediately after its offer has been rejected there are multiple equilibria in the infinite horizon game. In labour markets, though, it seems realistic to assume that an employer can make a last offer to match an outside option before a worker abandons the firm. Shaked calls such a market a *bazaar*. In such a market outside options cannot act as threat

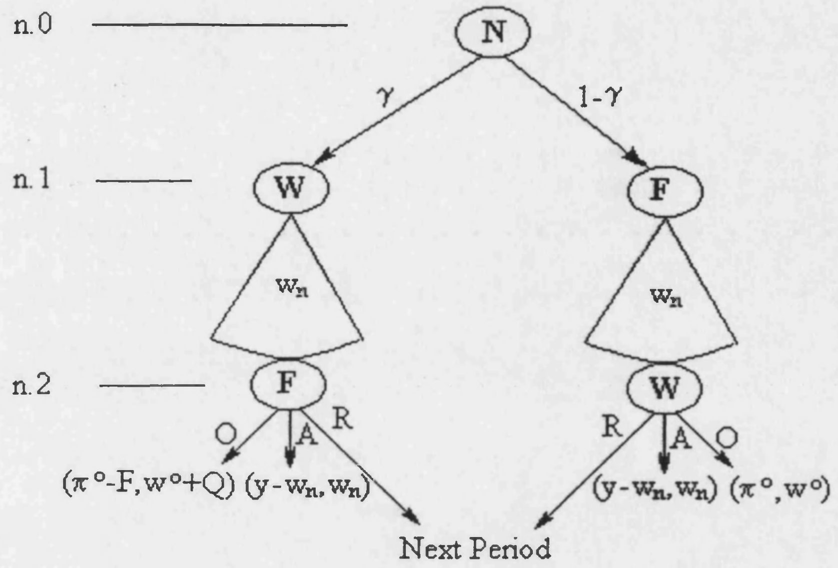

Figure 3-1: Bargaining game.

points ("take it or leave it" offers are not credible), but just provide a lower bound on the bargaining outcome.

3.3 Equilibrium payoffs and dismissal costs

3.3.1 The irrelevance of dismissal costs

We have seen that firing costs alter the firm's and worker's decision problem by reducing the firm's outside option. In order to understand how they may affect actions and payoffs it is essential to comprehend what determines the firm's outside option in situations in which wages are determined by bargaining.

Bargaining arises due to the existence of quasi-rents from the relationship. Effectively, if outside options are expressed net of (non-statutory) turnover costs, quasi-rents can only stem from past expenditures. Outside options, instead, depend on both past *and* future investment. In fact, one could break the total firm's return to the outside activity into two components: the residual value of past investment and the expected present value of future returns from investment in the outside activity net of investment costs. The second component is zero in equilibrium with free entry. It has to be non-negative since trade is voluntary and it cannot be positive as it would imply an infinite demand for the alternative activity[7]. Conversely, there is no reason, in general, for the resale or recycling value of the firm's tangible and intangible assets to be zero unless the initial cost of investment is fully sunk due to complete specificity or irreversibility.

In fact, most of the literature on dismissal costs assumes that search is the only investment activity (if any) firms are engaged in. This is equivalent to assuming full irreversibility, as search costs are a classic example of sunk cost. Allowing for firms' investment in human capital, as in Acemoglu (1997), does not make the point any less relevant. Investment in workers' human capital is also fully irreversible from the firm's point of view, as it is embodied in the worker.

This is crucial for the effect of dismissal costs on the equilibrium outcome. In all the models in which the firm's past investment is fully irreversible the firm's outside option is zero.

The effect of the availability of outside options on equilibrium payoffs in bargaining games has been explored in Stähl (1972), Binmore (1985) and Binmore,

[7]Familiriaty with markets and/or customers may imply that in equilibrium some firms, but not all, may enjoy strictly positive *ex ante* rents from a new business. Yet, these rents must stem from the previous activity and are thus included in the first component.

Rubinstein and Wolinsky (1986). In line with this literature, the equilibrium concept we use is that of subgame perfection. Strategies must be a Nash equilibrium in every subgame, thus ruling out incredible threats.

The Outside Option Principle implies that either a party's outside alternative exceeds what she could obtain in the original (without outside options) game or it does not provide her with a credible threat, hence cannot affect the equilibrium outcome. Put differently, an outside option matters only if, in equilibrium, it is binding with strictly positive probability.

If past investment is fully irreversible, firing costs cannot affect equilibrium payoffs, as π^o, the firm's return to the outside activity in the absence of firing restrictions, is zero. In this case, dismissal costs lower the firm's outside option but, since the latter is already not binding in the frictionless equilibrium, they cannot affect equilibrium payoffs. Intuitively, if there are gains from trade, the firm will never fire the worker and take its outside option. The following proposition characterizes the equilibrium.

Proposition 1 *If π^o, the firm's outside option in the absence of frictions, is zero and there are gains from trade ($y \geq w^o$), firing costs have no effect on the equilibrium payoffs. In the unique subgame perfect equilibrium trade takes place immediately and, as the time interval between offers goes to zero ($\Delta \to 0$), the equilibrium wage converges to*

$$w^* = \max \{\gamma y, w^o\}. \tag{3.1}$$

Proof. The result is just an application of the Outside Option principle to the random proposer model of Binmore (1987). Binmore proves that, in the absence of outside options, the equilibrium wage converges to γy for $\Delta \to 0$.

The outside option only acts as a constraint on the above solution[8] and, provided it can be taken only after rejecting an offer by the counterpart, the perfect equilibrium is unique. Osborne and Rubinstein p. 55-58 provide a detailed proof of this second result in the case of Rubinstein (1982) alternating offer model. ■

The first term in the curly bracket in equation (3.1), is the worker's payoff when the outside option in non-binding. The worker's share of the surplus increases with the probability that she is selected to make an offer. The second term implies that the worker's payoff cannot fall below the value of its outside option, since otherwise she could do better by quitting. In any case, if there are gains from trade the firm makes nonnegative profits and it has no incentive to fire the worker. Its (negative) outside option $-F$ is never binding.

The logic of the result is clear: a negative outside option cannot affect the equilibrium outcome whenever the continuation payoff is non-negative. Furthermore, dismissal costs cannot affect the equilibrium outcome even in case of separation. In fact, the following result can be proved.

Proposition 2 *If π^o, the firm's outside option in the absence of frictions, is zero and trade is not efficient ($y < w^o$), there exists a unique subgame perfect equilibrium in which the firm never fires the worker. As $\Delta \to 0$, the worker quits the firm to take her outside option with minimum delay.*

[8] Strictly speaking the firm would be indifferent between trading and not trading when $y = w^o$. We deal with ties by assuming throughout that whenever the parties are indifferent they do trade.

Proof. See Appendix 3.A ■

Proposition 2 has two implications. First, separation is always efficient[9], despite the existence of statutory dismissal costs. Secondly, severance of the relationship is costless to the firm, since it takes the form of a quit. The intuition is the following. When separation is efficient, it is not possible for the firm to make nonnegative profits and for the worker to obtain a wage no smaller than the alternative wage. The worker cannot impose negative profits on the firm, though, since the firm can secure a minimum payoff of zero by delaying forever. The best the worker can do is to take her outside option as soon as possible.

In brief, firing is an off-equilibrium strategy for the firm. So, when $\pi^o \leq 0$, firing costs cannot have any effect on the equilibrium outcome . They do not affect either payoffs or the separation decision in any equilibrium unemployment model in which firms' investment is completely irreversible and wages are determined by efficient bargaining[10].

The result that firing costs affect equilibrium payoffs and separation rates in models of voluntary separation with irreversible investment stems from *assuming* that they alter workers' and firms' threat points. This is equivalent to assuming that the firm can be obliged to play an off-equilibrium strategy.

[9]In general, if the firm employs more than one worker, bargaining over wages only (right to manage) results in inefficient underemployment or, equivalently, excessive separation. The outcome is efficient only if marginal returns to labour are constant. This is what most literature on search unemployment assumes. All our results would still apply to a decreasing returns to labour world under efficient bargaining. This would result if the firm bargained either with the marginal worker over wages only, or with a union with utilitarian preferences and identical members over both wages and employment. See Bertola and Caballero (1994) and McDonald and Solow (1988) for instances of each of the two cases.

[10]The conclusion would survive intact despite capital being employed in production in the rather unrealistic case in which there is a perfect market in capital and no costs of adjustment. In such a perfect world, capital does not affect the firm's opportunity cost of bargaining.

It has to be noted that in many cases the alternative to reaching an agreement which preserves employment of all the existing labour force is a reduction in the labour force. If layoffs are not associated with any scrapping of assets with positive resale value - i.e. if machines are idled rather than sold - the relevant outside option would be the non-positive return on the unused assets and propositions 1 and 2 would still apply.

The usual argument provided to support a positive effect of dismissal costs on wages even when the firm's investment is fully specific is that workers can oblige the firm to fire them by reducing effort, if the latter is not observable by third parties. Obviously, a necessary condition must be that trade concerns a *flow* of goods and services and an explicit contract is in place. This is equivalent to say that workers can renegotiate their contract wage up, as long as this is lower than their payoff in case of firing. By reducing their marginal product below the contract wage, they can increase the firm's cost of bargaining. This argument is subject to two objections. First, if the firm observes effort and lockouts are legal, the firm could respond to a reduction in effort by locking out and could not be forced by workers to fire them[11]. Namely, the workers' threat to inflict a cost onto the firm is not credible, as wages are not paid during a lockout. The bargaining process would still take the form outlined above and the outcome would still be the one described in propositions 1 and 2. A second and even more fundamental objection to the argument that workers can threat to reduce effort if an explicit

[11] The possibility that workers reduce effort and the firm's payoff during bargaining is analysed in Cramton and Tracy (1992) and Moene (1988). Both assume that workers cannot impose on the firm a negative payoff, or anyway one which is lower than the payoff during a lockout.

contract is in place is that, if this is the case, such a contract is not incentive-compatible and one is left wondering why a rational firm would agree to it in the first place. Some kind of performance-related pay would clearly be preferable from the firm's point of view and under symmetric information it would be unlikely to entail significant transaction costs[12].

It is hard to think of ways in which workers can oblige a firm to play a sub-optimal strategy when effort is perfectly observable, though only by the parties involved in the match.

3.3.2 Positive equilibrium effects of dismissal costs

Dismissal costs do not affect the set of equilibrium strategies in models of voluntary separation if the market value of past investment is not positive. So, for them to have any effect on equilibrium payoffs it has to be the case that either the firm can be forced to play a suboptimal strategy or that its outside option is strictly positive due a positive resale value of its assets.

Let us deal with the latter case first. If the firm's assets are only partly specific dismissal costs reduce their value outside the relationship if separation is initiated by the firm. Intuitively, in this case they must affect equilibrium payoffs.

Assume for simplicity that the firm's capital (including all tangible and intangible assets) does not depreciate and that the investment cost was paid upfront, so that the flow opportunity cost of capital does not affect the payoffs during

[12] MacLeod and Malcomson (1989) made it clear that if effort is non-verifiable, though perfectly observable by the parties involved, incentive compatibility requires that a positive severance payment is associated with a performance-related bonus.

bargaining. In terms of our set up, then, y would be the present discounted value of the surplus from the match gross of the opportunity cost of capital per worker installed and the outside option π^o would be the resale value of the stock of capital per head net of all shut-down expenditures (per head) excluding statutory dismissal costs.

It is clear that in this case firing costs do affect equilibrium payoffs when the firm's outside option is binding[13], since they reduce the net payoff in case the firm takes its outside option. Their effect is summarized in the following result.

Proposition 3 *Suppose π^o, the firm's outside option in the absence of frictions, is strictly positive and there are gains from trade ($y \geq \pi^o + w^o$) . In the unique subgame perfect equilibrium trade takes place immediately and, as the time interval between offers goes to zero ($\Delta \rightarrow 0$), the equilibrium wage converges to*

$$w^* = \begin{cases} \max\{\gamma y, w^o\} & \text{if } y - \max\{\gamma y, w^o\} \geq \max\{\pi^o - F, 0\} \\ y - \max\{\pi^o - F, 0\} & \text{if } (1-\gamma)y < \max\{\pi^o - F, 0\}. \end{cases} \tag{3.2}$$

Proof. This is exactly the same result as in proposition 1, but now π^o is strictly positive and provides a lower bound for the payoff to the firm in the frictionless equilibrium. Firing costs reduce the firm's outside option, but the firm can still secure a minimum payoff of zero by perpetual disagreement. ■

When the firm's assets have a strictly positive resale value, dismissal costs increase wages and reduce *ex post* profits over some states of nature through two

[13]I am grateful to Paola Manzini for pointing this out to me. Manzini and Snower (1995) analyse this case in a version of Shaked and Sutton (1984), but restrict their attention to stationary strategies.

effects. First, they reduce the firm's outside option, hence its payoff when the option is binding. Second, they reduce the probability that the firm's outside option is binding, thus obliging the firm to accept a lower payoff when the surplus from the match y is low relative to its frictionless outside option. Yet, there is a limit to the distributional effect of firing costs. They cannot reduce the firm's payoff below the disagreement payoff. Put differently, firing costs in excess of the resale value of capital per head have no marginal effect on the equilibrium shares of the surplus, as the firm can always secure a zero payoff through perpetual disagreement.

Note that as the worker's bargaining power - her probability of making an offer γ - converges to zero, we obtain the Walrasian solution: the wage in equation (3.2) converges to the reservation wage w^o. If workers are wage-takers, dismissal costs do not have any effect on the equilibrium outcome, despite reducing the firm's outside option. So, unless minimum wage constraints result in involuntary unemployment, dismissal costs cannot alter the equilibrium in the secondary labour market, at least in those sectors of it in which workers have no market power such as burger-flipping jobs. The claim that dismissal regulations may reduce job creation in low-paid, service jobs is inconsistent with workers' wage-taking behaviour.

The other instance in which firing costs can affect the equilibrium outcome is when the firm can be forced to play a suboptimal strategy. What section 3.3.1 has made clear, though, is that it cannot be the worker which forces the firm to act suboptimally under symmetric (even if possibly private) information about workers' effort. This difference between the end of negotiations because one party voluntarily takes its outside alternative and because some exogenous event may end the bargaining process at some point has been emphasized in Binmore, Rubinstein and Wolinsky (1986). It is the second possibility which may enlarge the set of credible threats with respect to the case in which separation can only be due to one party's decision to trade outside.

It is evident that a necessary and sufficient condition for firing costs to affect equilibrium payoffs when the firm's past investment is fully irreversible is that the firm fires the worker with a strictly positive probability. This can be easily seen by assuming that, if the firm and the worker do not reach an agreement within the first bargaining round, negotiations end with probability one and the outcome is deemed a firing. In the unique subgame perfect equilibrium agreement is immediate at wage

$$w = \gamma (y - \pi^o + F) + (1 - \gamma) (w^o + Q) . \tag{3.3}$$

If the parties do not agree *ex ante,* each can effectively make a take-it-or-leave-it offer if selected to propose in the unique bargaining round.

The wage in equation (3.3) coincides with the surplus sharing solution of Burda (1992) and Millard and Mortensen (1997). The result of insider-outsider models à la Lindbeck and Snower (1988) that workers are able to push firms to the firing barrier and $w = y + F$ is an even more special case when $\gamma = 1$. Lazear's (1990) result, that in a competitive labour market $w = w^o + Q$, can be obtained if $\gamma = 0$. In general, the cost F paid by the firm affects the wage only proportionally to the probability that workers have a chance to make a take-it-or-leave-it offer. If $\gamma = 0$, the wage is not affected by costs to the firm which do not accrue to workers.

This clearly exposes the implicit assumptions underlying all models that predict a positive effect of firing costs on equilibrium wages under symmetric information and irreversible firm investment. By assuming that dismissal costs alter threat points on a one-to-one basis, they effectively assume that the bargaining process is a one-shot game *and* that the disagreement outcome is a firing. The effect of dismissal costs on equilibrium payoffs would clearly be lower if the probability that negotiations end in each bargaining round were less than one. In fact, if the latter is the case, the impact of dismissal costs on wages diminishes the more impatient agents are and the lower is the probability that breakdown takes place soon, as it is shown in Appendix 3.B.

It is important to note that the model with exogenous breakdown is the only one which has the very strong implication that the impact of dismissal costs on wages is unboundedly increasing in their size. So, it cannot be simply treated as a convenient reduced form. Whatever stance one may have on strategic bargaining[14], the above result points out the importance, for a meaningful theory of firing costs, of explicitly analysing the mechanism which may force firms to fire workers despite it being suboptimal. For example, in the event that negotiations may be exogenously brought to an end by impatient creditors, the probability of break up may be initially negligible and payments in case of firing more heavily

[14]Kreps (1990) rightly points out that the non-robustness of the results of strategic bargaining models to small changes in the bargaining protocol should prevent from attaching too much weight to their conclusions.

discounted. Also, it is unclear whether in this case the firm would pay dismissal costs. If the firm had enough liquidity, it is difficult to understand how creditors could force it into liquidation. Conversely, if creditors could start a bankruptcy procedure, it is doubtful that the firm would pay firing costs.

All the above results on the effects of dismissal costs on equilibrium payoffs stem from the Outside Option Principle[15] rather than on a specific structure for bargaining costs. Also, allowing for agents to take their outside options at any time would just induce multiple equilibria, but would not change the insight that the only effect of firing costs is to increase the specificity of the firm's assets. If these assets are already fully specific, dismissal costs have no effect unless involuntary (from the firm's point of view) firing is possible.

3.4 Firing costs and separation

As we have argued above, if dismissal costs do not affect equilibrium payoffs they have no real effect. Workers would just quit whenever separation is efficient. Even when firing costs do have distributional consequences and reduce *ex post* profits, they do not have any real effect if they are a pure transfer and markets are perfect. As Lazear (1990) has noted, competition among workers would bid entrance wages down and leave *ex ante* profits unchanged. On the other hand, it is argued that firing costs should result in lower turnover if the cost born by the firm exceeds the payment to the worker. This prediction bears a clear resemblance to the theory

[15] Binmore, Shaked and Sutton (1989) provide some empirical evidence supporting the relevance of the Outside Option Principle.

of tax incidence. Suppose firing costs were a tax - a third party payment - on any separation, whoever initiated it. They could be fully translated onto workers only if the wage setting curve were perfectly inelastic. If the wage setting locus displayed a positive elasticity, a tax on firing would affect firm-level employment.

Dismissal costs differ from a tax on separation, though. Firms do not have to behave passively in the face of negative shocks. They can choose the least costly option between firing a worker and inducing a quit through a payment which compensates the worker for any mobility cost she has to bear.

Mobility costs can be exogenous relocation and search costs or may arise endogenously if workers had enjoyed rents in case of continuation. We concentrate on the latter kind of cost as the former is already included in w^o. If mobility involves no cost, the worker would just voluntarily quit whenever separation is efficient as we have shown in section 3.3.1. Workers can get no more than their outside option if they do not leave the firm.

Things are different if mobility is (endogenously) costly for workers. Mobility is costly, whenever the continuation wage w' lies strictly above w^o. In this case, the worker would never quit without a compensation payment. Yet, whenever separation is efficient (i.e. $y < w^o + \pi^o$), the firm would always prefer to compensate the worker and terminate the relationship. When the joint return within the match is lower than the joint return outside, the parties can always do better by separating. This can be easily seen by noticing that, when $y < w^o + \pi^o$, we have that

$$y - w' < w^o + \pi^o - w'. \tag{3.4}$$

If $w' > w^o$, workers would enjoy a rent $R = w' - w^o$ in case of continuation, and the firm has to pay R to induce a quit. Yet, even allowing for the voluntary severance payment, the firm's return in case of separation - the right hand side of equation (3.4) - is always strictly higher than the payoff from continuation $y - w'$.

It is useful to reflect on the mechanism through which workers' rents may not disappear when the firm would prefer to fire them. One would expect a worker, faced with a *credible* threat of redundancy, to accept a wage cut as long as her rent from continuation is positive. This should drive the rent R to zero, when the total surplus from the match is nonpositive, and mobility should be costless. For this to be the case, though, both the firm's firing threat and the worker's offer to accept a wage cut have to be credible.

The worker's offer may not be credible, if real wages are downward rigid. Real wage rigidity may be due to exogenous institutional constraints, such as minimum wage legislation, but it may also arise endogenously. This is the case if wages are set inefficiently by (right-to-manage) collective agreements or if incentive compatibility considerations require workers to enjoy state-independent rents, as in some efficiency wage models. For example, if firms pay efficiency wages in order to prevent shirking as in Shapiro and Stiglitz (1984), workers prefer a wage cut to being made redundant, but their offer is not credible, as it is not incentive compatible. Alternatively, the firm's threat to fire workers when profits are negative may not be credible, if dismissal regulations make firing costly. Either of these two elements is necessary for mobility to be costly for workers when the firm finds it optimal to terminate the employment relationship. Yet, it is only the second element - dismissal costs - which implies that costly workers' mobility may result in costly separation for firms[16]. Under employment-at-will the firm can fire the worker at zero cost whenever continuation is not profitable and does not need to induce a quit.

Dismissal regulations make separation costly for the firm if (and only if) mobility is costly for the worker. Yet, they have no marginal effect on the separation decision unless the statutory firing cost is binding, that is $R > F$ - the cost of inducing a quit (the worker's foregone rent) is higher than the statutory dismissal cost and the firm prefers to bear the latter. In fact, we can state the following result.

Proposition 4 *In the presence of dismissal regulations and independently from any assumption about market structure, the cost to the firm of terminating employment is given by* $B = \min\{R, F\}$, *where* F *is the statutory firing cost born by the firm and* $R = \max\{0, w' - w^o\}$ *is any rent that workers might enjoy in case of continuation at wage* w'. *Whenever* $B = R$, *the payment is a pure transfer, whether statutory dismissal costs involve third party payments or not, and separation takes the form of a quit.*

Proposition 4 makes it clear that firing costs in excess of R have no marginal effect on the cost of separation unless they increase R itself. So, whether firing

[16]Dismissal costs do not have to be exogenously imposed. As argued in Saint-Paul (1995), it is profitable for firms to tie their hands on firing when separation is involuntary, since workers are willing to accept lower wages in exchange for lower expected mobility costs.

costs result in lower turnover or not depends on the mechanism that generates workers' rents.

