Firm Training and Wage Rigidity[*]

Wolfgang Lechthaler - IFF-HSG[†]

November 29, 2005

Abstract

Based on overwhelming evidence suggesting that wages are in practice downwards rigid, we impose the restriction that wages of the current period are not allowed to be smaller than in the preceding period, whereas otherwise wages are negotiated freely via standard Nash-bargaining.

Thereby we are able to analyze the consequences of downwards wage rigidity on the determination of wages, on the degree of wage-compression and on firm's training investments. It turns out, that wage rigidity will increase wage-compression. However, this is not sufficient to increase firms' training investments. The reason lies in the endogeneity of separations, which become more frequent.

JEL-Classification: J24, J31, M53

Keywords: Human Capital, Wage Rigidity, Training

1 Introduction

Wage-rigidity is receiving more and more attention, especially in the RBC- and the empirical literature. It can be seen as a powerful device to bring in line the variability of vacancies and unemployment in RBC-models with the empirical facts. Models without rigidity usually produce unemployment-fluctuations which are way too low, because wages are adjusting too fast to macroeconomic shocks. Concerning the empirical literature, evidence on wage-rigidity comes from three different brands: Econometrics, surveys and experiments. Broadly speaking, all of them agree that there is a considerable amount of wage-rigidity, especially when it comes to downwards movements.

This work is built on the idea that wage-rigidity will have important effects on the wage-structure of an economy, potentially creating or increasing wage-compression.[1] According to the influential work of Acemoglu[2] wage-compression will induce firms to invest in the human capital of their workers because they can reap some of the returns to training. This is in contrast to the traditional theory on human capital based on Becker (1962). Becker argued that a firm would never pay for a worker's general human capital[3] because this kind of training would increase wages one-to-one. On the other hand, workers would invest efficiently in their human capital because they were the only beneficiaries. This is another difference to the models of Acemoglu where training-investments are inefficiently low due to various externalities which affect future employers and workers.

As will be discussed in the following section, the evidence on firm training and wage compression is in rare unison. As predicted by Acemoglu and Pischke (1999a), firms even pay for general training and wages are indeed compressed. But there is one subject in which the evidence is not so clear: Minimum wages. As discussed in Acemoglu and Pischke (2003) and described in more detail in the section on empirical literature below,

---

[1]A wage-structure is compressed if wages react less to changes in human capital than productivity does.

[2]See Acemoglu (1997) and Acemoglu and Pischke (1998, 1999a, 1999b).

[3]General human capital - as opposed to specific human capital - can be used without any restrictions in other firms.

the evidence pointing towards a positive relationship between minimum wages and firm training[4] is rather poor. This has lead Acemoglu and Pischke (2003) to adjust their models to bring back in line theory and evidence. In this paper I try to give an alternative explanation for a possibly negative relationship between inflexible wages and firm training.

In most aspects, my model is very similar to Acemoglu and Pischke (1999a), for instance there are two periods, only firms can invest in training and the first period is the training period. However, there are two important differences: First, the way in which wages are determined and second the endogeneity of separations. I stick to the usual assumption of Nash-bargaining but add the restriction of nominal wage-rigidity. Thus, I am able to show that wage-rigidity can indeed lead to a higher degree of wage-compression by altering the way wages are determined. Nevertheless, contrary to the models of Acemoglu and Pischke this is not sufficient to improve firm-training. This result is due to the effect of rigidity on separations. As confirmed by the empirical evidence these become more frequent as rigidity is introduced. In Acemoglu and Pischke (1999a) separations take place at an exogenous rate and therefore this effect is ruled out by assumption. Acemoglu and Pischke (1998) have endogenous separations. However, with this model is not possible to analyze minimum wages since all workers have the same wage.

The remainder of the paper is organized as follows. In the proceeding section the empirical literature on both, wage-rigidity and firm training will be discussed. Then I illustrate my own wage mechanism and compare it to other rules used in the theoretical literature. In section four I will present the general framework of the model before I outline a benchmark model in which wages are determined without any restriction. Section six discusses the model with rigid wages while section seven concludes.

---

[4]As suggested by the their models because minimum wages increase wage compression.

2 Empirical Literature

In this section I describe the empirical literature in more detail. As mentioned above there are three different branches dealing with the phenomenon of wage rigidity. I will start with the econometric evidence and then proceed with surveys and experiments. The chapter concludes with a discussion of econometric studies on firm training.

Baker, Gibs and Holmstrom (1994) use records of all management employees of a medium-sized US firm in a service industry over the period of 1969 to 1988. Excluding foreign employees[5] they arrive at a sample of 62.957 observations and try to determine the firms' wage policy. They report that nominal wage-cuts are extremely rare: Overall they observe less than 200 cases which is less than 0.32 % of all observations. On the other hand, zero nominal changes can be quite frequent, reaching up to more than 15% of observations in the year 1977. Real wage cuts are much more frequent than nominal wage cuts. For instance, 15% of workers of the 1975 cohort are receiving a lower real wage in 1985 than their starting wage. Not very surprisingly, real wage cuts are especially frequent in years of high inflation: In the high-inflation years around 1980, 40% of workers had to suffer declines in their real wages.

Fehr and Götte (2000) do a similar exercise for two Swiss firms (large and medium-sized) in the service industry.[6] In the large firm only 1.7 % of 35.779 observations are wage cuts. These are even less frequent in the medium sized firm: 0.4 % of 20.236 observations. Fehr and Götte do not only use firm-data but as well cross-sectional data from the Swiss Labor Force Survey and Social Insurance Files. Controlling for measurement error, they find that at most 5 % of workers received wage cuts, while for more than 50% of workers nominal rigidity prevented wage cuts. This wage rigidity does not vanish during periods of low GDP-growth.
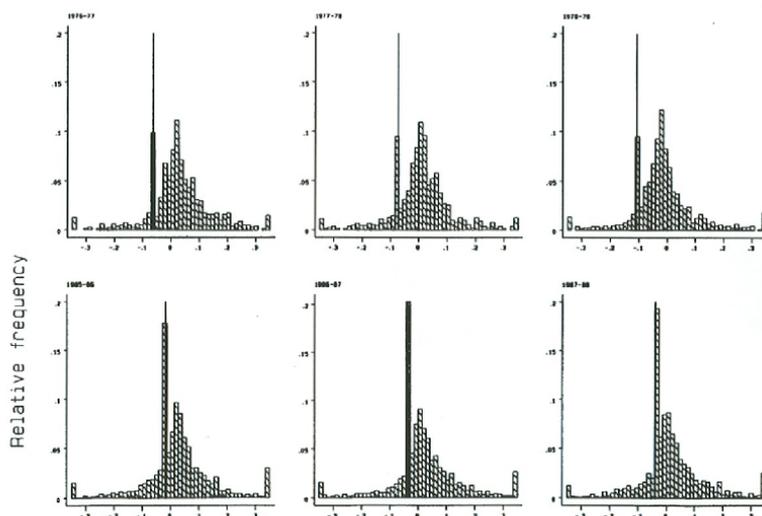
Card and Hyslop (1997) use the US Panel Study of Income Dynamics. Figure (1) - which is taken from their paper - illustrates the results for the years of relatively high

---

[5]Due to limited comparability.