If rents are not the result of dismissal costs, but stem from endogenously determined downward rigid real wages, as it is the case in models of involuntary separation, firing costs result in lower turnover[17]. Firing costs in excess of workers' mobility costs, though, have no marginal effect on the separation rate which is the same as in the case of a pure severance payment equal to R. The intuition is the following. In models of involuntary separation utility is non-transferable and, in the absence of firing restrictions, severance takes place whenever the firm's return is negative, even if the joint return within the match is higher than the joint return from separation. Job destruction is inefficiently high, as firms do not internalize the externality, in the form of foregone rent, that severance imposes on workers. Dismissal costs, by making firing costly, induce the firm to continue the match despite a negative return, as long as the loss from continuation is smaller than the cost of dismissal. Provided they do not increase workers' continuation rents[18] by more than they decrease firms' returns, firing costs efficiently imply that separation takes place when the joint rent from the match is lower than in the frictionless equilibrium. Firing costs no smaller than workers' continuation rents ensure efficient separation and have no marginal effect on job destruction,

[17]By involuntary separation, we mean that redundant workers would strictly prefer to be retained as in the dynamic versions of Shapiro and Stiglitz's model analysed in Saint-Paul (1997) and in chapter 4 of this thesis or the union model under right-to-manage bargaining in Booth (1997).

[18]Dismissal costs may affect workers' rents in models of involuntary separation. For example, this is the case in shirking models of efficiency wage, if courts are unable to distinguish between dismissals for economic and disciplinary reasons, or in right-to-manage collective bargaining models.

as there exists a side payment that *both* the firm is willing to pay *and* workers to accept in order to separate efficiently.

Intuitively, whenever utility is non-transferable the Coase theorem does not apply and efficiency depends on the initial distribution of property rights on the job. In the absence of dismissal costs, severance takes place whenever the firm's payoff is non-positive whether or not the joint return within the match is strictly lower than the joint return from separation. Dismissal costs reduce the firm's return from separation, by setting a minimum price for the exercise of the right to fire. So, provided they are high enough to ensure that the minimum trade price of the property right equals the value of employment continuation to workers, they induce the firm to terminate the relationship efficiently, when the joint rent from the match is zero. Since private trading of the property right is both possible and desirable when the statutory trade price is set too high, the latter ceases to have any effect and efficiency results, as predicted by the Coase theorem[19].

Rents may, instead, be the result of dismissal costs. This is the case in the class of models covered by proposition 3.

In these models, though, even if firing costs are high enough to result in costly severance, they can never alter the separation decision which is always efficient. In fact, the following result can be proved.

Proposition 5 *Suppose π^o, the firm's outside option in the absence of frictions, is strictly positive and separation is efficient ($y < \pi^o + w^o$) . In the unique subgame perfect equilibrium separation takes place immediately and, as the time interval*

[19] We are obviously assuming away any eventual external effect on third parties.

between offers goes to zero ($\Delta \to 0$), the worker's equilibrium payoff converges to

$$w^* = \begin{cases} \max\{\gamma(\pi^o + w^o), w^o\} & \text{if } (\pi^o + w^o) - \max\{\gamma(\pi^o + w^o), w^o\} \geq \max\{\pi^o - F, 0\} \\ (\pi^o + w^o) - \max\{\pi^o - F, 0\} & \text{if } (1-\gamma)(\pi^o + w^o) < \max\{\pi^o - F, 0\}. \end{cases} \tag{3.5}$$

Proof. *Since separation is efficient, it is Pareto optimal to bargain over the total payoff from separation $\pi^o + w^o$ rather than the joint return from continuation y. The result then follows from proposition 3 with the only difference that the total payoff from separation $\pi^o + w^o$ replaces y in equation (3.2).* ■

The main result of this section is clear. Firing costs, whether fully received by workers or not, are at worst welfare neutral and at best welfare enhancing, as far as the separation decision is concerned. In models in which separation is efficient they do not alter the separation rate. Conversely, if separation is involuntary, dismissal costs improve welfare by reducing job destruction from an inefficiently high level.

Furthermore, by reducing separation rates, they increase firms' marginal returns from investment[20]. *Coeteris paribus,* this enhances incentives to invest in human (and possibly physical) capital and increases the rate of growth. This further boosts welfare, as investment is too low in the decentralized equilibrium due to the inefficiently high separation rate.

While the above analysis has concentrated on the effect of dismissal costs on *ex*

[20]Whether they also increase workers' *marginal* returns to human capital investment depends on the mechanism which determines wages. In the simplest efficiency wage models wages are independent not only (and obviously) from individual productivity, but also from the productivity of the average worker.

post payoffs and job destruction, they can obviously affect *ex ante* payoffs, hence job creation. As emphasized by Lazear (1990), if markets are perfect the effect of firing restrictions on expected profits is fully undone by lower entrance wages and job creation is unaffected. So the allocation of labour is unchanged in models of voluntary separation. If entrance wages cannot fall by the necessary amount, though, dismissal costs depress job creation. Aggregate employment and welfare unambiguously fall if separation is voluntary.

The aggregate employment and welfare consequences of dismissal costs in the presence of wage rigidities are less straightforward and partly model-dependent. Yet, it is possible to get some general insight with the use of a simple diagram.

Suppose, along the lines of the model introduced in chapter 2, that firms in the economy cycle between a high and a low productivity state, represented by the two, constant-slope, labour demand curves in figure 3.2. The model is a simplified version of the stochastic, competitive model à la Lucas and Prescott (1974) analysed in Bertola and Ichino (1995b). The only difference is that we assume no discounting and that transitions between states are deterministic. There are 0.5 firms in each state and in every period all of them undergo a state reversal with probability one. Mobility entails an exogenous, positive relocation or search cost κ for workers. As in Lucas and Prescott (1974), in the frictionless competitive equilibrium, high productivity firms have to pay a positive wage premium $w_g - w_b$ which exactly compensates workers for the mobility cost and mobility is costless. In our simple setup the wage differential equals the mobility cost κ and employment at firms in the good and the bad state equals l_g and l_b respectively.

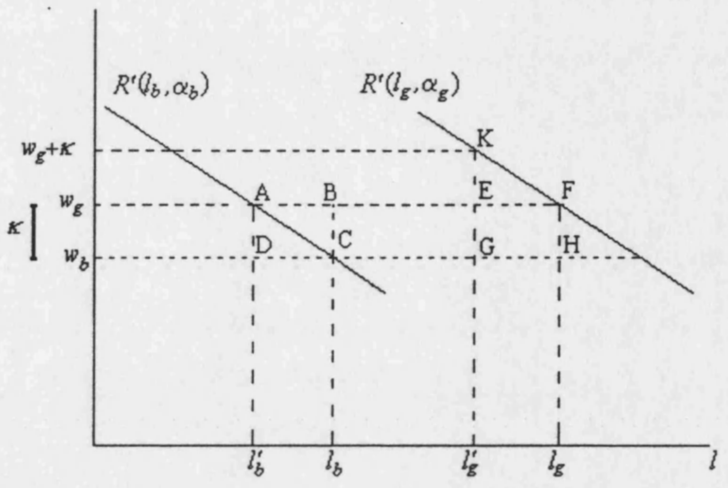

Figure 3-2: Welfare effects of severance payments in the presence of real wage rigidity

For simplicity we assume that this coincides with full employment of the labour force.

Assume now, that a minimum wage constraint or "equal-pay-for-equal-work" policy or some other form of real wage rigidity implies that the unique, economy-wide wage is w_g. Wage compression results in costly worker mobility and involuntary separation. In the absence of dismissal costs, the level of employment at contracting firms is too low at l_b' and there is too much wasteful relocation (and involuntary unemployment), as firing firms do not internalize workers' mobility costs.

The introduction of a severance payment equal to κ would be welfare improving in such a set up[21]. The payment would leave aggregate employment unchanged as the increase in employment $l_b - l_b'$ at contracting firms would be offset by the

[21] Obviously, the first best intervention would be a removal of the source of wage rigidity, but, as we show in chapter 4, the constraint on wage flexibility may well be endogenous.

fall in employment $l_g - l'_g$ at expanding ones. Yet, the fall in turnover would be welfare improving as, due to wage rigidity, the frictionless equilibrium features too much mobility. Since aggregate employment is unchanged, the change in aggregate welfare with respect to the frictionless equilibrium with wage rigidity is given by the change in total product minus the change in total relocation costs. The change in total product equals the difference between the trapezoids $KFl_gl'_g$ and $ACl_bl'_b$. Since the two labour demand curves have the same slope the net loss is represented by the area $EFHG$. Yet, total mobility costs fall by an amount equal to the sum of areas $EFHG$ and $ABCD$, as turnover fall from $l_g - l'_b$ to $l'_g - l_b$. This leaves the economy with a net gain equal to $ABCD$. Note that if wages at expanding firms were free to fall by the size of the future payment, employment at firing firms would be unaffected and aggregate welfare would be the same as in the first best equilibrium.

If the firing cost were a pure third-party payment, it would have the same effect on employment and turnover, but would be very likely to increase total mobility costs and reduce aggregate welfare. This is the case if it increases the total unit mobility cost - the sum of the firing cost and workers' mobility cost - at a higher rate than it reduces turnover. Only if the firing cost were higher than κ, would it have the same effect as the severance payment, as firms could induce workers to quit by offering them a voluntary severance payment equal to κ.

The important result is that what determines whether firing costs affect separation is not whether the cost is wasteful or not, but whether wages are rigid. Unless one is sure, that a wasteful firing cost is going to be undone by private negotiation[22], a pure severance payment is to be preferred. It involves no waste of resources and reduces the opposition of redundant workers' to redeployment as it makes mobility less costly.

Severance payments are efficient in the simple model above, yet they reduce profits. So, one would not expect firms to voluntarily introduce forms of job security. A lot of firms do offer substantial job security, though. In chapter 4 we show that severance payments may increase the value of firms if wage rigidity is endogenous rather than the result of exogenous institutional constraints.

3.5 Empirical evidence

Given the distinct implications of the two models about the desirability of dismissal costs one would like to ascertain their relative empirical relevance.

Indirect evidence giving some support to the view that inefficient separation is a relevant feature of real-world labour markets comes from studies such as Krueger and Summers (1988) which demonstrate the existence of non-competitive industry-specific rents. Case and survey studies by Raff and Summers (1987), Kaufman (1987) and Blinder and Choi (1992) find support to the causes of real wage rigidity emphasized by the efficiency wage literature.

As we have seen above, one testable prediction of models of involuntary separation is a negative correlation between the probability that the marginal job is destroyed and firing costs. A recent study by Nickell (1998) finds a strong

[22] This is the case only in the absence of significant informational asymmetries and other forms of transaction costs.

positive correlation in a cross-country regression between an index of job security and employees' tenure. One has to be careful in interpreting this result, though. First, it is based on a limited number of observations and an index which, by its own nature, has a certain degree of ad-hocness. Second, correlation does not necessarily imply causation. It could well be that countries in which workers have a lower propensity to quit are also characterized by higher dismissal costs.

A further testable implication, though, distinguishes the two classes of models. This is the response of job creation and destruction to a reduction in dismissal regulations. If separation is voluntary, a labour market reform that reduces firing costs should result in higher job destruction only if it boosts job creation. Since separation is efficient anyway, a reduction in dismissal costs can increase job destruction only if it increases hiring and workers' reservation wages. Enhanced job creation should lead (rather than lag) higher rates of job destruction. Conversely, if separation is involuntary, weakening of job security legislation should result in higher job destruction at firms for which dismissal costs are binding. Job creation would increase on impact only if market imperfections imply a negative relationship between job creation and dismissal costs *and* the policy change is credible. Again, the empirical evidence seems more supportive of this second scenario. Bertola and Ichino (1995a) provide evidence that the partial deregulation of the Swedish and Italian labour markets of the early Nineties was characterized by very high rates of job destruction by historical standards and very little job creation.

Of course, firms dislike dismissal costs, at least *ex post*. Even if they were welfare improving, firing restrictions alter the cost of separation for firms and, *ex post,* employers would always prefer not to have to pay them and would press for "increased flexibility" in downturns.

Clearly, dismissal costs are not necessarily a panacea even if separation is involuntary. First, if entry wages cannot fall by the full amount necessary to compensate for the negative impact of firing costs on firms' profitability, dismissal regulations reduce firms' value and investment. This is the case if either firing costs are binding and involve third party payments, so that workers are not willing to reduce wages by the full amount of the firm's cost[23], or if some form of market imperfection limits entrance wage flexibility. In this case, dismissal regulations may reduce job creation. Second, if firing costs increase workers' rents and wages, their impact on the separation rate is reduced. The net welfare effect would then be ambiguous, as one would expect in a second best world.

3.6 Conclusion

This chapter uses a strategic bargaining model to reassess the effect of firing costs in models of voluntary separation.

It shows that, in the presence of statutory redundancy costs, firing, as opposed to inducing workers to quit, is always an off-equilibrium strategy in this class of models. Dismissal costs can affect payoffs in only two cases: if some exogenous event may force the firm to fire the worker despite it being suboptimal, or if the firm's assets are only partly specific to the relationship. In this latter case, dis-

[23]See again chapter 4 in this thesis.

missal costs by increasing the specificity of the firm's assets, reduce the probability that its outside option is binding and *ex post* expected profits.

Most importantly, we show that, in any case, dismissal restrictions do not alter the separation rate in models in which separation is voluntary in the frictionless equilibrium, as firms always find it profitable to induce workers to quit whenever separation is efficient. Involuntary separation is an essential feature of a world in which firing costs result in a lower probability of separation. In such a world, they may be welfare improving, as the separation rate is inefficiently high in the absence of firing restrictions.

The empirical evidence on the effect of job security on tenure is broadly consistent with the involuntariness of separation.

Appendix 3.A: Proof of proposition 2

Proposition 6 *If π^o, the firm's outside option in the absence of frictions, is zero and trade is not efficient ($y < w^o$), there exists a unique subgame perfect equilibrium in which the firm never fires the worker. As $\Delta \to 0$, the worker quits the firm to take her outside option with minimum delay.*

Proof of proposition 2. From the result in proposition 1, when $y < w^o$, the firm cannot make positive profits and grant the worker a wage no smaller that her outside option. If the parties cannot agree *ex ante*, at stage n.1 in figure 3.1 the firm will propose a wage no greater than y and the worker will take her outside option. On the other hand, when responding to a proposal, the firm will reject with probability one any offer which results in negative profits, because, she can secure a payoff equal to zero by perpetual disagreement. For the same reason, the firm will never fire and pay the dismissal cost. So the supremum of the worker's *ex ante* expected payoffs at stage n.0 is

$$w = (1-\pi)w^o + \pi \max\left\{y, \delta^\Delta w\right\}. \tag{3.6}$$

The worker can obtain at most w^o if the firm is selected to make an offer in the following bargaining round and the higher between y, the highest payoff she can obtain by inducing the firm to trade (the firm making zero profit), and the highest payoff she can obtain by making an unacceptable offer in order to take her outside option if the firm is chosen to propose in the next bargaining round. In both cases, it is easy to check that immediate agreement is not possible as the firm's payoff

$y - w$ is strictly negative when $y < w^o$ and the firm would not accept. So, the parties will start bargaining and we know their optimal behaviour along the right part of the game tree in figure 3.1; viz. when the firm is selected to propose.

Let us now consider what happens when the worker is selected to make an offer. The supremum of the worker's *ex ante* expected payoffs takes two values. It is

$$w = (1 - \pi)w^o + \pi y, \tag{3.7}$$

if the surplus from trade y is not to low, that is if $y \geq \delta^\Delta w$. In this case, at stage n.1 the worker is better off making an offer that the firm will accept and obtaining y rather than waiting to take her outside option in a future round. If, otherwise, $y < \delta^\Delta w$ the supremum of the worker's payoffs is

$$w = \frac{(1 - \pi)w^o}{1 - \pi\delta^\Delta}, \tag{3.8}$$

and the worker prefers delaying and waiting to take her outside option. The value of y which partitions the set of optimal actions is $y^* = \delta^\Delta w$ where w is given by equation (3.8).

If $y < y^*$ it is optimal for the worker never to trade and wait for a chance to take her outside option. For $\Delta \rightarrow 0$, y^* converges to w^o. As the time interval between bargaining rounds tends to zero, the worker never finds it optimal to trade and takes its outside option with negligible delay .

The equilibrium can be supported by the following strategy:

1. At stage n.1 the firm offers a wage $w = y$ whenever chosen to make an offer

and the worker offers a wage $w = w^o$ when chosen.

2. At stage n.2 the firm rejects any offer of $w > y$ while the worker takes her outside option when offered anything less than w^o.

Given 1, 2 is always a best response if $y < w^o$. ■

Appendix 3.B: Equilibrium payoffs in the exogenous breakdown case

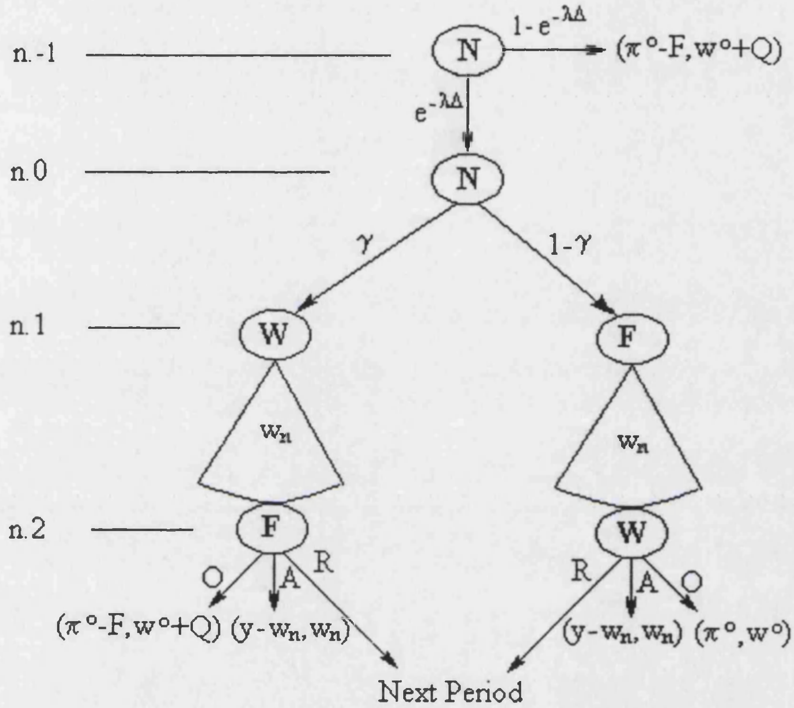

Figure 3-3: Bargaining game with exogenous breakdown.

To derive the effect of dismissal costs on equilibrium payoffs in the case in which the firm can be forced by some exogenous event to (suboptimally) fire the worker we adapt the exogenous breakdown model outlined in Sutton (1986). The structure of the bargaining game is represented in figure 3.3. In each bargaining round of length Δ, there is a probability $1 - e^{-\lambda\Delta}$ that the game is exogenously ended and the outcome is deemed a firing. This is represented by Nature's choice at node n.-1. The payoffs are $w^o + Q$ for the worker and $\pi^o - F$ for the firm. In what follows we assume the firm's gross return to the outside activity $\pi^o = 0$, but the assumption can be easily relaxed. With the complementary probability $e^{-\lambda\Delta}$ the game continues with an offer by one of the parties to the other as in the main text. If the parties fail to reach an agreement in the current round of negotiations, the game moves to a new bargaining round identical to the previous one.

Let w_λ be the infimum (supremum) of *ex ante* subgame perfect equilibrium payoffs to the worker in the modified game. It is

$$w_\lambda = (1-e^{-\lambda\Delta})(w^o+Q)+e^{-\lambda\Delta}\left[(1-\gamma)\max\left\{w^o,\delta^\Delta w_\lambda\right\}+\gamma\left(y-\delta^\Delta\max\left\{\Pi_\lambda,\Pi^d\right\}\right)\right], \tag{3.9}$$

were Π_λ is the highest (lowest) *ex ante* payoff to the firm defined by

$$\Pi_\lambda = -(1-e^{-\lambda\Delta})F+e^{-\lambda\Delta}\left[(1-\gamma)\left(y-\max\left\{w^o,\delta^\Delta w_\lambda\right\}\right)+\gamma\delta^\Delta\max\left\{\Pi_\lambda,\Pi^d\right\}\right]. \tag{3.10}$$

$$\Pi^d = -\frac{1-e^{-\lambda\Delta}}{1-e^{-(\lambda+r)\Delta}}F \tag{3.11}$$

is the firm's expected payoff in case of perpetual disagreement: with probability $1-e^{-\lambda\Delta}$, the firm faces an expected cost equal to the expected present value of the firing cost F. Symmetrically, the worker's expected payoff in case of perpetual disagreement is

$$W^d = \frac{1-e^{-\lambda\Delta}}{1-e^{-(\lambda+r)\Delta}}(w^o+Q). \tag{3.12}$$

It is easy to check that if the worker's outside option w^o and the firm's disagreement payoff are non-binding, the infimum and supremum of the worker's payoffs coincide and take the common value

$$w_\lambda = e^{-\lambda\Delta}\gamma y + \frac{1-e^{-\lambda\Delta}}{1-e^{-(\lambda+r)\Delta}}\left[(1-\gamma)(w^o+Q)+\gamma F\right] \qquad (3.13)$$

which for $\Delta \to 0$ converges to

$$w_\lambda = \gamma y + \frac{\lambda}{\lambda+r}\left[(1-\gamma)(w^o+Q)+\gamma F\right]. \qquad (3.14)$$

w_λ can also be recovered as the Nash bargaining solution to

$$w\max\left(w-w^d\right)^\gamma\left(y-\pi^d\right)^{1-\gamma},$$

where $\pi^d = -\lambda F/(\lambda+r)$ and $w^d = \lambda(w^o+Q)/(\lambda+r)$ are the limits for $\Delta \to 0$ of the disagreement payoffs Π^d and W^d in equations (3.11) and (3.12).The firm's outside option $-F$ is never binding if continuation is efficient, since it is strictly smaller than the disagreement payoff π^d if $y > w^o$. The worker's outside option instead still acts as a constraint on the worker's payoff.

As argued in the main text, the worker's payoff w_λ is decreasing in the discount rate and increasing in the instantaneous probability of breakdown λ. The equilibrium payoff converges from below to the one in the single round bargaining game with exogenous breakdown analysed in the main text for λ diverging to infinity or r converging to zero. In the first case, the probability of immediate breakdown converges to one. In the second one, the timing of the breakdown is irrelevant as the parties are infinitely patient.

The equilibrium can be supported by the following strategy:

1. At stage n.1 both the worker and the firm offer $w = w_\lambda$ if selected to make an offer.

2. At stage n.2 the worker (the firm) accepts any offer of $w \geq w_\lambda$ ($w \leq w_\lambda$) and rejects any offer strictly smaller (larger) that w_λ.

Given 1, 2 is a best response and viceversa and immediate agreement is optimal.

Chapter 4

Investment in general training with consensual layoffs

4.1 Introduction

The traditional theory of human capital as pioneered by Becker (1964) predicts that in a competitive labour market workers should bear the full cost of and capture the entire return to general training. In such an environment investment in general training is fully efficient, barring borrowing constraints. Yet, there is substantial evidence that firms share both costs and proceeds of general training. For instance, Harhoff and Kane (1994) document how German firms bear a substantial part of the cost of apprenticeship training despite that apprenticeship programmes are highly standardized and provide mostly general skills. For the US Barron, Berger and Black (1997) find that productivity growth associated with training exceeds ten times wage growth, even though most of this training is deemed general by the firms providing it[1]. Furthermore, there is a widespread consensus epitomized by Lucas (1987, p.53) that the Walrasian framework cannot capture crucial aspects of labour markets and that search frictions are crucial to explain unemployment.

Investment in general training is lower than socially optimal when costly search implies deviations from the benchmark competitive paradigm. Search costs drive a wedge between the return to a (profitable) match and the return to seeking another partner. They thus generate a quasi-rent to continuing employment. In the absence of contracts, then, bilateral bargaining determines the division of the joint surplus. This gives firms an incentive to invest in general training as long as they capture a positive fraction of the total surplus. On the other hand, the level of investment is inefficiently low as both firms and workers capture only a fraction of the return. This is the standard hold up problem of Williamson (1985).

There are two facets to hold up. First, even if complete contracting between the current employer and the worker is possible at the time of investment, part of the return to general training will be held up by *future* employers if there is a positive probability of separation. Since the future employer is unknown at the time of investment the first best can never be achieved, as argued in Acemoglu (1997). Second, in the absence of contracts, investment is held up also by the current partner further depressing incentives, as shown in Grout (1984).

Various simple and less simple contractual solutions to this second kind of spillover have been suggested. The existing literature, though, has concentrated

[1] Bishop (1996) provides extensive references to the empirical evidence on the issue.

on investment in assets that are either specific to the relationship or general, but of the "selfish investment" type. A general selfish investment is one that increases the investing party's benefit from trade both inside and outside the relationship (e.g. physical capital). General training does not fall in either of the above categories. It increases firm's revenues, but it is vested in the worker in case of separation.

This paper analyses non-contractible investment in general training in an equilibrium search model. It takes as a stylized fact firms' investment in general training and assumes that bargaining takes place according to a variant of Rubinstein's (1982) strategic bargaining model. Returns are determined by relative bargaining power if they exceed outside market opportunities, but are constrained by the binding outside return otherwise. We show that institutions that allow firms to terminate the employment relationship only with workers' consent, or that, in general, limit employers' ability to lay workers off, improve firms' incentives to invest in general training. The intuition is the following. Since human capital is vested in the worker, a firm's return in case of separation is independent from its investment in the current worker. So, its marginal return to training is zero in those states of nature in which its outside market opportunity is binding whether the match is severed or not. Nonetheless, as general training increases a worker's productivity also with other employers, the worker does capture part of the return in case of separation[2]. Consensual layoff arrangements prevent a firm from unilaterally terminating the employment relationship. Thus they oblige employers to bargain over the size of the payment - equivalently the share of the total payoff

[2]In case of separation, the remaining part of the return is reaped by the future employer as argued by Acemoglu (1997).

from separation - that induces workers to accept severance. Consensual layoffs, then, imply that firms and workers share the return to training in *all* states of nature in which workers do not quit voluntarily, thus improving employers' incentives to train. For the same reason, these arrangements also boost workers' incentives to carry out costly general investment which is vested in the firm. Examples of these investments are workers' effort to ensure product quality and the development of products that remain the intellectual property of the firm. On the other hand, with transferable utility, job security provisions can never result in inefficient separation.

Interestingly, there exist real world institutions that resemble the kind of optimal arrangements highlighted in this paper. In Germany, firms cannot legally carry out mass redundancies unless they have agreed with workers' representatives on a social plan covering procedures and compensation packages. Some US firms such as DEC, IBM, Eli Lilly contractually commit to a zero-firing policy that effectively prevents them from laying off workers unless by mutual consent. The institution of lifetime employment in Japan has the same effects. Legislated severance payments and other job security measures may achieve some or all of the efficiency gains associated with consensual layoffs depending on their size. Large enough statutory dismissal costs effectively prevent firms from unilaterally terminating the employment relationship. Yet, whenever separation is efficient the parties will bargain efficiently on a lower *voluntary* severance payment which induces the worker to agree on termination. Though, job security is often blamed for distorting the allocation of workers across firms, this paper shows that not only this is unlikely if wages are flexible, but that dismissal restrictions may actually induce both firms and workers to invest more in activities that benefit each other.

We also discuss the efficiency properties of the decentralized equilibrium we characterize. Independently, from underinvestment in training, the *laissez-faire* equilibrium is always inefficient for any given level of investment. Hosios (1990) has shown that the right value of the Nash bargaining parameter can decentralize the social optimum in search models with homogeneous agents. In our environment, the coexistence of skilled and unskilled workers implies that in addition to the sharing parameter ensuring equality between the social and private value of skilled workers, efficiency requires unskilled workers to bear the full cost of training. Since workers are heterogeneous along the job creation and the job destruction margins the sharing parameter alone cannot ensure efficiency.

This paper is related to a number of contributions in the literature. As in the literature surveyed in Acemoglu and Pischke (1999) it takes market imperfections as the reason why firms invest in general training. As in the incomplete contract literature it emphasizes contractual incompleteness *within* the current match as a source of underinvestment. Our result exploits the insight of Hart and Moore (1988) and further explored by MacLeod and Malcomson (1993), Che and Chung (1996) and Che and Hausch (1999). In all these papers breach remedies can restore efficiency under certain conditions. As noted above, though, they all restrict attention to investment which is either specific, or general but vested in the investor. On the other hand, the kind of investment we consider is general but vested in the non-investing party. The type of breach remedy proposed in the above articles is an unconditional tax on separation. Unlike the consensual layoffs arrangements discussed here, when investment is general and vested in the non-investing party such a tax would never allow the investor to capture a share of the return in case of separation.

The chapter is organized as follows. Section 4.2 introduces the model. Section 4.3 analyses the equilibrium and discusses the empirical predictions. Section 4,4 derives conditions for steady-state efficiency and discusses the sources of inefficiency in the *laissez-faire* equilibrium. Section 4.5 concludes.

4.2 The model

4.2.1 Economic environment

Time is discrete. We adopt the notational convention $x = x(t)$ and $x' = x(t+1)$ to denote the value of a variable x at the beginning of period t and $t+1$ respectively. Agents are risk-neutral and discount the future at the constant rate r. The total labour force is constant and there is a potentially unlimited supply of productive units. At the beginning of each period there are u searching unemployed workers and v firms with an open vacancy.

Production requires a fixed quantity of physical capital which has to be in place before the firm starts searching for a partner. The cost of the investment is κ and can be fully recovered in case of separation. Alternatively, one could think of κ as a one-off cost to the firm of entering the labour market. As shown in chapter 3, what is crucial for the result in this paper and for any effect of firing costs in a bargaining framework is that the firm's return to firing a worker is positive in the absence of employment protection legislation.

Because of uncertainty about the location of potential partners' agents have to search for one. Finding a match takes at least one period. Search frictions are modelled according to a constant returns to scale, strictly concave, matching technology. So, matching probabilities depend only on market tightness $\theta = v/u$. $q(\theta)$ and $p(\theta) = \theta q(\theta)$ are respectively the proportion of firms and workers who find a match by the end of the period. Both are restricted to lie in the unit interval.

The timing of events for a matched pair is illustrated in figure 4.1. At the end of period t a partner has been found. Before the quality of the match is discovered - at time $t.1$ - the parties can negotiate side-payments[3]. If the worker is untrained the firm trains her at time $t.2$. Training is fully general and takes place at a constant marginal cost normalized to one. Investment is instantaneous and third parties cannot verify neither its level nor the productivity of the match. This prevents a matched pair from writing a complete enforceable contract at time $t.1$ and implies that firms underinvest in training since investment is held up.

At the beginning of $t+1$ the pair draws a match-specific random productivity shock z. Shocks are independently and identically distributed across matches with support $[0, \infty)$ and continuous cumulative density function $G(z)$.

If the shock is favourable enough the pair bargains over a wage and produces in period $t+1$ a flow of output $zf(h)$ with $f(.)$ strictly increasing, strictly concave

[3]In section 4.3.3, we discuss the consequences of relaxing this assumption.

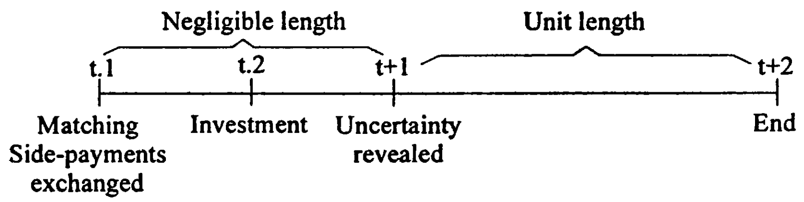

Figure 4-1: Timing of events.

and satisfying the Inada conditions. After one period of production the pair dies. To ensure stationarity of the environment it is assumed that every worker that is employed at the beginning of period $t+1$ begets a son/daughter that will enter the labour market and start searching at the beginning of the next period. Normalizing the total labour force at the beginning of each period to one implies that the flow of new entrants into the labour force at the beginning of the following period is

$$in' = 1 - u \tag{4.1}$$

If the shock is below a reservation level b the parties separate and start searching for a new partner. The firm has to pay a statutory severance payment[4] equal to F in case it fires the worker, but no payment is due if a worker quits. Clearly, our distinction is meaningful only if third parties can distinguish between quits and layoffs.

We assume that outside parties can verify: a) whether a worker shows up for work; b) if the firm allows the worker on the premises; c) any written communication between the two parties. A separation is deemed a dismissal if the firm gives

[4] As we have shown in chapter 3, allowing for part of the cost born by the firm to be wasted would not affect the result.

the worker written notice that it no longer wishes to continue the employment relationship. The end of the relationship is deemed a quit if the worker does not show up for work without providing a written justification (e.g. a medical certificate) or if the worker gives written notice that she no longer intends to continue in employment. Until one of these actions is taken the employment relationship is considered in existence. This seems broadly consistent with existing practices in most developed countries.

Carmichael (1983) has argued that severance payments cannot be conditioned on the identity of the party initiating separation: a firm that wanted to dismiss a worker could always induce her to quit by making her life difficult and viceversa. In practice, legislation often prescribes payments to employees in case of layoff, but workers are not entitled to (and in general do not receive) any payment if they quit. So it has to be the case that conditional severance payments are, if only imperfectly, enforceable. As noted in chapter 3 (pp. 62-63), MacLeod and Malcomson (1989) have shown that, if firms can observe effort, workers' moral hazard problem can be solved by a wage contract with a performance-related component. On the other hand, one would expect that, at least in the case of collective workforce reductions, it is difficult for an employer to convince a court that a claim of constructive dismissal filed by a works council or a group of workers is unfounded. Furthermore, if firms could easily disguise layoffs as quits dismissal costs, and the whole debate on their impact, would be irrelevant as firms would never pay them.

The fact that a proportion of trained matched workers becomes unemployed implies that the unemployment pool contains both skilled and unskilled workers. Since training is general and search costly, it also implies the presence of positive spillovers as in Acemoglu (1997).

For simplicity, I restrict attention to symmetric, steady-state, pure-strategy equilibria. To find such an equilibrium, suppose that (given the matching and bargaining process) the level of training of the representative skilled worker equals h^*. Then derive the individually-optimal entry decision of a single unmatched firm and the investment decision h of a single firm, matched to an unskilled worker, with the total number of vacancies v, unemployment stocks u and u_s and h^* taken as given. In equilibrium $h = h^*$.

4.2.2 Flows and unemployment

The stock of unemployed workers at the beginning of each period evolves according to

$$u' = u\left[1 - p(\theta)\left(1 - G(b)\right)\right] + in'. \tag{4.2}$$

u' equals the number of searching workers who were not matched in the previous period, plus those who found employment but whose job was destroyed plus the flow in' of new entrants into the labour force. Together with (4.1), equation (4.2) implies that steady state unemployment is given by

$$u = \frac{1}{1 + p(\theta)\left[1 - G(b)\right]}. \tag{4.3}$$

Equation (4.3) is the Beveridge curve. A higher job finding rate $p(\theta)$ and a lower rate of destruction of unproductive matches $G(b)$ decrease steady state unemployment.

Since all the workers who lose their job are trained, the stock of skilled unemployed workers evolves according to

$$u_s' = u_s\left[1 - p(\theta)\right] + up(\theta)G(b). \tag{4.4}$$

The mass of skilled unemployed workers u_s' equals the number of skilled workers who did not leave unemployment in the previous period plus those workers (all trained) who were matched but lost their job in the previous period. This implies a steady state proportion of skilled workers in the unemployed pool equal to

$$\frac{u_s}{u} = G(b). \tag{4.5}$$

4.2.3 Search

For simplicity, we assume there are no unemployment benefits and the utility of leisure is zero. So $U(h)$, the asset value of an unemployed worker with general human capital h at the beginning of the period, is

$$[r + p(\theta)]U(h) = p(\theta)E_a(h), \tag{4.6}$$

where $E_a(h)$ is the value of accepting a match.

Our set up implies that all skilled workers have the same level of training. So, in the symmetric equilibrium $h = 0$ if the worker is untrained and $h = h^*$ for a trained worker, where h^* is the optimal level of training for the representative firm.

V, the value of a searching firm, depends then on the expected level and incidence of training among the unemployed population and satisfies

$$[r + q(\theta)]\, V = q(\theta) \left[(1 - G(b))\, J_a(0) + G(b) J_a(h^*)\right], \tag{4.7}$$

where $J_a(0)$ and $J_a(h^*)$ are the values of accepting a match with an unskilled and trained worker respectively. Conditional on having contacted a worker the probability that she is skilled is $u_s/u = G(b)$. In equilibrium with free-entry the value V of posting a vacancy equals κ, the investment cost.

4.2.4 Bargaining

Because of search frictions a match which is formed and/or is not destroyed yields quasi-rents. We assume that the parties will bargain over the division of these quasi-rents according to a variant of alternating offer bargaining due to Binmore (1987).

At the beginning of each bargaining round, nature selects one of the two parties to make an offer, the worker being selected with probability β. The counterpart either accepts the offer, in which case production takes place and the game ends, or she rejects the proposal and the game moves to a new round after a delay equal to Δ. When responding to an offer each party can also unilaterally and irreversibly abandon the negotiations to trade outside (take her outside option, in the bargaining terminology), ending the game. We assume the parties cannot search for another partner during bargaining[5].

The solution to the general bargaining problem is given by the following proposition.

Proposition 1 *Be S the expected value of the total surplus from reaching an agreement and E and J respectively the worker's and firm's share of this surplus. Then:*

a)

$$S = \max\{C, U + \kappa\}, \tag{4.8}$$

where C is the expected value of the total surplus from continuation of the match;

b) the unique, subgame perfect equilibrium values of E and J satisfy

$$E = \begin{cases} \beta S & \text{if } U < \beta S < S + F - \kappa \\ U & \text{if } U > \beta S \\ S + F - \kappa & \text{if } \beta S > S + F - \kappa \end{cases} \tag{4.9}$$

and

$$J = S - E. \tag{4.10}$$

[5]Relaxing this assumption would not alter the qualitative nature of our result. Masters (1998) allows for search during bargaining in a similar set up. He shows that, unless the employment relationship is mediated by an intermediary who pays the parties their marginal product, the underinvestment result goes through.

Proof. See propositions 3 and 5 in chapter 3. ■

The first part of proposition 1 implies that the parties will bargain over the higher between the joint payoff from continuation and the total return from separation. With transferable utility, the separation decision is always efficient in the sense that it maximizes the total payoff, independently from the existence of legislated dismissal costs. This is just one more instance of the Coase theorem.

Part b) states that the parties share the joint payoff according to the relative bargaining power β unless either party can do better by abandoning the match and searching for a new one. In this latter case, the binding outside option determines the shares. If $F > 0$, firing costs reduce the firm's outside option and its payoff in those states in which its market return would be binding in the frictionless equilibrium.

Firing costs drive a wedge between the return to the firm's assets outside the relationship in case the worker unilaterally abandons the match and the same return if the firm fires the worker. This wedge increases the scope for bargaining not only over the surplus from continuation, but also over the total payoff from separation. As shown in chapter 3, the firm cannot sever the relationship unless it pays the firing cost or bargains with the worker over a voluntary side-payment that induces him to quit. On the other hand, workers are free to quit at any time.

When a match is formed at time t the ex ante expected surplus to split is

$$S_a(h) = S_p^e(h^*) - (h - h^*). \tag{4.11}$$

The ex ante surplus from meeting a worker with human capital h is given by $S_p^e(h^*)$, the expected ex post surplus from being matched with a trained worker at the beginning of $t+1$, minus the cost of training the worker. The cost is obviously zero for a trained worker with initial human capital h^*.

Using (4.7) and (4.10) we can then write the free-entry condition as

$$\kappa\left(1+\frac{r}{q(\theta)}\right) = S_p^e(h^*) - G(b)E_a(h^*) - (1-G(b))\left(h+E_a(0)\right), \tag{4.12}$$

where the expectation of the ex post surplus equals

$$S_p^e(h^*) = f(h^*)\int_b^\infty zdG + G(b)\left[U(h^*)+\kappa\right]. \tag{4.13}$$

The joint surplus coincides with the revenue from production if the match-specific shock is above the reservation productivity b and the total return from separation otherwise. In case the match is severed the joint payoff is given by the value $U(h^*)$ of being a trained unemployed worker plus κ, the value of search to the firm.

Given that all firms are identical the worker's outside option cannot be binding at $t.1$, as at best she will meet an identical firm one period later.

Similarly, the firm's outside option is not binding in case it is matched with a trained worker. In the best possible case, it will meet a similar worker with a one-period delay. Proposition 1 then implies

$$E_a(h^*) = \beta S_p^e(h^*). \tag{4.14}$$

Things are different in case a firm meets an unskilled worker. If the firm turns the worker down and searches for another match, with positive probability it will find a skilled worker after one period and will not have to bear the training costs. So the firm's share of the total surplus is the higher between the return to going back to search κ and a proportion $(1-\beta)$ of the surplus. That is

$$E_a(0) = \min\left\{\beta\left[S_p^e(h^*) - h^*\right], S_p^e(h^*) - h^* - \kappa\right\}. \tag{4.15}$$

Equation (4.15) shows that, though the firm invests in training non-cooperatively, an untrained worker shares the cost of the training that it is optimal for the firm to provide ex post. It needs to be pointed out that dismissal costs do not affect the firm's outside option at time $t.1$ since they are not due if a job applicant is turned down before starting employment.

Whatever the distribution of the ex post surplus at $t+1$ side payments ensure that the ex ante distribution satisfies (4.14) and (4.15).

We can then use equations (4.6), (4.13) and (4.14) to solve for the reduced-form asset value of a trained unemployed worker

$$U(h^*) = \frac{p(\theta)\beta}{r + p(\theta)\left[1 - \beta G(b)\right]}\left[f(h^*)\int_b^\infty z dG + G(b)\kappa\right] \tag{4.16}$$

and the ex post, expected joint payoff

$$S_p^e(h^*) = \frac{r + p(\theta)}{r + p(\theta)\left[1 - \beta G(b)\right]}\left[f(h^*)\int_b^\infty z dG + G(b)\kappa\right]. \tag{4.17}$$

Equations (4.12), (4.14) and (4.15) allow to solve for the reduced-form, free-entry condition

$$\kappa\left(1+\frac{r}{q(\theta)}\right) = (1-G(b))\max\left\{(1-\beta)\left[S_p^e(h^*)-h^*\right],\kappa\right\} + G(b)(1-\beta)S_p^e(h^*). \tag{4.18}$$

4.3 Investment and equilibrium

The firm invests in training non-cooperatively after side-payments have been exchanged and before uncertainty about the quality of the match is revealed. Optimality then requires equality between the marginal investment cost and the expected marginal return to the firm, or

$$1 = \frac{\partial J_p^e(h^*)}{\partial h}, \tag{4.19}$$

where $J_p^e(h^*)$ is the expected post-investment payoff to the firm.

At time $t+1$, once the quality of the match has been realized, the surplus from reaching an agreement is

$$S_p(z,h^*) = \max\left\{zf(h^*), U(h^*)+\kappa\right\}. \tag{4.20}$$

From proposition 1 we know that the parties bargain over $zf(h^*)$ as long as continuation is efficient or $z \geq b$, where the reservation productivity b satisfies

$$bf(h^*) = U(h^*) + \kappa. \tag{4.21}$$

In general, β determines the share of revenues that each party receives when revenues are high, but either party's outside return may become binding for low values of z. The following lemma establishes the conditions under which the firm's or the worker's market alternative is binding with positive probability.