[6]See Beissinger and Knoppik (2001) for a study on Germany.

Figure 1:   Log real wage changes

inflation 1976-1979 and the years of low inflation 1985-1988. Again there is sharp peak at zero nominal wage changes and nominal wage cuts are quite rare while real wage cuts are more frequent. The comparison of the periods of high and low inflation makes clear that nominal-wage rigidity is especially important during times of low inflation: In these years zero nominal wage changes are even more frequent.

Evidence of cross-sectional studies is not so clear. Parkin (2001) surveys ten panel studies and finds that between ten and twenty percent of wage changes are negative. However, reporting errors seem to be quite important in these studies, leading to over-statements of the frequency of wage cuts. A more detailed review of the econometric literature on wage rigidities can be found in Malcomson (1999) or Howitt (2002).

A different kind of evidence comes from Bewley (1999 and 2002) who asked US managers directly why they behave the way they do. He found an unusual deal of consensus that the most important factor inhibiting wage cuts is the fear that this might be interpreted as a hostile act and lead to a lower morale in the workforce, thus decreasing effort and productivity. For the same reason, firms do not replace workers by unemployed who

would be willing to work for less. On the other hand, managers are less reluctant to fire workers during a recession to improve productivity and profits. Although this will clearly lower the morale of the fired worker, she is no longer in the firm to spread the bad morale to other workers. A similar study has been done by Agell and Bennmarker (2003) for Sweden. They also find that the morale of the work force is an important obstacle to nominal wage cuts. Besides worker morale, the legal framework and institutions (unions) are important factors.[7]

Finally, a third piece of evidence on wage rigidity comes from experimental studies on reciprocal behavior. For instance, Fehr and Falk (1999) find that firms are not willing to hire underbidders and that workers who accept lower wages also provide lower effort if effort cannot be contracted. This clearly confirms the views of managers as reported by Bewley. However, as soon as effort can be contracted, underbidders are no longer refused.[8]

I next turn to the relationship between firm training and wages that has been the subject of numerous studies. For the US, Loewenstein and Spletzer (1998 and 1999) find that firms frequently pay for general training that general training usually increases productivity by more than wage incomes and that training is translated into higher wages if it was provided by a previous and not the current employer. This clearly suggests that wages are compressed and that training is general to a large degree - otherwise future employers would not pay higher wages. At the same time it can be neglected that workers pay indirectly for the training by receiving lower starting wages. Similar results are derived by Barron et al. (1999) for the US, by Booth and Bryan (2002) for the UK and by Gerfin (2003a and 2003b) for Switzerland.

A more direct test of the model in Acemoglu and Pischke (1999a) is provided by Bassanini and Brunello (2003). Using the European Community Household Panel (ECHP) they find a positive relationship between wage-compression and training.

---

[7]See as well Agell and Lundborg (1995).

[8]For other experiments on reciprocal behavior see for instance Camerer and Thaler (1995) or Falk, Gächter and Kovacs (1998).

To my knowledge there are no empirical studies directly relating the incidence of wage-rigidity and firm training. However, when it comes to minimum wages[9] the evidence is not so clear. Acemoglu and Pischke (2003) find no evidence that minimum wages reduce training. There is also little evidence that minimum wages increase training. Therefore, they create a new model in which workers are willing to accept wage cuts in order to finance part of their training. If this is prevented by the minimum wage, training might decrease. Thus, we have two countervailing effects. On the one hand the compression of the wage structure is increased, on the other hand workers cannot finance training by accepting lower wages. My own work might shed some light on this discussion. I show that results change considerably if we allow for endogenous separations.

3 Wage-setting

Since the wage-setting mechanism is the main feature of my model, this section discusses it in detail and compares it to other approaches in the literature. I assume that wages are determined according to Nash-bargaining[10] but with the restriction that wages cannot decline from one period to the other, so that the wage of the prior period is the minimum for the current period. Besides that, the wage of the prior period has no influence on the outcome of the bargaining and can be regarded as an outside option. I argue that - in line with the empirical literature discussed above - a decrease in the wage will be interpreted by the worker as a hostile act leading to the provision of zero effort. The firm foreseeing this "option" will never pay a lower wage than the one it has paid in the previous period.

This intuitive motivation is supported by Strand (2003) and Holden (2002). Holden discusses different ways of modelling wage rigidity that were used in the literature and compares them to empirical studies like the one by Fehr and Götte (2000) just described. Holden argues that so far no wage rule could explain the numerical results satisfactorily and suggests two alternatives that would fit better. In one of these alternatives he

---

[9]Minimum wages work similar as wage-rigidities because both prevent the wage from decreasing.

[10]See for instance Shaked and Sutton (1984).

combines a model with outside options with the adverse effects of bad morale: Wage cuts reduce efforts and this is modelled by a fixed cost that reduces output. Under the assumption that this extra cost is so large that production is no longer profitable, this results in exactly the same wage rule that I use.

Strand (2003) develops a synthesis between wage-bargaining à la Nash and efficiency wages. The firm sets the required effort in order to maximize its profits. Then wages are negotiated according to Nash-bargaining. However, if this were to lead to a wage, which is too low to assure that workers do not shirk, the firm will pay at least the non-shirking wage. In that way, the efficiency wage acts as a minimum wage in the Nash-bargaining process. In my model the wage of the last period plays the role of this efficiency minimum-wage. Whenever the worker receives less than in the past, the firm will expect her to shirk. Therefore, the firm will never lower the wage. In order to keep the model simple[11] I omit this micro-foundation and simply assume that the previous wage acts as a minimum.

Especially in the RBC-literature we find many other approaches to model wage rigidity. In this literature rigidity is used to amplify fluctuations of unemployment and vacancies over the business cycle because flexible wages usually adjust too quickly to shocks. Consequently, the variability of unemployment and vacancies is too low to match the empirical facts. The simplest method is used by Shimer (2004). who simply assumes that the wage is constant. By using this extreme rule he can demonstrate that a rigid wage can amplify the fluctuations of unemployment and vacancies sufficiently to better match the empirical facts. Hall (2003) also sets the wage constant but allows renegotiations in case a threat point of the two parties is violated, thus avoiding inefficient separations. Krause and Lubik (2003) stick to Nash-bargaining but with the modification that the actual wage is a weighted average of the Nash-wage and a wage norm. Another frequently used approach is the wage-staggering introduced by Taylor (1979), according to which wages are fixed for some periods until they can be renegotiated.