Lemma 2 *If* $(1-\beta)U(h^*) < \beta\kappa$ *in equilibrium, then for*

$$F < \kappa - (1-\beta)bf(h^*) \tag{4.22}$$

there exists $z_r \in [b, \infty)$ *satisfying*

$$F = \kappa - (1-\beta)z_r f(h^*) \tag{4.23}$$

such that $\forall z \leq z_r$, $J_p(z, h^*) = \kappa - F$.

Viceversa, if $(1-\beta)U(h^*) > \beta\kappa$*, then,* $\forall F$*, there exists* $z_r \in [b, \infty)$ *satisfying*

$$U(h^*) = \beta z_r f(h^*) \tag{4.24}$$

such that $\forall z \leq z_r$, $E_a(z, h^*) = U(h^*)$.

Proof. See appendix 4.A. ■

The condition $(1-\beta)U(h^*) < \beta\kappa$ implies that, when the match productivity is low, the firm's bilateral monopoly share of the highest between the surplus from production and that from separation falls short of the firm's payoff from firing the worker and trading outside. When the match productivity is low the firm receives its outside option since the threat to fire the worker is credible and is actually carried out when separation is efficient.

Viceversa, if the inequality is reversed, it is the worker's market return that becomes binding in bad states and independently from the size of firing costs. When separation is efficient, the worker quits the firm, since the share of the total payoff from separation he would obtain by bargaining is lower than her outside option.

In general, there is no reason to expect one condition rather than the other to prevail. In the presence of both idiosyncratic and aggregate uncertainty one would expect the first condition to prevail in recessions, when the value of being unemployed is low, and the reverse condition to prevail in booms, when market tightness and the expected surplus from a match are high.

Given that firing costs matter only in those states of nature in which the firm's outside option is binding in the *laissez faire* equilibrium, we will assume for simplicity in what follows that the first condition always holds.

4.3.1 Equilibrium with small severance payments

Proposition 1 and lemma 2 together imply that if firing costs satisfy $F < \kappa - (1 - \beta)bf(h^*)$, the expected ex post payoff to the firm will be

$$J_p^e(h^*) = (1-\beta)f(h)\int_{z_r}^{\infty} zdG + G(z_r)(\kappa - F). \qquad (4.25)$$

The firm receives a share $(1 - \beta)$ of total revenue if the match productivity is high enough and its outside option in all other states. The first-order condition for optimal investment is then

$$1 = (1-\beta)\frac{\partial f(h^{*})}{\partial h}\int_{z_r}^{\infty} z dG(z). \tag{4.26}$$

With small or no severance payments the privately optimal level of training is independent from external conditions. Since human capital is vested in the worker the firm's payoff when $z < z_r$ is independent from the level of training.

The level of investment is a decreasing function of z_r, as the higher z_r the higher the probability that the firm's outside return is binding. As equation (4.23) shows, severance payments reduce z_r. Hence they increase the range of states over which the firm shares the return from its investment and improve its incentives to train.

That breach remedies can improve the investor's incentives through the mechanism highlighted here was first suggested by Hart and Moore (1988) and further exploited in MacLeod and Malcomson (1993), Che and Chung (1996) and Che and Hausch (1999). The only difference is that while in those articles breach penalties cannot be conditioned on the identity of the party who refuses to trade, here severance payments are not due if it is the worker that quits the firm. The reason for this difference is twofold. First, this paper focuses on the employment relationship rather than general bilateral relationships. In practice workers do not receive any payment if they quit. Second, when firms invest in *general*, rather than specific, training it is not necessarily the case that imposing a lump-sum transfer on the firm if the worker quits improves the firm's incentives to train. For example, MacLeod and Malcomson (1993) show that if fixed-wage contracts can be written, firms capture the full marginal return to training in those states in which the contract is not renegotiated. Taxing firms on quits would increase the probability that workers capture part of the return and discourage investment[6].

We can now characterize the equilibrium with zero or small severance payments. Using equations (4.21) and (4.16) we can write the reduced form job destruction condition as

$$bf(h) = \kappa + \frac{p(\theta)\beta}{r + p(\theta)\left[1 - \beta G(b)\right]} \left[f(h^*) \int_b^\infty z dG + G(b)\kappa \right]. \tag{4.27}$$

Definition 3 *A stationary symmetric equilibrium with zero or small dismissal costs is a vector of allocations* $[\theta, u, u^s, h^*, b, z_r]$ *and a value function* $S_p^e(h^*)$ *such that: (i) the free entry condition (4.18) determines* θ*, (ii)* $S_p^e(z, h^*)$ *is given by equation (4.20), (iii) the two flow equilibrium equations (4.3) and (4.5) determine* u *and* u^s*, (iv)* h^* *solves the first order condition (4.26), and (v)* z_r *and* b *satisfy equations (4.23) and (4.27).*

For a given level of h, equilibrium can be represented graphically as the intersection of the job destruction (JD) and a job creation (JC) condition. Under the assumption that the firm's outside option is not binding[7] at $t.1$, one can write one version of the job creation condition by using equation (4.21) to replace $U(h^*)$ in

[6]On the other hand, in our model the worker's marginal return is lower outside than within the relationship in those states in which the the worker's frictioless outside option is binding. Taxing firms' on quits would further increase incentives to invest.

[7]The case in which the outside option is binding is qualitatively similar.

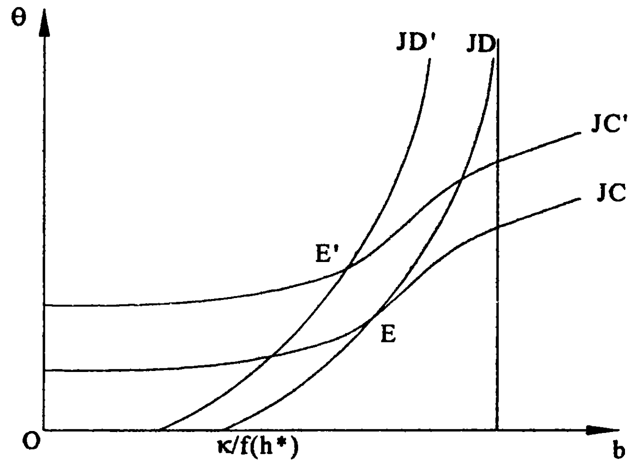

Figure 4-2: Effect of severance payments on equilibrium.

(4.18), (4.13) to obtain

$$\kappa\left(1+\frac{r}{q(\theta)}\right)=(1-\beta)\left[f(h^*)\int_b^\infty zdG+G(b)bf(h^*)-(1-G(b))\,h^*\right]. \quad (4.28)$$

Figure 4.2 plots the two curves in the (θ, b) space. The JC locus is upward sloping[8]. The JD curve - given by equation (4.27) - is upward sloping and convex, with a strictly positive horizontal intercept at $b=\kappa/f(h^*)$ and a vertical asymptote. Thus, provided JC lies above JD at $b=\kappa/f(h^*)$ - that is provided vacancy posting is positive when the value of unemployment is zero - an equilibrium exists[9].

The system is block recursive with equations (4.23) and (4.26) determining the level of training. An increase in severance payments, results in higher training. This induces firms to post more vacancies for given separation rate and to fire

[8]In Mortensen and Pissarides (1994) the JC curve is downward sloping due to the different bargaining solution adopted.

[9]It is not possible to prove that the equilibrium is unique, though numerical experimentation suggests that this is the case.

less for given market tightness. Suppose the economy is initially in equilibrium at E. An increase in severance payments then moves both the JD and the JC curves up. It can be easily shown that the horizontal shift in the JC locus always exceeds the shift in the JD curve. Assuming that the equilibrium is unique, severance payments unambiguously increase market tightness and the job finding rate $p(\theta)$, but have an ambiguous effect on the reservation productivity b and the separation rate. In case the job destruction rate increases, the net effect on equilibrium unemployment is ambiguous. Numerical simulations, though, indicate that whichever the direction of the movement in unemployment incidence, the increase in vacancy posting prevails and employment increases.

4.3.2 Equilibrium with consensual layoffs

Small severance payments reduce the probability that the firm's outside option is binding when continuation is efficient. Yet, they do not prevent the firm from firing the worker when the match is no longer viable. So, the firm does not capture any return to its investment in case of separation.

Suppose instead that a firm can sever the employment relationship only with the worker's consent. This effectively locks the firm in a bilateral monopoly situation. The firm does not only have to share the surplus from production. When separation is efficient the firm cannot unilaterally sever the relationship, though this would give it a larger share of the total separation payoff. Instead, it has to bargain over the size of the payment that induces the worker to agree on separation. The firm's ex post payoff at time $t+1$ is then

$$J_p(z,h^*) = (1-\beta)\max\{zf(h^*), U(h^*)+\kappa\}. \tag{4.29}$$

With the expected payoff at the time of investment given by

$$J_p^e(h^*) = (1-\beta)\left[f(h^*)\int_b^\infty zdG + G(b)\left(U(h^*)+\kappa\right)\right] \tag{4.30}$$

the firm would invest up to the point where

$$1 = (1-\beta)\left[f'(h^*)\int_b^\infty zdG + G(b)\frac{\partial U(h^*)}{\partial h}\right], \tag{4.31}$$

or, using equation (4.16)

$$1 = (1-\beta)\frac{r+p(\theta)}{r+p(\theta)\left[1-\beta G(b)\right]}f'(h^*)\int_b^\infty zdG. \tag{4.32}$$

Confronting equations (4.31) and (4.26) it is evident that the obligation to sever the employment relationship by mutual consent further increases investment for two reasons. First, the firm's outside option is never binding when production is efficient: z_r does not enter the investment condition any more. Second, the firm now captures a fraction $(1-\beta)$ of the marginal return to training outside the relationship. Consensual layoff arrangements reduce the firm's *total* return from separation, but by forcing employers to share the total outside payoff they increase their *marginal* return to training. This second effect is the new insight of this chapter. In so far as investment is general and vested in the non-investing party on separation, institutions or contractual arrangements that result in sharing of the total separation payoff improve incentives to invest.

Interestingly, institutions of the kind envisaged here do exist in practice. In Germany firms cannot carry out collective redundancies unless they have secured the works council's approval of a social plan detailing the conditions and terms of layoffs, including the size of severance payments. The institution of lifetime employment in Japan and the voluntary commitment to a zero-firing policy in certain firms such as DEC, IBM, Eli Lilly and others achieve the same result. Dismissals are still carried out but only on terms which meet the workers' consent. Note that, provided ex ante side payments, are unconstrained our model predicts that it is rational for firms to adopt such policies.

In other countries such as Spain and Italy, high explicit or implicit firing costs can achieve the same result. In fact, it can be shown that

Corollary 4 *If* $(1-\beta)U(h^*) < \beta\kappa$ *and* $F > \kappa - (1-\beta)bf(h^*)$ *then the firm's ex post payoff is given by*

$$J_p(z, h^*) = (1-\beta)\max\{zf(h^*), U(h^*) + \kappa\}. \tag{4.33}$$

Proof. See appendix 4.A ■

Large enough severance payments achieve the same effect as a consensual layoff clause by reducing the firm payoff from firing below the bilateral monopoly outcome. The firm is then better off paying the worker a share of the total separation payoff to induce her to quit rather than unilaterally severing the relationship.

We can now characterize the equilibrium with either large severance payments or consensual layoff provisions.

Definition 5 *A stationary symmetric equilibrium with consensual layoff is a vector of allocations* $[\theta, u, u^s, h^*, b]$ *and value function* $S_p^e(h^*)$ *such that: (i) the free entry condition (4.18) determines* θ*, (ii)* $S_p^e(z, h^*)$ *is given by equation (4.20), (iii) the two flow equilibrium equations (4.3) and (4.5) determine* u *and* u^s*, (iv)* h^* *solves the first order condition (4.32), and (v)* b *satisfies equation (4.27).*

The equilibrium can still be represented by the job destruction and job creation loci in figure 4.2, but the system is no longer recursive. The optimal level of investment in equation (4.32) now depends on aggregate variables. Yet, one can prove that the equilibrium with consensual layoff provisions features a higher training level than the one with small or no severance payments.

We can use a continuity argument exploiting the equivalence between consensual layoffs and large enough severance payments established in corollary 4. We know from the previous subsection that in the equilibrium with small severance payments the optimal level of training in equation (4.26) is independent from external conditions and increasing in the size of dismissal costs. As the severance payment F increases, both z_r and b change, but their distance decreases. For F converging to its critical value $\kappa - (1-\beta)bf(h)$ from the left, z_r converges to b. So, the integral on the right hand side of equation (4.26) is infinitesimally close to the first addendum in the bracket on the right hand side of (4.32). For F equal or larger than its critical value the right hand side of (4.32) equals the right hand side of (4.26) plus a strictly positive term in $U'(h^*)$.

So, an equilibrium with consensual layoffs features, coeteris paribus, a higher training level and job finding rate while, as in the previous section, no unambiguous analytical predictions can be made on the direction of the change in the separation and the unemployment rates.

4.3.3 Empirical predictions and discussion

The model is too crude to allow for convincing calibration. Yet, its main insight revolves around the internalization of the externality associated with human capital being vested in the worker on separation. This aspect would survive in largely unchanged form in a more realistic model .

It is then possible to work out the percentage change in the level of training stemming from the introduction of consensual layoffs in an economy in which severance payments are large enough to ensure that firms' outside return is never binding when continuation is efficient, but not so large as to induce consensual layoffs. When F is just below the level in (4.22) that results in consensual layoffs, z_r in equation (4.26) is infinitesimally close to b in equation (4.32). If we call h_L^* the optimal level of investment in the first case and h_H^* the investment level with consensual layoffs, we can take the ratio of (4.32) and (4.26) to obtain

$$1 = \frac{f'(h_H^*)}{f'(h_L^*)}\left(1 + \frac{p\beta G(b)}{r + p(1 - \beta G(b))}\right). \tag{4.34}$$

Assuming a Cobb-Douglas production function $f(h) = h^\delta$, the percentage change in investment associated with the introduction of consensual layoffs is then implicitly given by

$$\frac{h_H^*}{h_L^*} = \left(1 + \frac{p\beta G(b)}{r + p\,(1 - \beta G(b))}\right)^{\frac{1}{1-\delta}}. \tag{4.35}$$

The output elasticity δ can be recovered from empirical studies of the impact of training on wages. Under the assumption of rent sharing, the wage and revenues elasticity with respect to training coincide. Parent (1999), using the US National Longitudinal Survey of Youth estimates a wage semielasticity with respect to training equal to 0.12 which given a mean level of training equal to one quarter gives an elasticity equal to 0.03. An elasticity of 0.02 can be obtained based on a similar study by Loewenstein and Spletzer (1999). We chose an intermediate value of $\delta = 0.25$ and set the real interest rate r to 0.04. The value of the job finding rate p is not particularly crucial. It is clear from the above equation that, as long as p is relatively large with respect to r, it has little effect on the results. We set $p = 1$ which is consistent with an average unemployment duration not exceeding one year.

Table 4.1 presents the percentage change in the level of training associated with consensual layoffs for different values of the sharing parameter β and the layoff rate $G(b)$. The range for β reflects the empirically observed values for the share of labour income in total product. Since there is no sharing in the case of quits both in reality and in our model, the relevant rate to look for is the layoff probability rather than the total separation rate. Blanchard and Portugal's (2000) comparative study of job an worker flows in Portugal and the US identifies the

		Layoff rate		
		0.1	**0.2**	**0.3**
	0.5	0.05	0.11	0.17
β	**0.6**	0.06	0.13	0.22
	0.7	0.07	0.16	0.26

Table 4.1: Percentage increase in training level associated with consensual laoffs.

layoff rate with the rate of job destruction. They estimate the annualized (quarterly) layoff rate for Portugal to 16% and the same rate for the US to respectively 22% and 29% for the manufacturing sector and all sectors respectively.

As can be seen, the gains are small in countries with low layoff rates, but can be quite sizeable in countries in which firm-initiated turnover is higher. This is no surprise, as the extent of the externality is increasing in the rate of turnover. Also, the higher is β the higher is the fraction of the spillover accruing to the worker on separation and the larger the incentive that consensual layoffs provide.

The size of these effects suggests that the mechanism provided cannot be the main explanation for cross-country and cross-culture variation in training levels[10]. Yet, it is by no means negligible, at least in countries with higher layoff rates.

Some empirical support for this mechanism is provided by Bishop (1991) who finds that the likelihood and amount of formal training are higher at firms where firing a worker is more difficult.

[10]For example, the studies surveyed in Bishop (1996) document large differences in the incidence and duration of training between the US on the one hand and Germany and Japan on the other. Krafcik (1990) finds that newly hired assembly workers in the US receive an average of 48 hours of training in US-owned plants and 280 in Japanese-managed ones.

The insight of this chapter is not restricted to firm-provided training. A number of authors[11] have conjectured that job security measures may increase workers' contribution to firms' value. The mechanism studied here applies equally to investment carried out by employees which is vested in the firm and general in nature. For example, the reputation for high quality and reliability of German and Japanese cars is vested in the manufacturing companies, but is largely dependent on their labour force's effort. A programme developed for a software engineer employed by a firm is intellectual property of the employer. In all these cases, consensual layoff arrangements allow workers to capture part of the return to their investment on separation.

Of course, a measure which redistributes ex post payoffs from firms to workers must reduce firms' incentives along some other line. Provided side payments from workers to firms are not required or constrained, consensual layoffs arrangement do not alter ex ante bargaining power. So they have no direct effect on any investment carried out before a match is formed[12]. On the other hand, consensual layoff provisions do reduce firms' incentives to *re*invest in physical capital and other assets which may be general, but whose return is now partly captured by the worker in case of separation.

Relaxing the assumption of unconstrained entry fees, opens the possibility that consensual layoff arrangements may reduce firms' ex ante bargaining power. This would result not only in lower vacancy posting, as in Garibaldi and Violante

[11] See, for example, Nickell (1998) and Bean (1997).

[12] Though, they may have indirect effects if general training is a complement or substitute for other forms of investment.

(2000), but also in lower ex ante investment by firms. It has to be noted, though, that it is not obvious that side-payments from firms to workers are required in equilibrium. This depends not only on training costs, but also on whether it is workers' or firms' ex post bargaining power that exceeds its ex ante counterpart. In general this depends on the probability that each party's outside option is binding ex post[13].

Throughout our analysis, we have assumed that in the frictionless equilibrium firms rather than workers would like to unilaterally sever the relationship when productivity is low. The insight of the paper, though, applies equally to quits. Measures to prevent workers from quitting unless by mutual consent would further allow firms to capture part of the marginal return to their investment and improve incentives. We do not observe institutions of this kind, though. One would expect them not only to conflict with the natural law tenet that human capital cannot be alienated, but also to run into difficulties and possibly result in inefficient employment continuation in so far as workers are unable to buy out their jobs due to borrowing constraints. On the other hand, we do observe similar institutions when firms rather than workers are the non-investing party and natural rights or borrowing constraints are less of an issue. For example, top managers' effort is a typical example of general, worker-initiated investment which is vested in the firm of separation. It is highly common for firms to negotiate golden-handshakes when managers are removed.

[13]For example, under our assumption that workers' outside returns are never binding, it is firms that should pay an entry fee to skilled workers unless consensual layoff measures ensure that ex ante and ex post bargaining power coincide.

One point that this paper does not address is why firms invest in general training in the first place. Under realistic values for β the level of training would be higher if workers, rather than firms, invested. MacLeod and Malcomson (1993) have shown that simple fixed wage contracts allow the firm to capture the full marginal return to its investment with a very high probability. In such a set up it would be efficient for firms to carry out the investment provided that the probability that the workers' outside return is binding is low and the insight highlighted in this model would still apply. Extending the paper in this direction is a priority for future research.

4.4 Efficiency

It is well known that the decentralized equilibrium in a search environment without wage posting is not efficient unless the share parameter β happens to satisfy some variant of the Hosios (1990) efficiency condition and balance the thick market and congestion externalities. Apart from this special case, both job creation and job destruction are inefficient. As we have argued in the previous sections, the non-contractibility of investment introduces a further inefficiency. The theory of second best suggests that, even if consensual layoff provisions reduce the distortion associated with underinvestment in training, the end result may or may not be an increase in the flow of consumable resources.

As it turns out, independently from hold up issues, the mere coexistence of skilled and unskilled workers introduces a form of inefficiency that is absent from search models with homogeneous workers. For this reason, we will abstract from the investment decision in what follows and show that, conditional on any positive level of investment, the decentralized equilibrium is always inefficient. For ease of comparison we will assume that the level of investment is fully efficient in the decentralized economy of the previous section and will compare the decentralized and socially optimal job destruction and job creation decision. In appendix 4.B, we derive the condition for socially optimal investment and show that it always exceeds its decentralized counterpart.

The utilitarian social planner chooses a time path for the control variables, the beginning-of-period reservation productivity and market tightness pair (b, θ'), to maximize the present value of aggregate income. The corresponding value function solves the Bellman equation

$$\begin{aligned} L(u, u_s, \theta) &= \max_{b,\theta'} \frac{1}{1+r} \\ &\left\{ up(\theta) \left[f(h) \int_b^{\infty} z dG + G(b)\kappa \right] - (u - u^s) p(\theta) h - \left[\theta' u' - u(\theta - p(\theta))\right] \kappa - r\theta' u' \kappa + L' \right\} \end{aligned} \tag{4.36}$$

$$\begin{aligned} s.t. \qquad & u' = u\left[1 - p(\theta)\left(1 - G(b)\right)\right] + in' \\ & u'_s = u_s\left[1 - p(\theta)\right] + up(\theta)G(b) \end{aligned}$$

The social planner takes into account the evolution of the unemployment stock and of the number of skilled unemployed workers, but takes the demographics in' as given[14]. Aggregate income is defined as market output net of both investment

[14]Since the demographics in our model just ensures stationarity of the environment, it seems natural to assume that it cannot be controlled by the social planner.

costs and the opportunity cost $r\kappa$ of unfilled vacancies. Investment costs comprise the cost of training the number $(u - u^s)\,p(\theta)$ of untrained workers that find a match - the second addendum in equation (4.36) - plus the cost of opening new vacancies - equal to κ times the flow of new vacancies $\theta' u' - u(\theta - p(\theta))$. Note that θ, the lagged value of the control variable θ', enters the state space.

In what follows, rather than characterizing the social optimum for arbitrary initial conditions, we solve for the steady state.

The first order necessary conditions for the socially optimal reservation productivity and tightness are respectively

$$bf(h) = \kappa + L_u + L_{u_s} \tag{4.37}$$

and

$$\kappa\left(1 + \frac{r}{p'(\theta)}\right) = f(h)\int_b^\infty z dG + G(b)\kappa - (1 - G(b))\,(h + L_u)\,, \tag{4.38}$$

where L_u and L_{u_s} are the stationary partial derivatives of the value function. The above conditions are also sufficient under our assumptions of strict concavity and homogeneity of the matching function.

The first equation implies that separation is efficient when revenues from production fall below the value of physical capital κ plus the social value of a trained unemployed worker. The latter can be decomposed into the sum of the shadow price L_u of one more unskilled unemployed worker in the unemployment pool plus the value L_{u_s} of replacing one skilled for one unskilled worker, keeping the total size of the pool constant.

The second condition implies that the social cost of posting a vacancy, given by the investment cost κ plus the carryover cost - adjusted for the reduction in the duration of unemployment - must equal the expected social return. The latter takes into account the social opportunity cost L_u in case production takes place and the worker does not return to the unemployment pool[15].

By the envelope theorem, the steady state social value of a skilled and unskilled unemployed worker are respectively

$$L_{u_s} = \frac{p(\theta)}{r + p(\theta)} h \tag{4.39}$$

and

$$L_u = \frac{p(\theta)}{r + p(\theta)\left(1 - G(b)\right)} \left[f(h) \int_b^\infty z dG + G(b)\kappa - \kappa \left(1 + \frac{r}{q(\theta)}\right)\right] - \frac{p(\theta)}{r + p(\theta)} h. \tag{4.40}$$

At constant total unemployment, the only benefit from one more skilled worker in the pool is the saving of the cost h if the worker finds a job.

The social value L_u of an unskilled unemployed worker, instead, is the expected flow of output net of vacancy posting costs and of the cost of training her when she is matched with a firm for the first time.

It is useful to rewrite equation (4.38) by making use of the fact that $p(\theta) =$

[15]One may rightly note that it is a trained, not an untrained, worker that does not reenter the unemployment pool. Yet, keeping aggregate unemployment constant the steady state number of skilled unemployed is independent of market tightness as can be seen from equation (4.5).

$\theta q(\theta)$. If we call $\eta(\theta)$ the elasticity of $p(\theta)$ with respect to θ we can write (4.38) as

$$\kappa\left(1+\frac{r}{q(\theta)}\right) = f(h)\int_b^\infty z dG + G(b)\kappa - (1-G(b))\left(h+L_u\right) - \frac{r\kappa\left(1-\eta(\theta)\right)}{q(\theta)\eta(\theta)}. \tag{4.41}$$

Equations (4.40) and (4.41) together can then be used to rewrite the shadow value of an untrained unemployed worker as

$$L_u = \frac{\kappa\theta\left(1-\eta(\theta)\right)}{q(\theta)\eta(\theta)} - \frac{p(\theta)G(b)}{r+p(\theta)}h. \tag{4.42}$$

Let us write the private job creation condition in a form comparable to equation (4.41). To this purpose let us define the difference between the asset value of a skilled and unskilled matched worker as $e = E_a(h^*) - E_a(0)$. Using equations (4.6) and (4.12) privately optimal vacancy posting satisfies

$$\kappa\left(1+\frac{r}{q(\theta)}\right) = f(h^*)\int_b^\infty z dG + G(b)\left(\kappa + U(h)\right) - (1-G(b))\left(h^* - e\right) - \left(1+\frac{r}{p(\theta)}\right)U(h^*). \tag{4.43}$$

Under our assumption that investment in the decentralized equilibrium is socially optimal ($h = h^*$), we are now in a position to characterize the conditions for efficiency of the decentralized equilibrium conditional on a given level of training. Efficient vacancy posting requires the right hand sides of (4.41) and (4.43) to be equal, or

$$(1-G(b))\,L_u + \frac{r\kappa\left(1-\eta(\theta)\right)}{q(\theta)\eta(\theta)} = \left(1-G(b)+\frac{r}{p(\theta)}\right)U(h^*) - (1-G(b))\,e. \tag{4.44}$$

Efficient job destruction requires the private and social values of a skilled unemployed worker to be the same or, comparing equations (4.21) and (4.37),

$$U(h^*) = L_u + L_{u_s}. \tag{4.45}$$

One can use equations (4.16), (4.39) and (4.40) to rewrite (4.45) as

$$\Psi(\beta) = (r+p(\theta))(1-\beta)\left[f(h)\int_b^\infty zdG + G(b)\kappa\right] - (r+p(\theta)-\beta p(\theta)G(b))\left(1+\frac{r}{q(\theta)}\right)\kappa = 0. \tag{4.46}$$

It is easy to check that it is $\Psi(0) > 0$ and $\Psi(1) < 0$. Since $\Psi(.)$ is continuous the mean value theorem implies that (4.46) is satisfied for a value β^* of the sharing parameter strictly between zero and one. Let us assume that β takes exactly this value and derive the restrictions that this imposes on the differential e. This requires solving the system formed by (4.39), (4.45) and (4.44) for e as a function of L_u. The result is

$$L_u = \frac{\kappa\theta\,(1-\eta(\theta))}{q(\theta)\eta(\theta)} - \frac{p(\theta)}{r}\left[\frac{r+p(\theta)\,(1-G(b))}{r+p(\theta)}h^* - (1-G(b))\,e\right]. \tag{4.47}$$

Hence, the differential e has to ensure equality of (4.42) and (4.47). It can easily checked that this requires $e = h$.

With β taking care of job destruction, efficiency requires untrained workers to pay for the full cost of the training that the firm will provide. But, unless $\beta = 0$, unskilled workers pay for only a fraction of the total training cost h in the decentralized equilibrium, as can be seen from equation (4.15). With $\beta = 0$, though, $U(h^*)$ would be zero too and job destruction would be inefficiently high. The sharing parameter β alone is not sufficient to ensure full efficiency in this model. In the decentralized equilibrium the economy will always have either too little job creation or too much destruction or both. What we can rule out is the coexistence of inefficiently high duration and inefficiently high destruction.

For comparison, let us consider the case in which there is no investment in training; i.e. both h and L_{u_s} are zero and all workers are identical. Then, $e = 0$ is necessary and sufficient to equate the value of L_u in equation (4.42) and (4.47). If the sharing parameter is such as to ensure efficient separation then also job creation is efficient. This is the standard Hosios (1990) result[16]. With homogeneous agents all that is required to achieve efficiency on both the job creation and job destruction margins is that the sharing parameter β equates the private and social value of an unemployment worker.

The result that the Hosios condition is not sufficient to ensure efficiency in models with heterogeneous agents is not new. Bertola and Caballero (1994) show that when firms with heterogeneous productivity can choose the rate of vacancy posting at a convex cost, jobs creation at more productive units is inefficiently low in the absence of firm-specific subsidies. Shimer and Smith (1999) derive a similar result in a very general setting in which heterogeneous agents look for a match with endogenous search effort. They show that efficiency can only be achieved by subsidizing (taxing) the search effort of agents who are more (less) productive

[16]The only difference is that the optimal value of the sharing parameter β does not coincide with the elasticity of the probability of filling a vacancy due to the different bargaining solution adopted.

than average. In both these papers, the inefficiency stems from the inability of the share parameter alone to provide the correct incentives to search to heterogeneous agents.

Consider instead the case in which $h > 0$ and there are both skilled and unskilled workers in the unemployment pool. If the separation rate were exogenous, there would be just one active margin - the job creation one. It is easy, but tedious to show that there exists one value for the sharing parameter β that again ensures full efficiency. With workers' facing no active economic decision, apart from participation, efficiency only require that the private value of posting a vacancy coincides with its social counterpart.

In the present model there are two types of unemployed workers and two active margins: job creation and job destruction. As our previous discussion has shown, the reason is not workers' heterogeneity or the existence of more than one active margin per se. It is the heterogeneity *across* active margins that drives the result[17]. A searching firm can meet either a skilled or an unskilled workers, but all separations release a skilled unit of labour. For this reason, efficient entry of firms does not imply efficient reservation productivity. One way to achieve both is to ensure that both the private value of an unemployed worker equals its social counterpart and the differential between the value of a skilled and unskilled worker to equal its social value - the training cost.

Our model is different in so far as search intensities are exogenous and it is workers - who face no active decision- who are heterogeneous. The fact that the

[17]If training were fully specific, heterogeneity across active margins would disappear and an appropriate value for the share parameter would ensure efficiency.

inefficiency survives intact highlights in a simple way the general logic behind all these results. Efficiency in search models, requires private and social values to coincide across all *active* margins that affect matching probabilities. In so far as agents are heterogeneous across such margins the decentralized equilibrium cannot be efficient.

4.5 Conclusion

This chapter has analysed non-contractible firms' investment in general human capital in a model of frictional unemployment. General training increases workers' productivity with other employers but is vested in the worker on separation. This depresses investment as no return accrues to the firm on separation. We have shown that consensual layoffs, by obliging firms to share the total payoff from separation, improve employers' incentives to train. The mechanism applies to all forms of general investment that is vested in the non-investing party. It applies equally to workers' investment to improve product quality and develop new products that remain intellectual property of their employers.

We have also shown that, independently from underinvestment in training, the *laissez-faire* equilibrium is always inefficient for any given level of investment. The coexistence of skilled and unskilled workers implies that the Hosios (1990) condition fails to ensure equality between social and private values for both skilled and unskilled workers. Since workers are heterogeneous along the job creation and the job destruction margins the sharing parameter alone cannot ensure efficiency.

Appendix 4.A: Proofs of propositions

Lemma 2. *If* $(1-\beta)U(h^*) < \beta\kappa$, *then for*

$$F < \kappa - (1-\beta)bf(h^*) \tag{4.48}$$

there exists $z_r \in [b, \infty)$

$$F = \kappa - (1-\beta)z_r f(h^*) \tag{4.49}$$

such that $\forall z \leq z_r$, $J_p(z, h^*) = \kappa - F$.

Viceversa, if $(1-\beta)U(h^*) > \beta\kappa$, *then,* $\forall F$, *there exists* $z_r \in [b, \infty)$ *satisfying*

$$U(h^*) = \beta z_r f(h^*) \tag{4.50}$$

such that $\forall z \leq z_r$, $E_a(z, h^*) = U(h^*)$.

Proof. Suppose, by contradiction, there is no $z_r \in [b, \infty)$ such that either party's outside option is binding. Then, by proposition 1, it has to be $J_p(z, h^*) = (1-\beta)zf(h^*)$ $\forall z \in [b, \infty)$. By continuity of revenues in z then

$$U(h^*) = bf(h^*) \tag{4.51}$$

and

$$\kappa = (1-\beta)\, bf(h^*). \tag{4.52}$$

But then (4.51) and (4.52) imply $(1-\beta)U(h^*) = \beta\kappa$ which contradicts either assumption.

The inequality $(1-\beta)\,U(h^*) < \beta\kappa$ implies

$$\frac{\kappa}{1-\beta} > U(h^*) + \kappa = bf(h^*) \tag{4.53}$$

or

$$\kappa > (1-\beta)\,bf(h^*). \tag{4.54}$$

Hence, by continuity there exists $z_r > b$ such that

$$\kappa = (1-\beta)\,bf(h^*) \tag{4.55}$$

as long as $F < \kappa - (1-\beta)\,bf(h^*)$.

Symmetrically, it can be shown that the reverse inequality $(1-\beta)\,U(h^*) > \beta\kappa$ implies that it is the worker's outside option $U(h^*)$ which is binding for some $z_r \in [b,\infty)$. z_r is unaffected by severance payments in this case as the worker's outside option is not. ■

Corollary 3. *If $(1-\beta)U(h^*) < \beta\kappa$ and $F > \kappa - (1-\beta)bf(h^*)$ then the firm's payoff a time t.3 is given by*

$$J_p(z,h^*) = (1-\beta)\max\{zf(h^*), U(h^*) + \kappa\}. \tag{4.56}$$

Proof. The inequality $F > \kappa - (1-\beta)bf(h^*)$ implies that as long as $z \geq b$ the firm's is better of sharing the payoff from continuation rather than firing the worker. So, $z_r \notin [b,\infty)$. Remembering that $bf(h^*) = U(h^*) + \kappa$, it also implies that when $z < b$, it is optimal for the firm to negotiate a voluntary severance payment that leaves the worker a share β of the total payoff from separation $U(h^*) + \kappa$ rather than paying the legislated severance payment F. The inequality $(1 - \beta)U(h^*) < \beta\kappa$ implies that the worker would not leave voluntarily without such a payment. ■

Appendix 4.B: Socially optimal training

Let us define by $S_U(h)$ the social value of a trained unemployed worker and by $S_E^e(h)$ the expected social surplus associated with a matched skilled worker at time $t.2$. We can write

$$(p+r)S_U(h) = pS_E^e(h) \tag{4.57}$$

and

$$S_E^e(h) = f(h)\int_b^\infty zdG + G(b)\left(S_U(h) + \kappa\right). \tag{4.58}$$

Solving for $S_E^e(h)$ we can write

$$S_E^e(h) = \frac{r+p(\theta)}{r+p(\theta)\left(1-G(b)\right)} f(h)\int_b^\infty zdG + G(b)\kappa. \tag{4.59}$$

The socially optimal level of training satisfies $1 = \partial S_E^e(h)/\partial h$ or

$$1 = \frac{r+p(\theta)}{r+p(\theta)\left(1-G(b)\right)} f'(h)\int_b^\infty zdG. \tag{4.60}$$

It is straightforward to see that investment is always inefficiently low in the decentralized equilibrium as the right hand side of (4.32) is always smaller than the right hand side of (4.60) for any value of β.

Chapter 5

Dismissal costs and efficiency wages *

5.1 Introduction

The economic literature on job security has been mainly concerned with the possible costs of dismissal regulation: lower firms' profitability and distorted labour allocation. Yet, it is not obvious that profits have to fall as a consequence of dismissal costs. First, wages could fall at unchanged revenues, as in Lazear (1990) and Bertola (1996a). Second, revenues could increase more than wages. Job security could induce workers to invest more in firm-specific human capital or produce higher effort, thus increasing revenues. This view is quite common in the indus-

*An abridged version of this chapter is forthcoming in the European Economic Review. I am indebted to Charlie Bean for many useful comments. Giuseppe Bertola provided a crucial modelling suggestion and insightful remarks. Andrea Ichino, Claudio Michelacci, seminar participants at Research Strategy Seminars at LSE and two anonymous referees made useful comments.

trial relations literature[1] and receives indirect support by the fact that a number of firms voluntarily commit to forms of job security. Apart from the too notorious case of Japanese firms, a number of US firms such as DEC, Eli Lilly, Eastman Kodak, IBM have adopted a policy of no-layoffs. In the UK a significant proportion of contractual arrangements contain severance payment provisions well in excess of statutory minima. It seems likely that, at least for the firms which commit to it, job security should result in improved performance. So, it seems worth exploring possible mechanisms which can explain this stylised fact.

We have seen in the previous chapters that for dismissal costs to have any role, apart from a purely redistributive one, utility has to be non-transferable. This may be due either to exogenous constraints on *ex post* recontracting, as in Booth (1997) where right-to-manage collective agreements prevent redundant workers from renegotiating their wages, or market incompleteness, as in the efficient-risk-sharing model of Bertola (1996a). Yet, strong commitment to job security is not restricted to union firms and its introduction is not necessarily accompanied by wage concessions on the part of workers. Relatively recent agreements introducing strong forms of job security in the US automobile sector aimed at obtaining increased workers' cooperation and flexibility rather than extracting wage concessions[2]

Efficiency wages are not only a way to give formal content to many of the mechanisms so often emphasized by the industrial relations literature such as

[1] See, for example, Piore (1986) and Beuchtemann (1992).

[2] The concessions that Ford and General Motors required from the United Automobile Workers union were acceptance of flexibility in work organization and the right of the companies in making outsourcing and introducing new technologies. See Osterman and Kochan (1990).

long term attachment and participation[3]. They are also one of the few convincing explanations for the non-transferability of utility and inefficiency of separation in labour market relationships.

This chapter explores the effect of dismissal costs on effort[4]. It analyses the general equilibrium effect of severance payments in an economy featuring involuntary labour mobility, as firms pay efficiency wages in order to prevent shirking as in Shapiro and Stiglitz (1984). Though the empirical evidence on the bearing of monitoring considerations on wage rigidity is rather weak[5], its fully dynamic nature makes it an ideal candidate to analyse intertemporal issues, such as job security.

The model is a simplified version of Saint-Paul[6] (1995) insofar as it assumes, as in Bertola (1990), that the idiosyncratic productivity shock follows a two-state Markov process rather than a Poisson process with a continuum of states. All the results in the original model are preserved. Severance payments allow firms to commit to fire less in downturns. Since mobility has to be costly in order to promote effort, a lower dismissal probability results in a lower wage bill in good times. Also, a pure redundancy pay leaves hiring unaffected but reduces firing and so increases aggregate employment on impact.

[3]See Katz (1986) for a survey of efficiency wage theories. Salop and Salop (1976) and Akerlof (1984) argue that paying above market clearing wages may, respectively, reduce turnover and increase workers' cooperation.

[4]Given that the model assumes binary effort choices, job security can only affect the cost of promoting a *given* level of effort. Yet, the intuition is preserved.

[5]The survey studies by Kaufman (1984) and Blinder and Choi (1992) both point out that reducing turnover and buying workers' cooperation are the most common explanations for real wage rigidity that managers provide.

[6]I came across Saint-Paul's article only after completing a first draft of this chapter. The focus of the chapter is on aggregate employment and efficiency, while Saint-Paul's paper deals with the possibility that efficiency wages may generate employment persistence.

The simplified nature of the model, though, provides additional insight. Firing costs, even if they do not accrue to workers, unambiguously increase employment in general equilibrium. More interestingly, and this is the main contribution of the chapter, a pure severance payment monotonically increases welfare for any size up to the amount that maximizes the optimized value of hiring firms and equalizes wages across states. As incentive compatibility *requires* workers to enjoy state-independent rents, any offer by redundant workers to take a wage cut is not credible, since it is not *ex post* incentive compatible. Firms enjoying good business conditions would like to commit to lower firing *ex ante,* in order to reduce their wage bill. In the absence of reputational or other commitment devices though, their pledge is not time consistent and the separation rate is inefficiently high, as the externality, in the form of foregone rents, that firms impose on workers on severance cannot be traded either in spot or in future markets. A pure severance payment provides a commitment device. Acting as a Pigovian tax, it increases welfare, as it makes firms internalize the externality and fire less, while having no negative effect on hiring, as workers accept lower wages in good states in exchange for a lower mobility cost in case of economic (but not disciplinary) dismissal. Welfare may still increase even if statutory job security provisions depress hiring by introducing a deadweight loss between the cost to the firm and the indemnity received by workers.

The model also predicts that dismissal payments, as a proportion of the pre-displacement wage, should be increasing with the difficulty of monitoring workers. Rents account for a bigger proportion of wages for workers which are more difficult to monitor or motivate.

The chapter is structured as follows. Section 5.2 introduces the economic environment and describes the model. Section 4.3 analyses the equilibrium. Section 4.4 presents the normative result. Section 4.5 concludes.

5.2 Economic environment

We consider a labour market populated by a continuum of firms indexed by n, where $n \in [0,1]$, and a homogeneous (hence anonymous) labour force normalized to one. Firms are price-takers on the product market. As in Shapiro and Stiglitz, workers are risk-neutral and have an instantaneous utility function given by $u(c, e) = c - eI_s$, where c is consumption (assumed to be non-negative), e is effort, measured in units of consumption and I_s is an indicator function equal to 0 if the worker shirks or is unemployed and 1 otherwise. Each individual is endowed with one unit of labour. Effort choices are discrete: workers can decide whether to expend zero effort (shirk) or the positive amount e. Firms can only imperfectly monitor individual effort and cannot infer it from the total amount of output produced. If a worker shirks, she faces and instantaneous probability q of being caught and fired. Monitoring is independent across workers. So the probability that two workers are caught shirking in the same time instant is zero.

The assumptions relative to asset markets and the structure of uncertainty are the same as in chapter 2. We conveniently assume that capital markets are perfect. Equilibrium considerations then require the market real interest rate to coincide with the subjective discount rate and the timing of consumption is immaterial[7]. This allows us to write the utility associated with employment at firm n as $u(w_n, e) = w_n - eI_s$, where w_n is the wage at firm n. The utility flow associated with unemployment is zero. So is the output of a worker who shirks.

Output is produced using only labour according to the (strictly) concave production function

$$R(l_n; \alpha_n), \qquad\qquad R' > 0,\ R'' < 0.$$

where l_n is the employment level at firm n. α_n indexes the revenue function and, as in Bertola (1990), can take only two values: α_b if firms are hit by a bad productivity shock and α_g if a firm enjoys a favourable supply shock[8]. Suppose $\alpha_g > \alpha_b$ and $\partial R'/\partial\alpha > 0$, so that labour demand is higher for firms in the good state.