Danthine and Kurmann (2004a) try to incorporate rigid wages in an efficiency-wage

---

[11]Actually the focus is on firm training and not on wage bargaining.

model. However, this approach is rather ad hoc since it is simply assumed that the previous wage reduces the effort of the worker whereas the current wage improves it. Consequently, firms are more reluctant to lower wages because this will reduce effort. More elaborate appears the approach in Danthine and Kurmann (2004b). The authors stress that wage rigidity can be constructed in efficiency-wage models quite easily by moving the wage reference from external to internal. Usually the wage reference is assumed to be external[12] and wages will respond quickly to macroeconomic shocks. In contrast, Danthine and Kurmann suggest the firm's earnings per worker as wage reference. They demonstrate that this is sufficient to create a considerable amount of wage rigidity.

One common drawback of all the approaches discussed above is that they are creating rigidity in both directions: upwards and downwards. This is clearly at odds with the empirical literature showing that rigidity is mainly restricting downwards movements. It is an advantage of the approach used in this paper that it only prevents downwards adjustments of wages without affecting upwards movements. Additionally, it is able to create the large mass of zero-changes observed by the econometric literature. Another drawback of the efficiency wage literature is that it partly contradicts the evidence found by Bewley (2002). He argues that monetary incentives are only important when it comes to wage-cuts. If the management decides to increase the wage of its employees, they will improve their effort only temporarily. After some time they get used to the higher level of wages and perceive that they have earned it. Consequently, their effort goes back to normal. To the contrary, wage cuts have permanent effects on the morale of the workforce and therefore they are avoided. My model is clearly better able to take account of that fact.

Another motivation for sticking to Nash-bargaining is that it is most commonly used in the literature on firm training and specifically in the most influential work by Acemoglu and Pischke (1999a and also Acemoglu (1997)). In doing so I am able to allow direct comparisons between my model and the related literature.

---

[12]For instance, related to average earnings of the work force or unemployment benefits.

I believe that Nash-bargaining is predominant in the training literature not without good reason. It seems quite arbitrary to model training in an efficiency-wage model. How should this training affect the effort of the worker? If one believes that the worker will interpret training as sign of kindness and react by providing more effort, this might lead to wages declining with human-capital, which is clearly at odds with empirical findings. To my knowledge the only work directly incorporating firm training and efficiency wages is Katsimi (2003). In this model, training ties the worker and the firm more closely together. The worker knows that the firm will be more reluctant to fire her in case of recessions. This gives more credibility to the threat of firing as punishment to shirking. Consequently, a lower wage is sufficient to fulfill the no-shirking condition. Thus, this model creates a counterfactual negative relationship between training and wages.

4   General structure of the Model

The labor-market is described by a standard matching function,[13] such that unemployed workers searching for a new job face a certain probability of being successful that is dependent on the tightness of the labor market. Tightness θ is defined as the relation of posted vacancies v and searching workers u, so that an increase in labor market tightness implies that each worker has a better chance to find a job. Of course the opposite is true for firms which are posting a vacancy. The higher the tightness of the market, the lower the chances to find a worker. However, it is not necessary to explicitly model the labor market to draw meaningful conclusions. To keep things as simple as possible I will therefore not describe the value of a vacancy and the value of unemployment in detail.

Instead I concentrate on a firm and a worker who have already met. At a given time the firm can employ only one worker. After two periods the relationship is terminated and both parties return to the labor market. I could as well assume that the worker will die after this second period. However, since this does not change the results in any way

---

[13]See Pissarides (2000).

Figure 2: Time path
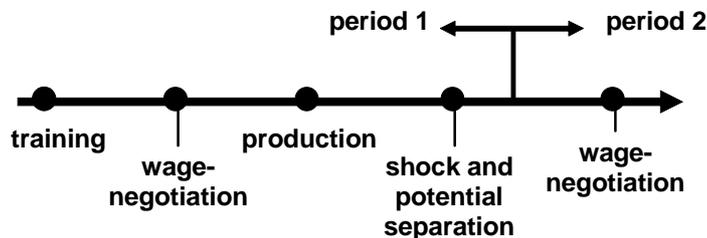
but does complicate the wage functions, I will not follow this approach.

At the beginning of the first period, the firm has the opportunity to train the worker, which improves her productivity $y(h)$ instantaneously.[14] In discrete-time models featuring firm-training it is usually assumed that the first period is the training period and productivity is improved only in the second period. However, in order to give a meaningful interpretation to the idea of wage rigidity, it is necessary that two wage-negotiations take place after the training-decision. Otherwise the wage rigidity would just be equivalent to a minimum wage and its effect on wage compression via altering the way wages are determined could not be analyzed. Wages are negotiated after the training decision but still at the beginning of the first period, i.e. before production takes place.[15,16]

At the end of period one the pair is hit by an idiosyncratic shock[17] π with distribution $g(\pi)$ which is joined additively to the productivity of the worker. In order to be able to derive analytical results regarding the comparison of wage compression in the benchmark and in the rigidity model it is necessary to assume uniformly distributed shocks. All the other results can be derived with a general distribution. Due to the shock the output of period two equals:

---

[14]Thus the worker produces with a higher productivity not only in the second period but in the first period as well.

[15]The timing of wage-negotiations is common in the training-literature - see for instance Acemoglu and Shimer (1999).

[16]We assume that the output of the worker is high enough to ensure positive wages.

[17]See for instance Pissarides (2000).

11

$$y_{t+1} = y(h) + \pi \tag{1}$$

In this way I am able to endogenize the separation decision. The match is only destroyed in case of shocks that are bad enough, i.e. below a certain threshold. This implies that high-skilled workers have a lower risk of getting fired than low-skilled workers, which makes sense. In case of a separation, both parties can search for a new partner at the beginning of period two. If they stay together, the wage will be newly negotiated, but now with the restriction that the wage cannot lie below the wage of the prior period. Of course this restriction is anticipated by the firm and reflected in the separation-decision. The timing of the most important events is illustrated in figure (2).

To be able to better illustrate the effects of wage rigidity, I first present a benchmark model, in which wages are negotiated without any restriction.

5 Benchmark

5.1 Value functions

In this section we assume that wages are fully flexible i.e. not rigid. The model serves as a Benchmark case to be compared with a model featuring rigid wages. Separations are endogenized via stochastic, idiosyncratic productivity-shocks. Whenever the shock lies below a certain threshold the worker and the firm will separate and return to the labor market. As discussed above, it is an advantage of this approach that training influences the probability of separations. It does not seem very plausible that all workers have the same risk of loosing their job, no matter how skilled they are.