α_n follows a continuous Markov chain with symmetric transition probabilities given by

$$\alpha_n(t+dt) = \begin{cases} \alpha_g \text{ with prob. } pdt \text{ if } \alpha_n\ (t) = \alpha_b, \text{ with prob. } (1-pdt) \text{ if } \alpha_n\ (t) = \alpha_g \\ \alpha_b \text{ with prob. } pdt \text{ if } \alpha_n\ (t) = \alpha_g, \text{ with prob. } (1-pdt) \text{ if } \alpha_n\ (t) = \alpha_b \end{cases}.$$

This implies that the ergodic probability that a given firm is enjoying good or bad business conditions is 0.5 and, since the number of firms is infinite, this is also the steady state proportion of firms in each state. Given that the total mass of firms is one, steady state aggregate employment equals $(l_g + l_b)/2$, where l_b and l_g are

[7] See footnote 10 in chapter 2.

[8] The shocks could be alternatively interpreted as changes in the relative demand for a differentiated product.

employment at the representative firm in the bad and the good state respectively.

We also assume $R'(l_g;\alpha_g) = R'(2-l_g;\alpha_b) = e$; i.e. the marginal product of labour at full employment equals its social cost. This ensures that full employment is efficient and jobs are rationed[9].

Firms cannot freely fire workers in downturns. They have to pay a redundancy cost F for each worker fired. As in the previous two chapters, a proportion $Q \leq F$ of the cost born by the firm accrues to the worker, thus allowing for red-tape administrative costs. For simplicity, we take both Q and F to be independent of workers' tenure. The payment is either explicitly contracted or imposed through legislation.

We assume that workers do not receive severance pay in case of disciplinary dismissal, as it is the case in most countries. As pointed out in MacLeod, Malcomson and Gomme (1994), the non-verifiability of workers' performance by third parties is necessary to ensure that an efficient contract is of the efficiency wage type rather than being performance-related. Carmichael (1983) notes that if performance is non-verifiable, the existence of a conditional severance payment introduces an incentive to misrepresent the reason for separation. In the present framework contracting firms would claim that the relationship was terminated because the worker was caught shirking. Viceversa, workers fired for malfeasance would claim to have been dismissed for economic reasons to obtain a severance payment. If a firm employs just one worker, a conditional severance cannot be enforced and the

[9]That the marginal product of labour at full employment equals (rather than being above) its social cost is sufficient, not necessary to have job rationing. Market anonimity and job rationing ensure that it is (constrained) efficient for workers to receive rents and post no bond as argued in MacLeod and Malcomson (1993c). This makes it clear that our result does not rely on any disguised bonding argument.

efficient contract has no severance payment, as argued in MacLeod, Malcomson and Gomme (1994). Under our assumptions that firms employ more than one worker and that the probability that two workers are contemporaneously detected shirking is negligible, such a payment is enforceable[10]. A firm which is firing more than one worker is automatically signalling that the dismissals have economic causes. If it tried to cut the wages of the redundant workers by such an amount as to induce them to shirk and fire them costlessly, this would be taken as unfair behaviour by the remaining employees who would reduce effort[11]. In practice, in so far as labour courts and arbitration bodies can only imperfectly ascertain whether a disciplinary dismissal is warranted or not, the expected severance payment for a worker caught shirking may be positive. This might result in higher workers' rents in the presence of constraints on side payments from workers to firms. We discuss this further in section 4.4.

5.2.1 Hiring and firing

In the presence of turnover costs the firms' optimization problem is intertemporal. So the shadow value of the marginal worker at firm n, J_n, must satisfy

$$rJ_n = R'\left(l_n; \alpha_n\right) - \varphi_n + \frac{E\left[dJ_n\right]}{dt}, \tag{5.1}$$

[10] In practice, at least for sufficiently large firms the scale of layoffs is a meaningful signal of the reason for termination.

[11] On the importance of fairness issues in this class of models see MacLeod and Malcomson (1998).

where $r \geq 0$ is the time-independent discount rate and φ_n is the marginal cost of employment; i.e. the derivative of the wage bill with respect to firm-level employment. For any firm the flow equivalent of the shadow value of the marginal worker equals the instantaneous marginal profit plus the expected capital gain/loss. Given that the assumed Markov process is time-invariant and firms are identical, n indexes business conditions.

Linear turnover costs determine an optimal inaction range, so $l_g \geq l_b$ with equality if firms find it profitable not to respond to a shock. We assume that the change in business conditions is such as to generate positive turnover; i.e. $l_g > l_b$. A firm which has just been hit by a negative shock will fire workers up to the point where the shadow value of the marginal job equals minus the dismissal cost; i.e. $J_b = -F$. Replacing in (5.1) we get

$$-rF = R'(l_b; \alpha_b) - \varphi_b + p(J_g + F). \tag{5.2}$$

The last term is the expected change in the value of the job. The firm turns good with instantaneous probability p and the value of the job changes from $-F$ to J_g.

For a firm in the good state the shadow value of the marginal worker J_g is given by

$$rJ_g = R'\left(l_g; \alpha_g\right) - \varphi_g - p(F + J_g). \tag{5.3}$$

As opening a new job is costless, firms in the good state will increase employment up to the point where $J_g = 0$. Replacing J_g in (5.2) and (5.3), labour demand at firms in the bad and the good state is then implicitly given by

$$\varphi_b = R'\left(l_b; \alpha_b\right) + \left(p + r\right) F, \tag{5.4}$$

$$\varphi_g = R'\left(l_g; \alpha_g\right) - pF. \tag{5.5}$$

As argued in chapter 2, equations (5.4) and (5.5) show that firing costs introduce a wedge between the marginal cost and the marginal product of labour. In "bad" firms the marginal cost is higher than the marginal productivity while the opposite is true in "good" firms. For a given marginal cost, firing costs increase employment in the former and reduce it in the latter through the well known option-value effect.

5.2.2 Workers' behaviour

Workers behave exactly as in Shapiro and Stiglitz. If we abstract from turnover costs for a moment, the only difference is that the separation rate is endogenous and depends on firms' transition probabilities.

Also workers' decisions are intertemporal and must satisfy asset-value equations. Subscripts indicate the firms' state and superscripts the worker characteristics (s if a shirker, n if not).

E_b^i, the value of being employed at a firm in the bad state for a worker of type i ($i = n, s$) is given by

$$rE_b^s = w_b + q\left(U - E_b^s\right) + p\left(\max\left[E_g^s, E_g^n\right] - E_b^s\right), \tag{5.6}$$

$$rE_b^n = w_b - e + p\left(\max\left[E_g^s, E_g^n\right] - E_b^n\right), \tag{5.7}$$

where U is the expected utility of being unemployed. According to (5.6) and (5.7), the value of being employed at a firm which is in a bad state, for both shirkers and non-shirkers, is given by the current wage plus the expected capital gain if the firm turns good with instantaneous probability p. When the firm's productivity changes, workers adopt the most profitable behaviour between shirking and not shirking. Their values differ insofar as non-shirkers bear the cost e of supplying the effort while shirkers are subject to a capital loss if they are detected shirking, an event which has an instantaneous probability q.

Equations (5.8) and (5.9) below are the equivalent of (5.6) and (5.7) for the workers employed at a firm in the good state. We assume workers are randomly laid off.

$$rE_g^s = w_g + q\left(U - E_g^s\right) + p\left(1 - \frac{l_b}{l_g}\right)\left(U + Q - E_g^s\right) + p\frac{l_b}{l_g}\left(\max\left[E_b^s, E_b^n\right] - E_g^s\right) \tag{5.8}$$

and

$$rE_g^n = w_g - e + p\left(1 - \frac{l_b}{l_g}\right)\left(U + Q - E_g^n\right) + p\frac{l_b}{l_g}\left(\max\left[E_b^s, E_b^n\right] - E_g^n\right) \tag{5.9}$$

are the values of being employed at a firm experiencing good business conditions for a worker of type i. Given that on transiting into the bad state a firm will hold only a proportion l_b/l_g of its original labour force, all workers, shirkers and non-shirkers alike, face an instantaneous probability $p(l_b/l_g)$ of keeping their job but at a firm in the bad state and a probability $p(1 - l_b/l_g)$ of becoming unemployed.

The permanent income of an unemployed worker is

$$rU = h\left(\max\left[E_g^s, E_g^n\right] - U\right), \tag{5.10}$$

where h is the exit rate from unemployment. As only firms in the good state hire new workers, the value of being unemployed equals the expected capital gain if hired by a firm whose business conditions have improved.

In each time interval the number of firms which switch from the good to the bad state is $p/2$ and each of them will fire $(l_g - l_b)$ workers. In steady state the outflow from unemployment must equal the inflow into it. So the exit rate from unemployment is implicitly given by

$$hu = \frac{p}{2}(l_g - l_b), \tag{5.11}$$

where u is the stock of unemployed workers which equals

$$u = 1 - \frac{l_g + l_b}{2} \tag{5.12}$$

given that in steady-state there are 0.5 firms in each state.

So we can rewrite (5.10) as

$$rU = h\left(\bar{l}_b, \bar{l}_g\right)\left(\max\left[E_g^s, E_g^n\right] - U\right). \tag{5.13}$$

where bars over variables indicate that the exit rate from unemployment is a function of average firm-level employment that firms take as given.

It is easy to verify that E_j^n and E_j^s $(j = b, g)$ are linear increasing functions of w_j with E_j^n having a lower intercept and higher slope than E_j^s ; i.e. they satisfy the single-crossing property. So, to prevent workers from shirking, firms have an incentive to raise wages up to the point where $E_j^n = E_j^s$. Subtracting respectively (5.6) from (5.7) and (5.8) from (5.9) and rearranging we get

$$E_g - U = E_b - U = \frac{e}{q}. \tag{5.14}$$

Equation (5.14) implies that the punishment from being caught shirking and fired must equal the expected effort a shirker would save before being caught. Keeping in mind that $E_j^n = E_j^s$ we can replace in (5.7) and (5.9) using (5.13) and (5.14) to get the incentive compatible wages

$$w_b = e\left(1 + \frac{r}{q}\right) + h\left(\bar{l}_b, \bar{l}_g\right)\frac{e}{q} \tag{5.15}$$

and

$$w_g = e\left(1 + \frac{r}{q}\right) + h\left(\bar{l}_b, \bar{l}_g\right)\frac{e}{q} + p\left(1 - \frac{l_b}{l_g}\right)\left(\frac{e}{q} - Q\right). \tag{5.16}$$

The intuition behind equations (5.15) and (5.16) is the following. The second term in both equations is the permanent income from unemployment. The first one is the compensation for providing the effort. For workers not to shirk they must be compensated not only for the effort put forth but also for the flow opportunity cost of shirking. The third term in equation (5.16) is the premium that firms in the good state have to pay in order to compensate their workers for the mobility cost they bear in case they are made redundant. The premium is positive if $Q < e/q$. Mobility *has* to be painful for disciplinary dismissals to provide an effective punishment. In equilibrium no worker shirks, but the fact that mobility has to be costly to prevent shirking imposes a negative externality on workers who are dismissed for economic reasons. The wage differential - the last term in equation (5.16) - has to compensate for this cost. Only when p tends to 0 - i.e. when idiosyncratic shocks tend to be permanent - do all firms pay the same wage. In this limit case no worker is dismissed whatsoever in equilibrium and so no compensation is required.

The redundancy pay Q lowers w_g, with respect to the frictionless case, as it reduces the expected loss from being dismissed. Workers at firms in the good state are willing to accept a lower wage now in exchange for the severance payment when fired. The transfer lowers the mobility cost for workers who are dismissed for economic reasons while leaving unaffected the punishment for potential shirkers. This reduces the wage differential.

The wage differential disappears when the statutory cost that the firm has to bear when firing a worker exceeds the worker's mobility cost e/q. If this is the case, it is $F = Q = e/q$. As argued in chapter 3, when the statutory cost to the firm exceeds the mobility cost for the worker, it has no marginal effect on the parties' decision to continue the relationship as there exists a jointly optimal side payment - in this case e/q - that induces workers to quit efficiently whenever their shadow value falls below their rent. Unless otherwise specified, we assume throughout that the statutory dismissal cost is lower than workers' mobility cost.

5.2.3 Wage setting

Equation (5.15) shows that firms in the bad state are effectively wage-takers. The exit rate from unemployment and the unemployed permanent income depend on average variables that firms take as given. So the marginal cost of employment for a firm in the bad state is given by

$$\varphi_b = w_b = e\left(1 + \frac{r}{q}\right) + h\left(\bar{l}_b, \bar{l}_g\right)\frac{e}{q}. \tag{5.17}$$

Firms in the good state, instead, control the premium they pay over wages in the bad state. Their employment choice affects the probability that a worker enters unemployment when business conditions turn bad, as can be seen from equation (5.16). Their wage bill is

$$w_g l_g = \left[e\left(1 + \frac{r}{q}\right) + h\left(\bar{l}_b, \bar{l}_g\right)\frac{e}{q}\right] l_g + p\left(\frac{e}{q} - Q\right)\left(l_g - l_b\right). \tag{5.18}$$

The last term looks exactly like a cost of turnover and is conceptually equivalent. Firms in the good state have to compensate their workers for the risk of being dismissed. So turnover imposes a cost on firms in the form of a higher wage bill. Their marginal cost of employment, $\partial\left(w_g l_g\right)/\partial l_g$, is then

$$\varphi_g = e\left(1+\frac{r}{q}\right) + h\left(\bar{l}_b, \bar{l}_g\right)\frac{e}{q} + p\left(\frac{e}{q} - Q\right). \tag{5.19}$$

As in Lazear (1990), the marginal employment cost for hiring firms falls by the full size of the payment.

Note also that, for given l_g, the risk of becoming unemployed is lower for higher l_b. So employment at firms in the bad state has a positive externality on firms in the good state. The marginal benefit is $p\left(e/q - Q\right)$ as can be seen from equation (5.18). This important point was made by Saint-Paul (1995). Higher employment in downturns reduces the current wage bill in the good state by decreasing the dismissal probability $p\left(1 - l_b/l_g\right)$. The value of the firm is improved up to the point where the marginal current benefit from firing less in a downturn equals the marginal expected discounted loss from having to hoard labour in the bad state. As for the optimized value of firing firms, it is reduced by labour hoarding for a positive discount rate. At the moment of transiting into the bad state, any reduction in the wage bill in the good state is already sunk. The marginal expected gain accrues in the future and is discounted while the current marginal loss is not.

A time-inconsistency problem exists. Firms in the good state would like to commit *ex ante* to fire less, but, in the absence of reputational or other commitment devices, their pledge is not credible as it would not be optimal from the firm's point of view *ex post.*

Severance payments allow a hiring firm to credibly commit to fire less in the event that business conditions turn bad[12].

Saint-Paul (1995) demonstrates that the level of redundancy pay which maximizes the optimized value of a firm in the good state is $Q = F = e/q$. Equations (5.15) and (5.16) show that this is also the amount that equalizes wages across states by eliminating the mobility cost involuntarily born by dismissed workers. Efficiency wage considerations do not require workers to actually bear any mobility cost, but only the *threat* of facing one in the event of underperformance.

5.3 Market equilibrium

5.3.1 Definition

In equilibrium firms choose employment so that its (shadow) marginal product, the right-hand-side of equations (5.4) and (5.5), equals its marginal cost, the right-hand-side of (5.17) and (5.19). So replacing for φ_b and φ_g using (5.17) and (5.19) in (5.4) and (5.5) respectively, results in the two equilibrium conditions

$$R'(l_b; \alpha_b) = e\left(1 + \frac{r}{q}\right) + h(l_b, l_g)\frac{e}{q} - (p + r)F \tag{5.20}$$

and

$$R'(l_g; \alpha_g) = e\left(1 + \frac{r}{q}\right) + h(l_b, l_g)\frac{e}{q} + p\left(\frac{e}{q} - Q + F\right), \tag{5.21}$$

[12]Obviously, firms and workers could achieve the same result by writing an explicit state-contingent employment contract. However, in a more realistic set up in which shocks to productivity can take more than just two values, writing an explicit contract that fixes a unique, state-independent, level of the severance payment is much easier than writing a contract that specifies the number of workers to fire in each possible state of the world.

where h is given by equations (5.11) and (5.12) and bars over l_b and l_g have been dropped as all firms in the same state are identical.

5.3.2 The effect of dismissal costs

Equation (5.20) shows that dismissal costs induce labour hoarding at contracting firms (on impact). As argued in Chapter 3, since mobility is costly for workers in the frictionless equilibrium, there is no room for mutually beneficial trade to undo the effect of dismissal costs as long as the legislated firing cost is smaller than workers' mobility cost. Since firing a redundant worker is cheaper that inducing her to quit, firing costs are binding and employment in downturns is higher, on impact, than in the frictionless equilibrium.

Lazear's point that, with perfect asset markets, dismissal costs have real effects only if a part $(F - Q)$ of the cost born by the firm does not accrues to workers, applies only to hiring. As can be seen from equation (5.21), dismissal costs involving third-party payments depress employment at firms in the good state, provided the statutory cost does not exceed workers' mobility cost e/q.

Dismissal costs are always non-neutral when real wages are downward rigid, though. F, the cost born by the firm, unambiguously reduces firm-level turnover and the inflow into unemployment - the right-hand-side of equation (5.11). For equilibrium to be reestablished, the outflow from unemployment, the left-hand-side of (5.11), has to fall. So either the unemployment rate u or the exit rate h or both must fall. Suppose first that $F > Q = 0$ (i.e. F is a pure red-tape administrative cost). Firing costs have only small and ambiguous effects on average labour demand as proved by Bertola (1990, 1992). Intuitively, since the effect on the inflow into unemployment is an order of magnitude higher than the impact effect on the unemployment rate, the exit rate from unemployment has to fall to reestablish equilibrium. This further boosts employment as it reduces the permanent income from unemployment and wages in both states. If $F = Q$ instead, it is clear from equations (5.20)-(5.21) that severance payments not only reduce turnover, but have a positive impact effect on aggregate employment. The direction of the general equilibrium change in the exit rate from unemployment is, then, ambiguous and so is the effect on wages. We show in Appendix 4.A that the positive impact effect always prevails and dismissal costs always result in higher aggregate employment, whether totally ($F = Q$) or partially ($F > Q$) received by workers[13].

Yet, higher employment does not necessarily result in higher aggregate welfare. In a first-best world severance payments cannot improve on the decentralized outcome. If mobility is involuntary costly for workers, dismissal costs are potentially welfare improving, as argued in chapter 3. In efficiency wage models in which profits are entirely distributed to workers the *laissez faire* equilibrium is not, in general, Pareto optimal - unemployment is inefficiently high. This is the case, for example, in Shapiro and Stiglitz's model and has been proved for the general case in Greenwald and Stiglitz (1988). In our model aggregate welfare depends also on the cross-sectional distribution of employment. An employment-increasing

[13]The result does not depend on the simple stochastic structure of the model. It does not rely on the fact that firms in the bad state do not carry out redundancies, but on the fact that severance payments do not push up wages. Also, allowing for entry/exit of firms is likely to strengthen the result, as the value of hiring firms is increased while the fall in the value of firing firms is small if r is, realistically, close to zero.

intervention may or may not improve welfare depending on how it affects the distribution of employment across firms.

5.4 Efficient redundancy pay

For simplicity we discuss efficiency in terms of steady state undiscounted optimization. By continuity the qualitative result is unchanged for a small enough positive r.

The social planner maximizes steady state consumer utility. Given linear preferences this is tantamount to maximizing the utility of working plus any income transfer - namely profits and the severance payment Q. This is equivalent to maximizing the difference between total output and its social cost of production minus any waste of resources associated with redundancy pay; i.e.

$$W = \frac{1}{2}\left(R\left(l_b;\alpha_b\right) - el_b\right) + \frac{1}{2}\left(R\left(l_g;\alpha_g\right) - el_g\right) - \frac{p}{2}(l_g - l_b)(F - Q). \qquad (5.22)$$

Provided they are fully received by workers ($F = Q$), redundancy payments have no direct effect on aggregate welfare because they are just a redistribution[14]. Obviously they do affect welfare through their effect on the allocation of labour. Since severance payments cancel out in (5.22) and leave the value of firms unaf-

[14]The fact that the gains may accrue to individuals who are different from the ones who bear the losses does not matter *ex ante* and, given no discounting, not even *ex post* since, as the individual distribution coincides with the cross-sectional one, all individuals have the same expected utility, .

fected[15], they do not alter the social planner problem and we can harmlessly set them to zero in the welfare analysis.

The social planner has no monitoring advantage over private agents. She has to pay incentive compatible wages. Furthermore she cannot enforce allocations which are not profitable from the private point of view. So she maximizes (5.22) subject, not only to the wage constraints (5.15) and (5.16), but also to

$$\frac{1}{2}\left[R\left(l_b;\alpha_b\right)-w_bl_b+R\left(l_g;\alpha_g\right)-w_gl_g\right]\geq 0. \tag{5.23}$$

The intuition behind the expression in (5.23) is the following. The social planner can at best tax and redistribute pure profits. So total aggregate profits (or equivalently the value of firms) cannot be negative.

Since wages are a pure transfer, but reduce the total amount of resources out of which employment taxes/subsidies can be financed, the social planner will never pay wages in excess of the incentive compatible level. So constraint (5.23) can be rewritten as

$$\Pi=\frac{1}{2}\left[R\left(l_b;\alpha_b\right)-w_b(l_b,l_g)l_b+R\left(l_g;\alpha_g\right)-w_g(l_b,l_g)l_g\right]\geq 0, \tag{5.24}$$

where $w_b(\cdot,\cdot)$ and $w_g(\cdot,\cdot)$ are the incentive compatible wages in equations (5.15) and (5.16).

[15]The optimized value of the firms in the two states satisfies the Bellman equations $rV_g = R\left(l_g;\alpha_g\right)-w_gl_g+p\left(V_b-V_g\right)-pF\left(l_g-l_b\right)$, $rV_b = R\left(l_b;\alpha_b\right)-w_bl_b+p\left(V_g-V_b\right)$. If $F=Q$, replacing for w_g confirms that the value of firms is unaffected by dismissal costs at unchanged employment levels.