The value of a firm with an employee is described by $J(y)$ while a vacancy is denoted by V. The value of a worker occupying a job is $W(y)$ and the value of an unemployed worker is $U(h)$, where h denotes the amount of human capital of that worker. The value

12

of unemployment is dependent on human capital if training is assumed to be general. I do not consider any other forms of human capital such as, for instance, education.

The value of a filled job at the beginning of the first period is:[18]

$$J(y_t) = y(h) - w_t^b(h) - c(h) + \rho \int_{y_q}^{\max} J(y_{t+1}) f(y_{t+1}) dy_{t+1} - \rho \int_{\min}^{y_q} V f(y_{t+1}) dy_{t+1} \quad (2)$$

The output of the worker $y(h)$ is an increasing function of her human capital. The cost of training $c(h)$ is assumed to be rising as well. Either the output of training has to be growing at a declining rate or the cost of the training has to be increasing at an enhancing rate to assure an interior solution.

The first three terms show the revenues and expenditures of the current period (output minus wages and training costs), while the integrals give the expected value of the firm next period, which has to be discounted by factor ρ. Note that the integral is over the actual output of the worker y_{t+1} and not over the productivity shock π. I have chosen this kind of notation since it implies more intuitive definitions of the thresholds - actually it is the output of the worker that the firm cares about and not the value of the shock. The additivity of the shock implies that the PDF of the second-period output equals:

$$f(y_{t+1}) = g(y_{t+1} - y(h))$$

so that the PDF of y_{t+1} is nothing else but the PDF of the shock shifted in the mean by the output of the first period $y(h)$. y_q is the threshold-productivity:[19] If output turns out to be lower than this threshold, the partnership will be terminated and the firm will get the (constant) value of a vacancy (second integral). If output is above y_q, the match

---

[18]The notation is very much in line with Pissarides (2000) or the Appendix in Acemoglu and Pischke (1999a).

[19]We have called the separation-threshold y_q, referring to quits, because in the benchmark model both parties agree to separate. This will be different in the rigidity model and thus there we refer to a separation as firing and use y_f to denote the threshold.

will continue. In this case the firm-value is $J(y_{t+1})$ which is dependent on the actual realization of the shock π. All these cases have to be weighted with their respective probabilities and added up over the domain of the probability distribution $[\min, \max]$. The threshold y_q is defined by:

$$J(y_q) = V \tag{3}$$

or alternatively by:

$$W(y_q) = U(h)$$

This implies that both parties are indifferent between continuing the partnership (in which case they would get $J(y_{t+1})$ respectively $W(y_{t+1})$) and terminating it (in that case they would get V respectively $U(h)$. For any shock lower than y_q, both parties agree to separate because their values in the outside labor-market are higher. For the remainder of the paper I assume free entry of firms, so that the value of a vacancy is zero at any time - if it were positive, new firms would enter the market, lowering the probability of all firms finding a worker and thereby driving down the value of a vacancy. To the contrary, if the value of a vacancy were negative, some firms would exit the market, the chances of the remaining firms to find a worker would go up and thereby the value of a vacancy until it has reached its equilibrium level zero - only then will there be no incentives for further adjustment.

The value of an employed worker in period one is very similar to the value of a firm. I write it as:

$$W(y_t) = w_t^b + \rho \int_{y_q}^{\max} W(y_{t+1}) f(y_{t+1}) dy_{t+1} - \rho \int_{\min}^{y_q} U(h) f(y_{t+1}) dy_{t+1} \tag{4}$$

The income of the current period is just equal to the wage, since the costs of the training are paid by the firm. The integrals again illustrate the expected value of the

Figure 3: Value functions of the second Period - Flexible Wages
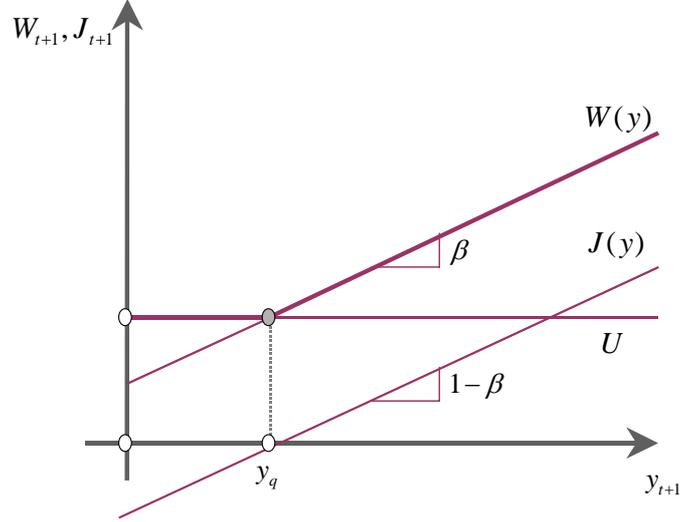
worker in the second period. If the output lies above y_q the match will continue and the worker will get value $W(y_{t+1})$, if output lies below the threshold-productivity she will quit and have the value of an unemployed worker $U(h)$.

Since the match terminates with certainty after the second period, the second-period value functions are just equal to the incomes during that period plus the respective values at the labor market that is:

$$J(y_{t+1}) = y_{t+1}(h) - w_{t+1}^b(h) + \rho V = y_{t+1}(h) - w_{t+1}^b(h) \tag{5}$$

$$W(y_{t+1}) = w_{t+1}^b(h) + \rho U(h) \tag{6}$$

Figure (3) shows the quitting threshold and the value functions of workers and firms for period two in dependence of the productivity shock. The slope of the value functions of the match is equal to $1 - \beta$ respectively β[20] because the threat-point of both parties

---
[20]β denotes the bargaining strength of the worker in Nash-bargaining - see the section on wages.

is independent of the idiosyncratic shock.[21] Thus the outcome of the shock is shared according to the respective bargaining powers of the parties. Due to the same reason, the value of unemployment is characterized by a horizontal line. The worker prefers the state with the higher value. Therefore she chooses unemployment for any shock lower than y_q. In the graph this is illustrated by the thick line. The firm has the alternative between $J(y_{t+1})$ and the value of a vacancy which is equal to zero. Of course, whenever the value of the job lies below zero, the firm will prefer to terminate the relationship. As can be seen in the picture, the firm and the worker will agree on whether to stay together or whether to separate. The worker's value function intersects the value of unemployment at the same value of output at which the job's value function turns negative. This is due to the way wages are determined. Nash-bargaining is always assuring that a positive rent to the match is shared between both partners according to their respective bargaining-powers. No matter how small this rent is, both partners get a positive share of it and therefore prefer to continue the match - as long as it is positive. At the threshold y_q the rent of the match is exactly zero and therefore nothing is to be shared - both the firm and the worker are indifferent between separating and continuing the match.

5.2   Wages

As was mentioned above, wages are determined by Nash-bargaining,[22] according to which the surplus of the match over the threat-points[23] of both parties is shared corresponding to their bargaining strength. From the perspective of the worker this means that her surplus over the threat-point $(W - U)$ has to be equal to the rent of the whole match

---

[21]Figure (3) shows the special case where $\beta = 1 - \beta = 1/2$. In all other cases the value functions would not be parallel.

[22]See for instance Shaked and Sutton (1984) for a game-theoretic foundation or Pissarides (2000) for an application to the matching framework.

[23]The threat-point or fall-back position in a bargain is the value a party would get in case of a brake-down of negotiations. As is standard, I assume that both parties are able to turn back to the labor-market instantaneously.

$(W + J - U - V)$ multiplied by her bargaining-power β. We then have:

$$W - U = \beta(W + J - U) \tag{7}$$

The bargained wage assures that the surplus of the match is shared according to the rule above. For both periods this results in the following wage formula:[24]

$$w_t^b = (1 - \rho)\, U(h) + \beta\, (y_t(h) - (1 - \rho)\, U(h)) \tag{8}$$

This is a standard result: The worker gets at least the value according to her threat-point[25] plus a share β of the surplus over that threat-point.

Using this wage function together with the value-functions of the firm in the definition of the quitting threshold (equations (8), (2) and (5) in (3)) we find that the threshold is given by the sum of both threat-points (where the threat-point of the firm V is zero):

$$y_q = (1 - \rho)\, U(h)$$

Thus the two parties agree to separate whenever the output of the second period lies below the value of unemployment. In this case the negotiated wage is so low that it is more profitable for the worker to look for another job.[26] But still the wage is so high that it lies above the output of the worker and the firm is making losses. In fact, Nash-bargaining would assure that the loss is shared between both parties. Therefore both, the firm and the worker are better off in case of a separation.

---

[24] Here we can see the advantage of assuming that workers do not die after the second period. Otherwise, we would have a different wage-rule for both periods and both wages would be more complicated since $(1 - \rho)\, U(h)$ would have to be replaced by $U_t(h) - \rho U_{t+1}(h)$.

[25] The threat-point has to be adjusted due to discounting.

[26] Remember that the shock is idiosyncratic, so that the output of the worker in an alternative firm is not affected.

5.3 Wage compression

The degree of wage compression can be determined by taking the derivative of wages (equation (8)) with respect to productivity respectively firm-training. We thus have:

$$\frac{\partial w_t^b}{\partial h} = \beta \frac{\partial y_t}{\partial h} + (1-\beta)(1-\rho)\frac{\partial U(h)}{\partial h} < \frac{\partial y_t}{\partial h} \tag{9}$$

The equation illustrates how the wage is reacting to changes in productivity. It might seem surprising that this equation is not so simple. The reason is that the wage is not only affected by the output but also by the outside alternatives of both parties. The first term in equation (9) is the direct effect of training on wages. Since the worker always gets a share β of the value of production, she also gets a share β of the training's value. But there is an additional, more indirect effect of training on wages which is working via the bargaining position of the worker. If training is assumed to be general, it will increase the value of unemployment to the worker since she will earn a higher wage when she finds a new job. This improves the threat-point of the worker and thereby her bargaining position. In consequence, her wage is higher. This effect is captured by the second term of the equation.

As is illustrated in equation (9) and is proved in Appendix A, wages react to training less than output. In other words, the wage structure is compressed. According to Acemoglu and Pischke (1999a) this is sufficient and necessary to induce firm-sponsored training.

5.4 Training

Before wages are negotiated the firm decides privately about the amount of training. The optimal decision is found by taking the derivative of the first-period value function with respect to training and setting it equal to zero:

$$\frac{\partial J(y_t)}{\partial h} = \frac{\partial y(h)}{\partial h} - \frac{\partial w_t^b(h)}{\partial h} - \frac{\partial c(h)}{\partial h} +$$

$$\rho \int_{y_q}^{\infty} \frac{\partial J(y_{t+1})}{\partial h} f(y_{t+1}) dy_{t+1} - \rho \frac{\partial y_q}{\partial h} J_{t+1}(y_q) f(y_q) = 0$$

The third term is the marginal cost of additional training, the first and the second term illustrate the effect on current profits. The output increases but at the same time the wage increases as well. The fourth term shows the effect of training on the firm-value next period while the last term illustrates the effects of a change in the quitting-threshold. However, this last term drops out since the value of the firm at this threshold is zero by definition (see equation (3)). [27]

By using the value function of the firm (equation (5)) and the wage rule (equation (8)), taking account of the fact that (due to the additivity of the shock) the output $y(h)$ can be taken out of the integral and rearranging, we arrive at the following - more meaningful - equation:

$$\frac{\partial c(h^b)}{\partial h} = \left[\frac{\partial y(h^b)}{\partial h} - \beta \frac{\partial y(h^b)}{\partial h} - (1-\beta)(1-\rho) \frac{\partial U(h^b)}{\partial h} \right] \left[1 + (1-F(y_q)) \right] \qquad (10)$$

where we have marginal costs on the left-hand side and marginal revenues on the right-hand side. The terms inside the first square brackets show the marginal revenue per period while the second square brackets gives the expected number of cases in which the firm and the worker will stay together. The first 1 stands for the first period. There cannot be a separation during that period so the match survives with probability one. However, it continues into period two only if the realization of the shock lies above the threshold y_q. $F(y_{t+1})$ is the cumulative distribution function of $f(y_{t+1})$. Therefore $F(y_q)$ is the probability of separation and $1 - F(y_q)$ the probability of staying together.

Let's get back to the terms in the first square brackets on the right-hand side of equation (10): Overall this illustrates the effect of training on the value of the firm each period (given that the match continues), which is the increase of the output of the worker

---

[27]It should be noted that this last result is NOT due to the assumption that the value of a vacancy is equal to zero. If it were not zero, there would be an additional term in the value function, as illustrated in equation (2). In this case those two terms would cancel out yielding the same result.

net of the increases in wages as defined in equation (9). By inspection of equations (10) and (9) it becomes clear that the firm will invest in the worker's human capital if and only if the wage-structure is compressed. Otherwise, the term inside the first square brackets - and thereby the marginal revenues to training - would be zero (or even negative). Considering this point I am able to confirm the results of Acemoglu and Pischke (1999a).