The social planner chooses l_b and l_g to maximize the Lagrangean

$$\mathfrak{L} = W + \lambda \Pi. \tag{5.25}$$

We know from Shapiro and Stiglitz (1984) that the "invisible hand" of the market does not achieve the social optimum in their model. Employment is inefficiently low in the decentralized equilibrium as the private cost of labour is above its social cost. Given that they assume an exogenous separation rate, this means that hiring and, hence, the exit rate from unemployment are inefficiently low in the decentralized equilibrium. In the present model, firms heterogeneity implies that aggregate welfare depends also on the cross-sectional employment allocation. We define an allocation as locally efficient if it is not possible to increase aggregate welfare by exchanging l_b and l_g at a rate which leaves aggregate profits unchanged. By letting λ in equation (5.25) vary, we obtain a pseudo-contract[16] curve - the set of locally efficient allocations, one for each particular level of aggregate profits chosen. The first best allocation is the one associated with the λ that maximizes (5.25).

Points on the pseudo-contract curve must satisfy the two first order conditions

$$R'(l_b;\alpha_b) - e + \lambda\left(R'(l_b;\alpha_b) - \varphi_b - (l_b + l_g)\frac{\partial h}{\partial l_b}\frac{e}{q} + p\frac{e}{q}\right) = 0 \tag{5.26}$$

[16]The locus of efficient allocations is not a true contract curve if market incompleteness prevents the parties from contracting to achieve local efficiency.

and

$$R'\left(l_g;\alpha_g\right)\ -e+\lambda\left(R'\left(l_g;\alpha_g\right)-\varphi_g-\left(l_b+l_g\right)\frac{\partial h}{\partial l_g}\frac{e}{q}\right)=0. \qquad (5.27)$$

Since the matrix of partial derivatives of (5.22) and (5.24) with respect to (l_b, l_g) is full-rank, (5.26) and (5.27) are necessary conditions for an efficient solution. So we can use the Lagrange multiplier associated with the level of aggregate profits which prevails in the decentralized equilibrium to evaluate the *laissez faire* allocation.

As in Shapiro and Stiglitz the decentralized equilibrium features two kinds of inefficiencies. First, employment in both states is too low, as the private cost of labour is above its social cost e. While the marginal private benefit $R'(l_i, \alpha_i) - \varphi_i = 0$, the marginal product of labour, evaluated at the decentralized allocation, is strictly higher that its social cost e and the first terms in equations (5.26) and (5.27) are strictly positive. Second, firms do not take into account that by increasing employment they alter wages at all firms by affecting the exit rate from unemployment: the thick market externality associated with the terms in $\partial h/\partial l_i$ in (5.26) and (5.27). But the endogeneity of the separation rate introduces a further inefficiency not present in the original model. As argued in section 5.2.3, employing one more worker in the bad state reduces the wage bill in the good state by pe/q - the last term in equation (5.26) - but downsizing firms do not take it into account since the benefit is sunk for them.

The externalities associated with the non-vanishing terms in the brackets in (5.26) and (5.27) look like pecuniary externalities as they are all reflected in wages. So, one would expect them to be internalized by the price mechanism in the decentralized equilibrium. As argued by Greenwald and Stiglitz (1986), though, pecuniary externalities may not disappear - i.e. may become real externalities - if markets are not complete (or other real distortions are present). This is a case in point and can be most easily seen in the case of the externality associated with the last term in equation (5.27). The externality is not reflected in the marginal employment cost of any firm. The reason, as argued in section 5.2.3, is market incompleteness. Separation is inefficient in the *laissez faire* equilibrium, as contracting firms do not take into account the externality, in the form of foregone rents, that layoffs impose on workers. As incentive compatibility requires workers to enjoy state-independent rents, the externality cannot be traded in a spot market. Even if redundant workers would volunteer to take a wage cut to avoid being laid off, *ex post* they would have an incentive to shirk. On the other hand, firms enjoying good business conditions would be willing to trade the externality *ex ante* and commit to fire less in downturns in exchange for lower current wages. The inefficiency would obviously disappear if there existed a market in which firms in the good state could express their demand for labour in the bad state. Yet, in the absence of reputational or commitment devices their pledge is not time-consistent. Moral hazard on the part of workers and the time-inconsistency of firing decisions prevent the externality from being traded either in spot or future markets.

The *laissez faire* levels of both l_b and l_g are lower than their first-best counterparts. As long as the total resource constraint (5.24) is slack, its shadow price is zero and the social planner would increase employment in both states until all profits have been taxed away[17], as can be seen from equation (5.26) and (5.27). So, a first-best intervention would consist of a hiring subsidy and a tax on redundancies financed through a lump-sum tax on firms' value.

On the other hand, the above discussion suggests that the decentralized equilibrium may not even be locally efficient. Severance payments do not address the inefficiency stemming from a too low employment level at hiring firms, but by reducing job destruction they are Pareto improving. In fact, the following proposition can be proved.

Proposition 1 *If $Q = F$, the change in the frictionless allocation induced by severance payments is Pareto improving for a size of the severance payment lower or equal to the privately optimal size e/q.*

Proof. See appendix 5.B.

Given that the decentralized equilibrium features different forms of inefficiencies the result is not obvious. As a Pigovian tax, dismissal costs make firms internalize the externality, in the form of foregone rents and higher wages, that they impose on dismissed workers and expanding firms. The privately optimal size of the dismissal cost determines the same allocation of labour which would prevail if the externality could be traded in either spot or future markets[18]. Yet, in a second-best world there is no reason to expect that tackling one source of inefficiency alone results in a welfare improvement. Intuitively, one can explain

[17]The first best allocation can only be an interior solution. The exit rate from unemployment and the incentive-compatible wages would be infinite at full employment .

[18]Failure of reputational mechanims and costs of writing private contracts may partly explain the existence of legislated dismissal regulations. The time inconsistency problem may explain how political pressures to scrap job security provisions may arise in response to changes in the balance between the number of firms in the good and bad state.

the result in the following way. A pure redundancy pay reduces workers' mobility cost in case of economic dismissal, hence resulting in lower wages in good states. The fall in wages fully offsets the partial equilibrium effect and the firing cost works as a tax on separation, but does not depress hiring on impact. By reducing the inflow into unemployment and the size of the unemployment pool on impact, redundancy pay depresses the exit rate from unemployment and wages thus reducing also the inefficiency associated with the thick market externality in job creation.

The model also predicts a negative relationship between the efficient level of severance payments as a proportion of wages and the detection probability q. The ratio between the optimal severance payment and the pre-displacement wage[19] is given by

$$\frac{e/q}{w_b} = (q + r + h)^{-1}. \tag{5.28}$$

The lower the probability of detection, the higher is the rent as a proportion of the wage. Workers' in professions which are more difficult to monitor should receive higher proportional severance payments.

The mechanism highlighted here is relevant not only for this particular model, but applies to all efficiency wage models and, in general, to all cases in which separation is inefficient because of endogenous real wage rigidity. Severance payments not only reduce the inefficiently high rate of job destruction but, by reducing workers' expected mobility cost, they increase the expected utility from continuing

[19] Since mobility for redundant workers is costless, if severance payments are efficient, wages are uniform across states.

employment with the firm until the firm carries out redundancies. As employment is *ex ante* more attractive, workers are willing to accept lower wages, or reduce turnover, or increase cooperation with the firm. So dismissal costs are likely to efficiently increase employment at contracting firms, but have limited negative impact effects on job creation.

If the potential benefits of redundancy payments apply to the general class of models in which separation is inefficient, one aspect which we have overlooked is specific to this model. Redundancy pay is unambiguously efficient only in the case of job losses associated with shocks which are orthogonal to individual effort. Any limit to the firm's ability to discipline non-compliant workers can only be efficiency-reducing in the present set up if it results in higher workers' rents, hence lower employment. In so far as third parties, such as labour courts and arbitration bodies, can only imperfectly ascertain whether disciplinary dismissals are warranted or not, any unfair dismissal compensation is likely to reduce somehow the expected cost from malfeasance, *coeteris paribus*, and result in higher labour costs[20]. As shown in MacLeod and Malcomson (1989), though, a firm could still retain its ability to punish shirkers by introducing a performance-related component into the wage package which exactly offsets the expected payment that a shirker may obtain on dismissal. An equal reduction in the fixed component of the wage package would leave marginal profits and workers' rents unchanged. In the extreme case in which disciplinary dismissal are impossible, the whole rent should

[20]The impact of measures which increase the cost of firing shirkers on workers' rents is reduced, though, in so far as the stigma associated with job loss is higher when disciplinary dismissals are costly.

accrue to the worker in the form of a performance-related bonus and the fixed component of the wage should be zero[21]. In the presence of constraints on the minimum explicitly contracted wage, though, unfair dismissal compensation may increase workers' rents and reduce employment at both hiring and firing firms, thus reducing efficiency.

Outside from the case of a pure severance payment, the effect of dismissal costs on aggregate welfare is ambiguous. Government intervention in the form of mandatory redundancy pay is likely to entail significant deadweight losses insofar as it is binding and introduces red-tape and enforcement costs. In this case $Q < F$ and the net outcome is the result of three different effects. First, the reduction in firing enhances welfare. Second, hiring falls as the net implicit turnover cost increases with $(F - Q)$. Despite that aggregate employment unambiguously increases, as we have argued in section 5.2, this reallocation of labour from high to low productivity firms is likely to reduce output at unchanged aggregate employment. Finally, as part of the cost born by firms is just a waste of resources, firing costs entail a deadweight loss given by the term $p(l_g - l_b)(F - Q)/2$ in equation (5.22). These two latter effects are obviously welfare reducing.

The net result is more likely to be negative the higher the variability of marginai productivity and employment across states; i.e. the higher is the fall in output at unchanged employment and the higher is the deadweight loss associated with turnover.

To illustrate the effect of firing costs on aggregate welfare we simulate the

[21]The rest of the workforce would be the third party which ensures that the firm does pay the bonus to the marginal worker.

model for different values of the deadweight loss. The production function is assumed to be Cobb-Douglas with labour share equal to 0.64, which is standard in the Real Business Cycle literature. The Cobb-Douglas specification allows to normalize e to 1. In fact, if shocks are multiplicative and severance payments are realistically modelled as a linear function of the wage, the system (5.20)-(5.21) is homogeneous of degree zero in e. The multiplicative productivity shocks are $\alpha_g = 2.7$ and $\alpha_b = 1.9$, which imply that revenues increase by 35 percent at unchanged employment as a consequence of a positive productivity shocks. We also assume $r = 0$, $q = 1$ and $p = 0.15$. On average, a shirker can expect not to be caught for one year and firms experience shocks of the size we consider every six and a half years. The chosen parameter values imply a significant fluctuation in employment between states. In the decentralized equilibrium, firm-level employment falls by fifty per cent in response to a negative shock: $l_g = 1.2$ and $l_b = 0.6$. The qualitative results are unchanged for any parameterization.

The results are reported in figure 5.1. On the horizontal axis is the size of the payment F born by the firm. Its maximum value of one corresponds to e/q, the value which ensures efficient separation.

The highest curve in figure 5.1 corresponds to the case of a pure severance payment ($F = Q$). The others correspond, in decreasing order, to the cases in which the unit deadweight loss equals respectively one fourth, one third, one half and the whole severance payment Q which accrues to the worker. Apart from the highest curve, all curves have a discontinuity at $F = e/q$. When the cost born by the firm equals workers' mobility cost, the parties can economize on third-party

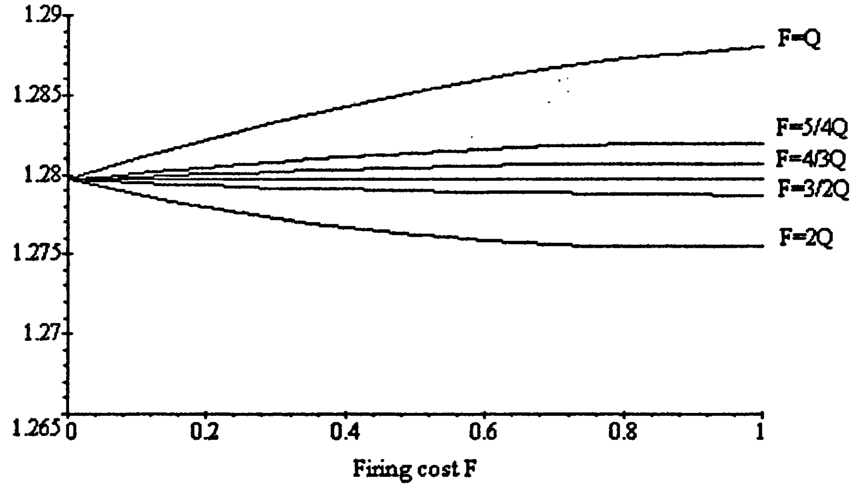

Figure 5-1: Effect of firing costs on welfare for different deadweight component

payments by labelling the separation a quit. If this is anticipated by the agents involved, aggregate welfare is the same as under a pure severance payment equal to e/q.

It is apparent that even a small deadweight component significantly reduces the efficiency enhancing effect of severance payments. Yet, even for the very high employment variability implied by our parameterization, firing costs reduce welfare below its *laissez faire* level only if they entail a rather sizeable unit deadweight loss in the range of half the severance payment Q or higher.

5.5 Conclusion

The simple model in this chapter highlights important dynamic implications of efficiency wage theories. If firms cannot commit on future firing, job destruction is inefficiently high in the decentralized equilibrium, as firms do not take into account the negative externality, in the form of foregone rents, that they impose on dismissed workers. The price mechanism cannot internalize the externality, due to both spot and future markets incompleteness, as moral hazard prevents workers from credibly offering to take a wage cut despite a positive rent from continuation and firms' *ex ante* pledges to restrain firing in downturns are not time-consistent.

Firing costs not only unambiguously increase employment, but, acting as a Pigovian tax on firing, they make firms internalize the externality. Thus, they can be Pareto improving, if they at least partly accrue to workers.

The intuition is relevant for all efficiency wage models and, in general, for all models that generate endogenous wage rigidity.

Appendix 5.A: Employment effect of dismissal costs

The equilibrium effect of dismissal costs on aggregate employment can be easily analysed graphically if one describes the equilibrium in terms of the unemployment rate u and the exit rate h. It is convenient to assume that the marginal product of labour has the linear form

$$R'(l_i; \alpha_i) = \alpha_i - \sigma l_i.$$

From (5.20) and (5.21) we can recover

$$l_b = \frac{\alpha_b - e\left(1 + \frac{r}{q}\right) - h\frac{e}{q} + (p + r) F}{\sigma} \tag{5.29}$$

and

$$l_g = \frac{\alpha_g - e\left(1 + \frac{r}{q}\right) - h\frac{e}{q} - p\left(\frac{e}{q} + F - Q\right)}{\sigma}. \tag{5.30}$$

Replacing in (5.11) and (5.12) results in the two equations

$$hu = \frac{p}{2\sigma}\left(\alpha_g - \alpha_b - p\left(\frac{e}{q} + F - Q\right) - (p + r) F\right) \tag{5.31}$$

and

$$u = 1 - \frac{1}{2\sigma}\left(\alpha_g + \alpha_b - 2e\left(1 + \frac{r}{q}\right) - 2h\frac{e}{q} - p\left(\frac{e}{q} - Q\right) + rF\right). \tag{5.32}$$

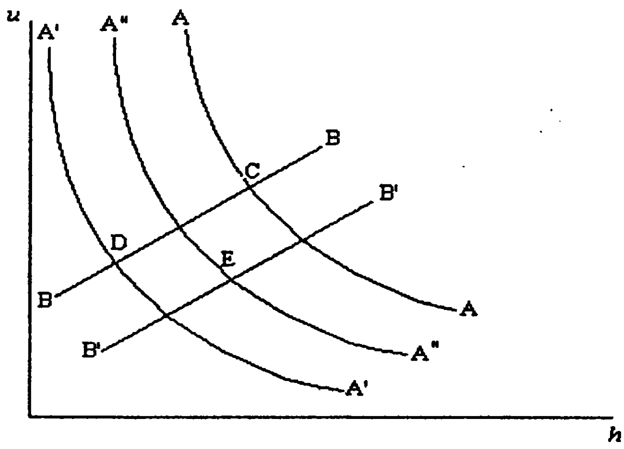

Figure 5-2: Effect of the firing cost F and the indemnity Q on equilibrium.

These correspond respectively to the curves AA and BB in figure 5.2. The AA curve is downward sloping as an increase in unemployment requires an offsetting fall in the exit rate to keep the outflow from unemployment constant. The BB curve slopes upward since an increase in h, by increasing the permanent income from unemployment, results in higher wages and lower employment[22].

An increase in F, the firing cost born by the firm, shifts the AA curve inward. As a higher F reduces the inflow into unemployment, either the unemployment or the exit rates or both have to fall. Viceversa an increase in Q, the indemnity which accrues to workers, shifts the curve out, as it increases employment at firms in the good state and so turnover.

Both Q and F shift the BB curve down as they increase employment at given

[22]The shape of the two curves and the qualitative results are unaffected if the marginal product of labour has a more general form provided its slope at equilibrium is lower in the good state than in the bad one. If this is not the case, the AA curve can have an upward sloping part for u low enough: at constant unemployment, an increase in the exit rate may have a bigger impact on the inflow into unemployment than on the outflow. Numerical simulations with a Cobb-Douglas production function show that if an equilibrium exists the intersection will always be to the left of this part.

wages. It has to be noted, though, that the positive effect of F on the BB curve is small and functional form dependent. As argued in Bertola (1990,1992), if the marginal product of labour is a linear function and if shocks are additive, firing costs have an (positive) effect on aggregate employment only if the discount rate is positive - the term rF in equation (5.32). Conversely, if the marginal product of labour curve is convex enough (e.g. Cobb-Douglas), firing costs can reduce aggregate employment but this effect is second order as it is a consequence of Jensen inequality. In order not to muddle the results with second order, specification dependent effects, we assume that the BB curve does not shift with F (this is the case, for example, if $r = 0$). Any second order shift of the BB curve in either direction does not imply any qualitative change in the following results.

An increase in the redundancy cost F, at constant Q, unambiguously reduces unemployment and the exit rate, moving the economy from C to D in figure 5.2. By reducing turnover, firing costs lower the exit rate from unemployment and wages in both states. So aggregate employment increases.

The inward shift in the AA curve is smaller in the case of a pure severance payment ($F = Q$), since the impact on turnover is smaller (the shift is to AA") On the other hand, the indemnity Q shifts BB down to BB', as it increases average employment. The resulting equilibrium is at point E in figure 5.2. Employment may be higher or lower than in the case in which $Q = 0$ (point D), but it unambiguously increases with respect to the frictionless equilibrium (point C). The effect on the exit rate from unemployment is ambiguous, though.

So firing costs, whether totally ($F = Q$) or partially ($F > Q$) received by workers, unambiguously increase aggregate employment.

Appendix 5.B: Proof of proposition 1

Proposition 1. *If $Q = F$, the change in the frictionless allocation induced by severance payments is Pareto improving for a size of the severance payment lower or equal to the privately optimal size e/q.*

Proof. The condition for local optimality can be rewritten by taking the ratio of equations (5.26) and (5.27).

$$\frac{R'(l_b;\alpha_b) - e}{R'(l_g;\alpha_g) - e} = \frac{A(l_b + l_g)\frac{e}{q} - \left(R'(l_b;\alpha_b) - \varphi_b + p\frac{e}{q}\right)}{B(l_b + l_g)\frac{e}{q} - \left(R'(l_g;\alpha_g) - \varphi_g\right)}, \tag{5.33}$$

where

$$A = \frac{\partial h}{\partial l_b} = \frac{2p\left(l_g - 1\right)}{\left(2 - l_g - l_b\right)^2} \tag{5.34}$$

and

$$B = \frac{\partial h}{\partial l_g} = \frac{2e\left(1 - l_b\right)}{\left(2 - l_g - l_b\right)^2}. \tag{5.35}$$

Equation (5.33) states that at the social optimum a social indifference curve must be tangent to an isoprofit line in the (l_b, l_g) space. The left-hand-side of (5.33) is the social marginal rate of substitution and the right-hand-side the slope of an isoprofit line. If the right-hand-side of (5.33) were greater than the left-hand-side, it would be possible to increase welfare by exchanging l_b for l_g at a rate that left profits unchanged. By evaluating the two sides of (5.33) at the privately chosen allocation in the presence of firing cost we can prove the result. We can replace for $\varphi_g = \varphi_b + pe/q$, $R'(l_b;\alpha_b)$, $R'(l_g;\alpha_g)$, φ_b using equations (5.20), (5.21), (5.17) and remembering that $Q = F$ and $r = 0$. We guess that the left-hand-side of (5.33) is not smaller than its right-hand-side; i.e.

$$\frac{w_b - e - pF}{w_b - e + p\frac{e}{q}} \geq \frac{A(l_b + l_g)\frac{e}{q} - p\left(\frac{e}{q} - F\right)}{B(l_b + l_g)\frac{e}{q}}. \tag{5.36}$$

Replacing for w_b, A and B and after a little of algebraic manipulation, the inequality (5.36) can be rewritten as

$$p\left(\frac{e}{q} - F\right) \geq -p\left(\frac{e}{q} - F\right)\left(\frac{2 - l_g - l_b}{l_g + l_b}\right) \tag{5.37}$$

which proves that our guess is right for $F \leq e/q$. So the exchange of l_b for l_g induced by firing costs is welfare improving for any value of F up to e/q. ■

Chapter 6

Conclusion

The first generation of models on dismissal costs concentrated on the welfare reducing effect of the introduction of an exogenous firing cost in a first best environment. Dismissal costs, it was argued, reduce profits and turnover and distort labour allocation. Also, since only wasteful dismissal costs, but not pure severance payments, affect separation rates, even the possible welfare gains from increased tenure and human capital investment are outweighed by the deadweight losses that firing costs impose.

This thesis has challenged this view. If utility is transferable and information symmetric, as all these models assume, the Coase theorem would predict that dismissal costs, whether they accrue to workers or not, can have at most distributional effects, but not real ones. In fact, we show that separation is always efficient in this class of models. If necessary, firms would induce workers to quit by offering them a voluntary severance payment, but they would never fire them. Dismissal costs may still reduce job creation, if they reduce *ex post* profits and entrance wages do not fall enough to leave the *ex ante* value of firms unchanged, but they cannot affect tenure. Yet, the empirical evidence points to a strong correlation between measures of job security and tenure.