6 Rigidity model

6.1 Value functions

As already mentioned above, the rigidity model differs from the benchmark only with respect to the wage-negotiations of the second period. In principle, these are the same with the only restriction that the wage is not allowed to drop from the first period to the second. As we will see later, the restriction to wage-negotiations in the second period has consequences for the wages of the first period as well, although these are still freely negotiated. To account for the possibility that rigidity becomes binding, we have to add another state to the description of the second period, so that I can distinguish between situations in which the wage of the previous period is restricting the negotiations of the last period and situations in which it is not. To make things clear I add a superscript r to denote value functions describing the restricted case and a superscript u to denote the reverse. We thus have:

$$J^u(y_{t+1}) = y_{t+1}(h) - w_{t+1}(h) \tag{11}$$

$$J^r(y_{t+1}, w_t) = y_{t+1}(h) - w_t(h) \tag{12}$$

Again the firm-value of the second period is straight forward, it is just production minus wages. The unrestricted value function $J^u(y_{t+1})$ is exactly the same as in the

benchmark model (equation (5)), whereas the restricted value function $J^r(y_{t+1}, w_t)$ has another state-variable which is the wage of the previous period. The values of workers are straight forward as well and can be written as:

$$W^u(y_{t+1}) = w_{t+1}(h) + \rho U(h) \tag{13}$$

$$W^r(w_t) = w_t(h) + \rho U(h) \tag{14}$$

It should be noted that, in contrast to the firm value, the restricted value function of a worker no longer depends on the output of the match - the worker receives the same wage in any case - of course, only so long as the match is not destroyed.

Clearly, this distinction between two different value functions for the second period has consequences for the values of the first period as well: [28]

$$
\begin{aligned}
J(y_t) \quad = \quad & y(h) - w_t(h) - c(h) + \rho \int_{y_b}^{\max} J^u(y_{t+1}) f(y_{t+1}) dy_{t+1} + \\
& \rho \int_{y_f}^{y_b} J^r(y_{t+1}, w_t) f(y_{t+1}) dy_{t+1} - \rho \int_{\min}^{y_f} V f(y_{t+1}) dy_{t+1}
\end{aligned} \tag{15}
$$

$$
\begin{aligned}
W(y_t) \quad = \quad & w_t(h) + \rho \int_{y_b}^{\max} W^u(y_{t+1}) f(y_{t+1}) dy_{t+1} + \\
& \rho \int_{y_f}^{y_b} W^r(w_t) f(y_{t+1}) dy_{t+1} - \rho \int_{\min}^{y_f} U(h) f(y_{t+1}) dy_{t+1}
\end{aligned} \tag{16}
$$

The interpretation of the value-functions is analogous to the benchmark-model. Again the terms without an integral give the earnings of the first period, while all the integrals taken together constitute the expected value of the second period. The distinction between

---

[28] The ordering of the thresholds (i.e. the fact that $y_f < y_b$) will become clear by introspection of figure (4) resp. the section further below discussing the thresholds in more detail.

21

Figure 4: Value functions of the second Period - Rigidity-model
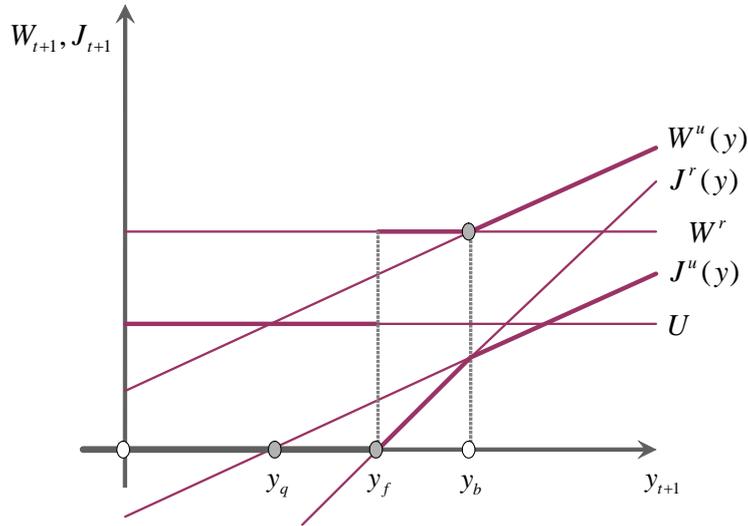
the cases where the first-period wage is restricting and where it is not, necessitates another threshold discriminating between these two cases. I call this threshold y_b, to make clear whether the wage is binding or not. So whenever the (partly) random y_{t+1} lies above this threshold, the freely negotiated wage of the second period will be higher than the wage of the first period and the restriction will not be binding. In this case both parties receive the unrestricted value of period two, $J^u(y_{t+1})$ respectively $W^u(y_{t+1})$. Whenever output is below y_b, the negotiated wage would lie below the wage of the previous period so that wages would have to be cut down. However, the management fears the bad effects of wage-cuts on the morale of the work-force and therefore prefers to keep the wage constant. Thus in this case the firm and the worker get the restricted values of period two $J^r(y_{t+1}, w_t)$ respectively $W^r(w_t)$.

As in the benchmark case there is a separation threshold (y_f), such that the match will be terminated for lower shocks. Just as in the benchmark-model in such a case the parties receive the values V and $U(h)$. The thresholds and their relations to one another are discussed in more detail further below.

Again the value functions and thresholds are illustrated graphically (see figure (4)).

The unrestricted value functions have the same slopes ($1 - \beta$ respectively β) as the value functions in the benchmark model (see figure (3)). The restricted value function of the worker is horizontal: Since in these cases the worker gets a fixed wage, the value will be independent of actual productivity. In turn, the restricted value function of the firm has slope one. The wage is fixed and therefore any increase in output will lead to an one to one increase in firm value.

The thick line indicates the actual value of the worker respectively the firm for all levels of the shock. The worker prefers the larger of the unrestricted and the restricted value so whenever output lies below the threshold y_b the wage-restriction will become binding. For the firm it is just the other way around: Due to the fear of bad morale it will always get the lower alternative. However, the firm has the possibility of firing the worker and it will do so whenever the value of the firm becomes negative - this implies the second kink of the thick line at the threshold y_f, which lies clearly above y_q. From the worker's perspective this means a jump form the horizontal line W^r to the value of unemployment, which is horizontal as well. It is clear that - in contrast to the benchmark - the worker would always prefer to stay employed. For shocks between y_f and y_q the worker would even be willing to accept a wage cut in order to stay employed. However, the firm will not accept a wage cut because it fears that the worker will provide no effort in such a case.