For dismissal costs to have real effects when workers are homogeneous utility has to be non-transferable. We concentrate on one particular source of non-transferability: real wage rigidity. Real wage rigidity may arise both because of exogenous institutional constraint or because of right-to-manage wage setting or efficiency wages. In all these case, in the absence of dismissal costs separation rates are inefficiently high as firing firms do not take into account the externality, in the form of foregone rents, that they impose on workers on separation. Since utility is non-transferable efficiency depends on the distribution of property rights. Dismissal costs, whether received by workers or not, induce firms to internalize the externality and increase efficiency. We show in a simple model of exogenous mobility costs that a pure severance payment is unambiguously welfare improving in such a setup. What is crucial for firing costs to affect the separation rate is not their wastefulness, but the non-transferability of utility.

The positive relationship between tenure and job security implied by models of real wage rigidity is consistent with the empirical evidence. Our result also implies that human capital investment is likely to be inefficient in the absence of dismissal costs, as tenure is inefficiently short.

Possibly the most widely accepted rationale for the existence of real wage rigidities are efficiency wage considerations. Many industrial relations scholars argue that job security may significantly increase profits by improving workers' cooperation and openness to flexibility and reducing voluntary turnover. These are exactly some of the reasons that justify paying efficiency wages. We analyse the effect of dismissal costs in a dynamic version of Shapiro and Stiglitz (1984). We endogenize the separation rate by assuming that firms' productivity is subject to idiosyncratic uncertainty.

We show that severance payments unambiguously improve both the value of hiring firms and aggregate efficiency. Since mobility has to be costly for unemployment to be a credible punishment, redundant workers bear an involuntary mobility cost. This raises incentive compatible wages at expanding firms, as workers have to be compensated for the expected mobility cost. Firms control the firing probability and can reduce their wage bill by committing to fire less in downturns. Yet, at the time of firing, their pledge is not time-consistent as the benefit from lower wages is sunk. In the absence of commitment or reputational devices, the externality, in the form of foregone rents, that firms impose on workers cannot be traded. Workers' moral hazard prevents it to be traded *ex post,* while the firm's moral hazard rules out *ex ante* trading. Dismissal costs not only unambiguously increase aggregate employment, but allow firms to credibly commit to fire less in downturns. By acting as a Pigovian tax on firing, they induce firms to internalize the externality. Welfare may still increase even if only part of the cost born by the firm accrues to workers.

The intuition applies to all efficiency wage models and in general to all models that generate endogenous wage rigidity. Whenever mobility is involuntary costly, job security is a substitute for higher wages, so it either reduces the total wage bill, or it increases profits by reducing turnover costs or increasing workers' effort and cooperation. Firms should find it profitable to offer job security, but, in times of crisis, when the benefits are sunk employers may have an incentive to renege on it.

Wage rigidities are not the only market imperfection that generates a useful role for job security provisions. It is well known that, in the absence of contracts, search frictions imply that investment is inefficiently low due to hold up. We show that consensual layoff provisions increase firms' investment in general training in a model of frictional unemployment, through the following mechanism. Under employment-at-will in case of separation the firm does not capture any return to training, as the latter is vested in the worker. If the firm cannot unilaterally severe the relationship, it has to bargain over the total payoff from separation whenever job termination is efficient. This reduces the firm's *total* payoff from separation, but increases its *marginal* return to training as it now shares the return that accrues to the worker. The same intuition applies also to workers' investment in activities that increase firms' goodwill, such as effort to improve product quality or to develop products that remain the intellectual property of the firm. It also provides support for the often heard conjecture that job security should increase workers' investment in human capital that is vested in the firm.

It has to be noted that, though the models in this thesis predict a welfare-improving role for job security provisions, they imply that the efficient outcome can be achieved by credible, private *ex ante* contracting on employment protection, be it the size of severance payments or consensual layoffs. They do not provide any rationale for government-imposed discipline. One possible rationale for mandated job security is the existence of economies of scales in reaching credible agreements. Implicit contracts may be unenforceable as reputation is likely to be worthless when layoffs are associated with takeovers, liquidation or relocation to other countries. Individual workers may find it too costly to verify whether written contracts offered by firms would be upheld in courts. These costs may exceed the individual gains associated with job security, but could become negligible when shared across a large enough number of workers. If these economies of scales are significant then there can be scope for the government to provide a standardized, fill-in-the-details contract. Alternatively, the observed legislated discipline may achieve a similar effect in so far as mandatory measures can be negotiated away whenever it is efficient to do so, as shown in chapter 3.

Bibliography

Abraham, Katharine G. and Houseman, Susan N. (1993), *Job Security in America: Lessons from Germany*, The Brookings Institution, Washington D.C.

Acemoglu, Daron (1996), 'A Microfoundation for Social Increasing Returns in Human Capital Accumulation', *Quarterly Journal of Economics* **111**, 779–804.

Acemoglu, Daron (1997), 'Training and Innovation in an Imperfect Labour Market', *Review of Economic Studies* **64**, 445–464.

Acemoglu, Daron and Pischke, Jörn-Steffen (1999), 'The Structure of Wages and Investment in General Training', *Journal of Political Economy* **107**, 539–572.

Akerlof, George (1984), 'Gift Exchange and Efficiency Wage: Four Views', *American Economic Review* **74**, 79–83.

Azariadis, Costas (1975), 'Implicit Contracts and Underemployment Equilibria', *Journal of Political Economy* **83**, 1183–1202.

Baily, Martin (1974), 'Wages and Employment under Uncertain Demand', *Review of Economic Studies* **41**, 37–50.

Barron, John M. , Berger, Mark C. and Black, Dan A. (1997), *On-the-Job Training*, W. E. Upjohn Institute for Employment Research, Kalamazoo.

Bean, Charles R. (1994), 'European Unemployment: A Survey', *Journal of Economic Literature* **32**, 573–619.

Bean, Charles R. (1997), 'Comment to: Exploring the Political Economy of Labour Market Institutions by Gilles Saint-Paul', *Economic Policy* **23**, 265–292.

Becker, Gary (1964), *Human Capital*, The University of Chicago Press, Chicago.

Bentolila, Samuel and Bertola, Giuseppe (1990), 'Firing Costs and Labor Demand: How Bad is Eurosclerosis?', *Review of Economic Studies* **57**, 381–402.

Bentolila, Samuel and Saint-Paul, Gilles (1994), 'A Model of Labor Demand with Linear Adjustment Costs', *Labour Economics* **1**, 303–326.

Bertola, Giuseppe (1990), 'Job Security, Employment and Wages', *European Economic Review* **34**, 851–886.

Bertola, Giuseppe (1992), 'Labor Turnover Costs and Average Labor Demand', *Journal of Labor Economics* **10**, 389–411.

Bertola, Giuseppe (1994), 'Flexibility, Investment and Growth', *Journal of Monetary Economics* **34**, 215–238.

Bertola, Giuseppe (1996*a*), Uninsurable Risk, Efficiency and Labor Market Rigidities, mimeo Università di Torino.

Bertola, Giuseppe (1996*b*), Wage Differentials and Labor Mobility with Uninsurable Risk, mimeo Università di Torino.

Bertola, Giuseppe and Caballero, Ricardo (1994), 'Cross-Sectional Efficiency and Labor Hoarding in a Matching Model of Unemployment', *Review of Economic Studies* **61**, 435–456.

Bertola, Giuseppe and Ichino, Andrea (1995*a*), 'Crossing the River: A Comparative Perspective on Italian Employment Dynamics', *Economic Policy: A European Forum* **21**, 359–415.

Bertola, Giuseppe and Ichino, Andrea (1995*b*), 'Wage Inequality and Unemployment: U.S. versus Europe', *NBER Macroeconomics Annual*.

Binmore, Ken G. (1985), Bargaining and Coalitions, *in* Alvin E. Roth, ed., 'Game-Theoretic Models of Bargaining', Cambridge University Press, Cambridge: UK.

Binmore, Ken G. (1987), Perfect Equilibria in Bargaining Models, *in* Ken G. Binmore and Partha Dasgupta, eds, 'The Economics of Bargaining', Blackwell, Oxford: UK.

Binmore, Ken G. , Rubinstein, Ariel and Wolinsky, Asher (1986), 'The Nash Bargaining Solution in Economic Modelling', *Rand Journal of Economics* **17**, 176–188.

Binmore, Ken G. , Shaked, Ariel and Sutton, John (1989), 'An Outside Option Experiment', *Quarterly Journal of Economics* **104**, 753–770.

Bishop, John (1991), On-the-Job Training of New Hires, *in* David Stem and Josef Ritzen, eds, 'Market Failure in Training', Springer Verlag, New York, pp. 61–96.

Bishop, John H. (1996), What We Know about Employer-provided Training: A Review of the Literature, Center for Advanced Human Resources Studies Working Paper 96-09, Cornell University.

Blanchard, Olivier J. and Portugal, Pedro (2000), What Hides Behind an Unemployment Rate: Comparing Portuguese and U.S. Labor Markets, April 2000, Department of Economics, MIT, Accessed 20 September 2000, < http://econ-wp.mit.edu/RePEc/2000/blanchar/paper31.pdf>.

Blanchard, Olivier J. et al. (1986), Employment and Growth in Europe: A Two Handed Approach, *in* Olivier J. Blanchard et al., ed., 'Restoring Europe's Prosperity', The MIT Press, Cambridge: Massachussets.

Blinder, Alan S. and Choi, Don H. (1990), 'A Shred of Evidence on Theories of Wage Stickiness', *Quarterly Journal of Economics* **105**, 1003–1015.

Booth, Alison L. (1987), 'Extra-Statutory Redundancy Payments in Britain', *British Journal of Industrial Relations* **25**, 401–419.

Booth, Alison L. (1997), An Analysis of Firing Costs and Their Implications for Unemployment Policy, *in* Dennis Snower and Guillermo de la Dehesa, eds, 'Unemployment Policy: Government Options for the Labour Market', Cambridge University Press, Cambridge: UK.

Buechtemann, Christopher F. (1992), Employment Security and Labor Markets: An Introduction, *in* Christopher F. Buechtemann, ed., 'Employment Security and Labor Market Behavior: Interdisciplinary Approaches and International Evidence', ILR Press, Ithaca Cornell University.

Burda, Michael C. (1992), 'A Note on Firing Costs and Severance Benefits in Equilibrium Unemployment', *Scandinavian Journal of Economics* **94**, 479–489.

Cabrales, Antonio and Hopenhayn, Hugo A. (1996), Job Dynamics, Correlated Shocks and Wage Profiles, mimeo Universitat Pompeu-Fabra.

Carmichael, Lorne H. (1983), 'Firm-Specific Human Capital and Promotion Ladders', *Bell Journal of Economics* **14**, 251–258.

Chang, Chun and Wang, Yijiang (1996), 'Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revesited', *Journal of Labor Economics* **14**, 505–519.

Che, Yeon-Koo and Chung, Tai-Yeong (1996), Incomplete Contracts and Cooperative Investments, mimeo, Department of Economics, University of Wisconsin-Madison.

Che, Yeon-Koo and Hausch, Donald B. (1999), 'Cooperative Investment and the Value of Contracting', *American Economic Review* **89**, 125–147.

Cramton, Peter C. and Tracy, Joseph S. (1992), 'Strikes and Hold Outs in Wage Bargaining: Theory and Data', *American Economic Review* **82**, 100–121.

Davis, Steven J. and Haltiwanger, John (1992), 'Gross Job Creation, Gross Job Destruction and Employment Reallocation', *Quarterly Journal of Economics* **107**, 819–863.

Feldstein, Martin (1976), 'Temporary Layoffs in the Theory of Unemployment', *Journal of Political Economy* **84**, 937–957.

Felli, Leonardo and Harris, Christopher (1996), 'Learning, Wage Dynamics, and Firm-Specific Human Capital', *Journal of Political Economy* **104**, 838–868.

Garibaldi, Pietro and Violante, Giovanni L. (2000), Severance Payments in Search Economies with Limited Bonding, mimeo.

Greenwald, Bruce C. and Stiglitz, Joseph E. (1986), 'Externalities in Economies with Imperfect Information and Incomplete Markets', *Quarterly Journal of Economics* **101**, 229–264.

Greenwald, Bruce C. and Stiglitz, Joseph E. (1988), 'Pareto Inefficiency of Market Economies: Search and Efficiency Wage Models', *American Economic Review: Papers and Proceedings* **78**, 351–355.

Grout, Paul A. (1984), 'Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach', *Econometrica* **52**, 449–460.

Grubb, David and Wells, William (1993), 'Employment Regulation and Patterns of Work in E.C. Countries', *OECD Economic Studies* **21**, 7–58.

Hamermesh, Daniel S. (1989), 'Labor Demand and the Structure of Adjustment Costs', *American Economic Review* **79**, 674–689.

Hansen, Gary (1985), 'Indivisible Labor and the Business Cycle', *Journal of Monetary Economics* **16**, 309–328.

Harhoff, Dietmar and Kane, Thomas (1993), Financing Apprenticeship Training: Evidence from Germany, NBER Working Paper No. 4557, Cambridge, Massachusetts.

Hart, Oliver and Moore, John (1988), 'Incomplete Contracts and Renegotiation', *Econometrica* **56**, 755–86.

Holtz-Eakin, Douglas and Rosen, Harvey S. (1991), 'Municipal Labor Demand in the Presence of Uncertainty: An Econometric Approach', *Journal of Labor Economics* **9**, 276–293.

Hopenhayn, Hugo and Rogerson, Richard (1993), 'Job Turnover and Policy Evaluation: A General Equilibrium Analysis', *Journal of Political Economy* **101**, 915–938.

Hosios, Arthur J. (1990), 'On the Efficiency of Matching and Related Models of Search and Unemployment', *Review of Economic Studies* **57**, 279–298.

Katz, Lawrence (1986), 'Efficiency Wage Theories: A Partial Evaluation', *NBER Macroeconomics Annual.*

Kaufman, Roger T. (1984), 'On Wage Stickiness in Britain's Competitive Sector', *British Journal of Industrial Relations* **22**, 101–112.

Krafcik, John F. (1990), Training and the Automobile Industry, Contractor Report prepared for the Office of Technology Assessment under contract N3/1910.

Kreps, David M. (1990), *Game Theory and Economic Modelling*, Oxford University Press, Oxford.

Krueger, Alan B. and Summers, Lawrence H. (1988), 'Efficiency Wages and the Interindustry Wage Structure', *Econometrica* **56**, 259–293.

Lazear, Edward P. (1990), 'Job Security Provisions and Employment', *Quarterly Journal of Economics* **105**, 699–726.

Leontief, Wassily (1946), 'The Pure Theory of the Guaranteed Annual Wage Contract', *Journal of Political Economy* **54**, 76–79.

Lindbeck, Assar and Snower, Dennis (1988), *The Insider-Outsider Theory of Employment and Unemployment*, The MIT Press, Cambridge: Massachussetts.

Loewenstein, Mark A. and Spletzer, James R. (1999), 'General and Specific Training: Evidence and Implications', *The Journal of Human Resources* **34**, 710–733.

Lucas, Jr., Robert E. (1987), *Models of Business Cycles*, Basil Blackwell, Oxford.

Lucas, Robert E. Jr. and Prescott (1974), 'Equilibrium Search and Unemployment', *Journal of Economic Theory* **7**, 188–209.

MacLeod, W. Bentley and Malcomson, James M. (1989), 'Implicit Contracts, Incentive Compatibility and Involuntary Unemployment', *Econometrica* **57**, 447–480.

MacLeod, W. Bentley and Malcomson, James M. (1993*a*), 'Investments, Holdup and the Form of Market Contracts', *American Economic Review* **83**, 811–837.

MacLeod, W. Bentley and Malcomson, James M. (1993*b*), 'Specific Investment and Wage Profiles in Labour Markets', *European Economic Review* **37**, 343–354.

MacLeod, W. Bentley and Malcomson, James M. (1993*c*), 'Wage Premiums and Profit Maximization in Efficiency Wage Models', *European Economic Review* **37**, 1223–1249.

MacLeod, W. Bentley and Malcomson, James M. (1995), 'Contract Bargaining with Symmetric information', *Canadian Journal of Economics* **28**, 336–367.

MacLeod, W. Bentley and Malcomson, James M. (1998), 'Motivation and Markets', *American Economic Review* **88**, 388–411.

MacLeod, W. Bentley , Malcomson, James and Gomme, Paul (1994), 'Labor Turnover and the Natural Rate of Unemployment: Efficiency Wage versus Frictional Unemployment', *Journal of Labor Economics* **12**, 276–315.

Manzini, Paola and Snower, Dennis J. (1996), On the Foundations of Wage Bargaining, CEPR discussion paper n. 1514.

Masters, Adrian M. (1998), 'Efficiency of Investment in Human and Physical Capital in a Model of Bilateral Search and Bargaining', *International Economic Review*.

McDonald, Ian M. and Solow, Robert M. (1981), 'Wage Bargaining and Employment', *American Economic Review* **71**, 896–908.

Merz, Monika (1996), A Market Structure for and Environment with Heterogeneous Job Matches, Indivisible Labor and Persistent Unemployment, mimeo Rice University.

Millard, Stephen P. and Mortensen, Dale T. (1997), The Unemployment and Welfare Effects of Labour Market Policy: A Comparison of the US and the UK, *in* Dennis Snower and Guillermo de la Dehesa, eds, 'Unemployment Policy: Government Options for the Labour Market', Cambridge University Press, Cambridge: UK.

Millward, Neil et al. (1992), *Workplace Industrial Relations in Transition*, Aldershot, Dartmouth.

Moene, Karl O. (1988), 'Unions' Threats and Wage Determination', *Economic Journal* **98**, 471–483.

Mortensen, Dale T. and Pissarides, Christopher A. (1994), 'Job Creation and Job Destruction in the Theory of Unemployment', *Review of Economic Studies* **61**, 397–415.

Nickell, Stephen (1997), 'Unemployment and Labor Market Rigidities: Europe versus North America', *Journal of Economic Perspectives* **11**, 55–74.

Nickell, Stephen J. (1998), Job Tenure and Labour Reallocation: a Partial Overview, Paper presented at the conference on Job Tenure and Labour Reallocation, 24-25 April 1998, CEP, London School of Economics.

Nickell, Stephen J. and Layard, Richard (1997), Labour Market Institutions and Economic Performance, *in* Oliver Ashenfelter and David Card, eds, 'Handbook of Labour Economics, Vols. 3, 4', North Holland, Amsterdam. Forthcoming.

Osborne, Martin J. and Rubinstein, Ariel (1982), *Bargaining and Markets*, Academic Press, San Diego: California.

Osterman, Paul and Kochan, Thomas A. (1990), Employment Security and Employment Policy: An Assessment of the Issues, *in* Katharine Abraham and Robert McKersie, eds, 'New Developments in the Labor Market: Towards a New Institutional Paradigm', The MIT Press, Cambridge: Massachussets.

Parent, Daniel (1999), 'Wages and Mobility: The Impact of Employer-Provided Training', *Journal of Labor Economics* **17**, 298–317.

Piore, Michael J. (1986), 'Perspectives on Labor Market Flexibility', *Industrial Relations* **25**, 146–166.

Raff, Daniel M. G. and Summers, Lawrence H. (1987), 'Did Henry Ford Pay Efficiency Wages?', *Journal of Labor Economics* **5**, S57–S86.

Roberts, Mark J. (1996), Employment Flows and Product Turnover in Three Developing Countries, *in* Mark J. Roberts and James R. Tybout, eds, 'Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure', Oxford University Press, Oxford.

Rubinstein, Ariel (1982), 'Perfect Equilibrium in a Bargaining Model', *Econometrica* **50**, 97–109.

Saint-Paul, Gilles (1995*a*), 'Efficiency Wage, Commitment and Hysteresis', *Annales d'Economie et de Statistique* **37-38**, 39–53.

Saint-Paul, Gilles (1995*b*), 'The High Unemployment Trap', *Quarterly Journal of Economics* **110**, 527–550.

Saint-Paul, Gilles (1997), *Dual Labor Markets: A Macroeconomic Perspective*, MIT Press, Cambridge: Massachussets.

Salop, Joanne and Salop, Steven C. (1976), 'Self-Selection and Turnover in the Labor Market', *Quarterly Journal of Economics* **90**, 619–627.

Shaked, Avner (1987), Opting out: Bazaars versus High Tech Markets, STICERD Theoretical Economics Discussion Paper 159, London School of Economics.

Shaked, Avner and Sutton, John (1984), 'Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model', *Econometrica* **52**, 1351–1364.

Shapiro, Carl and Stiglitz, Joseph E. (1984), 'Equilibrium Unemployment as a Worker Discipline Device', *American Economic Review* **74**, 433–444.

Shimer, Robert and Smith, Lones (1999), Matching, Search and Heterogeneity, Mimeo, Princeton.

Siebert, Horst (1997), 'Labor Market Rigidities: At the Root of Unemployment in Europe', *Journal of Economic Perspectives* **11**, 37–54.

Ståhl, I. (1972), *Bargaining Theory*, Economics Research Institute Stockholm School of Economics, Stockholm.

Stiglitz, Joseph (1986), Theories of Wage Rigidity, *in* J. L. Butkiewicz et al., ed., 'Keynes' Economic Legacy: Contemporary Economic Theories', Praeger, New York.

Sutton, John (1986), 'Noncooperative Bargaining Theory - An Introduction', *Review of Economic Studies* **53**, 709–724.

Topel, Robert (1990), Specific Capital and Unemployment: Measuring the Costs and Consequences of Job Loss, *in* Olivier J. Blanchard et al., ed., 'Carnegie-Rochester Conference Series on Public Policy', North-Holland, Amsterdam.

Topel, Robert (1991), 'Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority', *Journal of Political Economy* **99**, 145–176.

Williamson, Oliver (1985), *The Economic Institution of Capitalism*, Free Press, New York.