6.2   Wages

Since the wages are determined for the first time at the beginning of the first period there is no previous-period wage that could become a restriction. Consequently, wage negotiations are unrestricted. Nevertheless, wage rigidity will play a role in these negotiations since the prospect of a binding restriction alone is enough to alter the outcome of the bargain. It is this modification that gives rise to enhanced wage compression, as will be shown further below.

The wage is again found by plugging in the value functions into the sharing rule of Nash-bargaining (equation (7)) which results in: [29]

$$w_t = (1 - \rho) U(h) + \beta \left(y(h) - (1 - \rho) U(h) \right) + \frac{\beta \rho \int_{y_f - y_t}^{y_b - y_t} \pi g(\pi) d\pi}{1 + \rho \left(F(y_b) - F(y_f) \right)} \qquad (17)$$

Compared to the wage of the benchmark model (see equation (8)) there is one additional term on the right-hand side.[30] Besides that, the wage outcomes are equivalent. But what is this additional term? It is the compensation of the firm for the possibility that it might have to pay a wage "too high", i.e. not according to the unrestricted bargaining rule. In the numerator we can find the deviation of the output of the second period from the output of the current period in all those cases that the wage restriction is binding but the worker not fired. Remember that the output of the second period is equal to $y(h) + \pi_t$, while the output of the first period is just $y(h)$ - thus π_t is the deviation from one period to the other. This deviation is irrelevant for all those cases that the wage is freely negotiated, because in these cases the wage is adjusted accordingly. But for all those states that the wage would have to be cut and this is hindered by wage rigidity, the adjustment is not possible. This benefits the worker because she gets a wage higher than she would get otherwise (under the condition of free bargaining), but hurts the firm. These possibilities are foreseen by both parties and reflected in the value-functions. Thus because the worker will have an advantage over the firm in the second period and both parties are foreseeing this, the worker has to compensate the firm by accepting a lower first-period wage, compared to the benchmark. In this sense, wage rigidity can be interpreted as an insurance against wage cuts. The firm provides the insurance to the worker and pays a wage that is at least as high as the wage of the current period. The difference between the first-period wage in the benchmark and the rigidity model is the insurance premium that the worker has to pay to the firm.

---

[29] See Appendix B for a proof.

[30] It should be noted, that the value of the integral is negative since both of the boundaries of the integral are negative.

The only term that remains to be explained is the denominator, which is equal to one plus the probability that the wage is binding in the second period. Thus the denominator gives the expected number of cases in which the currently negotiated wage will be paid. To interpret this, it is useful to rearrange the wage-equation by multiplying both sides by the denominator. We then have: [31]

$$(1 + F(y_b) - F(y_f))w_t = \beta \left[(1 + \rho F(y_b) - \rho F(y_f))y(h) + \rho \int_{y_f - y_t}^{y_b - y_t} \pi g(\pi) d\pi \right]$$

Now we can see on the left-hand side of the equation the wage payments of the firm for all those cases that the currently negotiated wage has to be paid. On the right-hand side inside the square brackets we can see the expected output in these cases. The output y_t plus the expected deviations from this output. According to Nash-bargaining, the worker should get a share according to her bargaining strength β and therefore this term has to be multiplied by β. Summarizing, it can be said that the bargaining of the first period assures that, overall, both parties are compensated according to their respective bargaining-strengths. If one party is expected to have a future advantage over the other party, the Nash-bargaining in the present period assures that the aggrieved party is compensated by the profiting party.

Since the match ends after the second period for sure, wage-rigidity can no longer have such an effect on negotiations. Instead, the freely negotiated wage is exactly equal to the wage of the benchmark (equation (8)). By assumption the wage will only be freely negotiated if the outcome lies above the wage of the first period. Thus the actual wage of period two is given by the maximum of the two:

$$w_{t+1} = Max[(1 - \rho) U(h) + \beta (y_{t+1}(h) - (1 - \rho) U(h)), w_t]$$

Using the assumption of uniformly distributed productivity shocks I am able to show that in the rigidity model the wage structure of the first period is more compressed than in the benchmark:[32]

---

[31] To save notation we have left out the terms related to the threat-point of the worker.

[32] For a proof see Appendix C.

$$\frac{\partial w_t}{\partial h} < \frac{\partial w_t^b}{\partial h}$$

Thus an increase in firm training will have a smaller effect on wages when wages are rigid. According to the predictions of Acemoglu and Pischke (1999a) this should lead to higher firm-training. I will come back to this question further below but first I discuss the thresholds in more detail.

6.3   Discussion of thresholds

The binding-threshold (separating the states where the wage rigidity is relevant and where it is not) can be defined in three equivalent ways:

$$W^u(y_b) = W^r(w_t) \tag{18}$$

$$J^u(y_b) = J^r(y_b, w_t)$$

$$w_{t+1}(y_b) = w_t$$

where $w_{t+1}(y_{t+1})$ denotes the freely bargained wage. Thus, as the last equation illustrates, at the quitting-threshold the bargained wage is just equal to the wage of the previous period. Therefore the worker is indifferent between the old wage and the freely negotiated wage and, consequently, both the restricted and the unrestricted values are equal to each other. Alternatively, the threshold could be defined by using firm-value functions.

The separation threshold of the rigidity-model y_f is not equal to the separation threshold of the benchmark model y_q due to the inflexibility of wages. The letter f is used to denote the firing which will take place if output is lower than this threshold. It is referred to as a firing because in such a case the worker would prefer to keep up the relationship since she would always get the same wage $w_{t+1} = w_t$ and therefore never has any interest

to terminate the relationship. In figure (4) this is illustrated by the horizontal line W^r which lies above the value of unemployment U for any value of output.

Nevertheless, for the firm a termination is more profitable and it therefore fires the worker. This is in contrast to the benchmark-model where both parties will agree to separate if output lies below y_q. It might be criticized that this implies inefficient separations. Both parties could be better off, if they agreed on a lower wage. However, this is exactly what the evidence tells us. Managers prefer layoffs to wage cuts because it moves the problem of bad morale outside of the firm. In other words, the firm does not favor the wage cut because it anticipates that this will induce the worker to shirk and this would be even worse than a separation. In that sense, we cannot call the separation inefficient: Actually the firm is acting rational.

Although the separation-thresholds are different, they are defined in a very similar way as:

$$J^r(y_f, w_t) = V \tag{19}$$

Similar to the quitting-threshold, the firing-threshold is found by setting the value of the firm equal to its threat-point, the value of a vacancy. For values of output below y_f, the firm's value of keeping up the match is lower than terminating it and thus it fires the worker. Of course, if output is so low that a separation occurs, output will be low enough to make wage rigidity binding (if the parties were not separating). Therefore, we have to use the restricted value function J^r. The use of the restricted value-function explains the difference between y_f and y_q. In the rigidity model the firm is not able to lower the wage from one period to the other, whereas in the benchmark model the wage can go down to zero. As long as the wage in the first period was not negative, it follows that for low (bad) shocks the wage of the second period in the rigidity model has to be higher than in the benchmark model.[33] Due to this higher wage, the firm is less reluctant to fire the

---

[33]This is true for all states below the binding-threshold.

worker and consequently separations are more frequent. Summarizing, we can state the following about the order of thresholds:

$$y_b > y_f > y_q \implies F(y_b) > F(y_f) > F(y_q)$$

Consequently, we can distinguish four different intervals for the second period. From max to y_b, wages and values are the same in both models. From y_b to y_f the wage-rigidity becomes relevant so that the wage in the benchmark is lower. From y_f to y_q a worker is fired in the rigidity model but not in the benchmark model. Below y_q workers in both models get unemployed.

By plugging in the value and wage functions into the definitions of the thresholds we can easily find that:

$$y_b = y_t + \frac{\rho \int_{y_f - y_t}^{y_b - y_t} \pi g(\pi) d\pi}{1 + \rho F(y_b) - \rho F(y_f)} \tag{20}$$

$$y_f = w_t \tag{21}$$

These equations can be interpreted as follows. Because the wage adjustment[34] is no longer necessary in the second period (the relationship will be terminated afterwards), for equal productivities the bargained wage of the second period is generally higher than the wage of the first period. Therefore, the output of the worker has to fall by the value of that adjustment-term in order to make rigidity binding. The interpretation of the second threshold is straight forward: Since the worker gets w_t for sure if the output is below the binding-threshold y_b, the firm will get the residual of the output over that wage. This residual turns negative as soon as output lies below the wage and then the firm will fire the worker.

---

[34]The additional term in equation (17).

6.4 Training

Again the optimal amount of training is found by setting the derivative of the value function with respect to human capital equal to zero:[35]

$$\frac{\partial J(y_t)}{\partial h} = -c'(h) + \frac{\partial y(h)}{\partial h} - \frac{\partial w_t(h)}{\partial h} + \rho \int_{y_b}^{\infty} \frac{\partial J^u(y_{t+1})}{\partial h} f(y_{t+1}) dy_{t+1} +$$

$$+\rho \int_{y_f}^{y_b} \frac{\partial J^r(y_{t+1}, w_t)}{\partial h} f(y_{t+1}) dy_{t+1} - \rho \frac{\partial y_b}{\partial h} J^u_{t+1}(y_b) f(y_b) +$$

$$+\rho \frac{\partial y_b}{\partial h} J^r_{t+1}(y_b, w_t) f(y_b) - \rho \frac{\partial y_f}{\partial h} J^r_{t+1}(y_f, w_t) f(y_f)$$

$$= -c'(h) + \frac{\partial y(h)}{\partial h} - \frac{\partial w_t(h)}{\partial h} + \rho \int_{y_b}^{\infty} \frac{\partial J^u(y_{t+1})}{\partial h} f(y_{t+1}) dy_{t+1}$$

$$+\rho \int_{y_f}^{y_b} \frac{\partial J^r(y_{t+1}, w_t)}{\partial h} f(y_{t+1}) dy_{t+1} = 0$$

After plugging in the definition of values and wages (as given in equations (17), (12) ,(11)) and rearranging we arrive at the following equation which is very similar to the benchmark:

$$c'(h^r) = \left[\frac{\partial y(h)}{\partial h} - \beta \frac{\partial y(h)}{\partial h} - (1 - \beta)(1 - \rho)\frac{\partial U(h^r)}{\partial h}\right][1 + (1 - F(y_f))] \qquad (22)$$

Actually, the only difference between the optimality condition of the benchmark given in equation (10) and the equation above is the separation probability inside the second square-brackets at the right-hand side. The reason for this surprising result lies in the flexibility of wage-bargaining in the first period. As already discussed in the section above, it assures that the value of the whole match is shared according to the respective bargaining powers of the parties. Mathematically, the additional term in the wage of the first period implies that the effect of wage rigidity on the values of workers and firms just cancels out.

Consequently, the only difference between the two models, when it comes to firm training, is the difference in separation rates. A higher separation rate makes it less likely

---

[35]Note that $J^u_{t+1}(y_b) = J^r_{t+1}(y_b, w_t)$ and $J^r_{t+1}(y_f, w_t) = 0$ by definition of the thresholds.

for the firm to get a return on its investment by paying wages below productivity in the second period. Of course the firm can still get some return on the training during the first period, but a higher risk of termination in the second period will lower the profitability of the second period. As already discussed in the section above, this separation probability is higher in the rigidity model, compared to the benchmark. It follows that training in the rigidity model is in fact lower, not higher as predicted by Acemoglu and Pischke (1999a) :

$$y_f > y_q \implies F(y_f) > F(y_q) \implies h^r < h^b$$

The difference in results is due to the endogeneity of separations. In Acemoglu and Pischke (1999a) separations occur at an exogenous rate. Thus training can have no effect on the probability of separations. This is not very plausible, given that training is improving the output of a worker in any state of the world. Instead a worker with higher productivity should be better able to overcome bad times. The concept of exogenous separation-rates seems even more problematic in the context of minimum wages or wage rigidity. Both phenomena restrict the flexibility of the firm in a severe way. They do not allow the firm to cut wages below a certain level. It appears only natural that firms react by being less reluctant to fire workers - and indeed this is confirmed by empirical surveys like the ones of Bewley (1999) or Agell and Lundborg (1995). Therefore, it seems even more important to allow for endogenous separations in the context of such restrictions. It might be interesting to explore the potential of endogenizing separations in models of minimum wages to bring the theoretical results back in line with empirical findings.

7   Conclusion

By endogenizing the separation decision I was able to show that higher wage compression does not necessarily lead to more training investments as implied by the model of Acemoglu and Pischke (1999a) assuming that jobs are destroyed at an exogenous rate. This

assumption implies that all workers face the same risk of loosing their job no matter how skilled they are. This is not only implausible but also at odds with the empirical literature on firm training which is pointing towards a negative relationship between a worker's training and her turnover-rate.[36] By assuming that the productivity of the match is hit by an idiosyncratic shock I am able to endogenize the separation decision so that workers are only fired if the shock lies below a certain threshold. The higher the human capital of the worker the lower is this separation-threshold implying a higher risk of getting unemployed for untrained workers in comparison to trained (or better trained) workers.

Wage rigidity is modelled by implying the restriction that wages of the current period are not allowed to fall below wages in the preceding period, whereas otherwise wages are negotiated freely via standard Nash-bargaining. Worker and firm - foreseeing this - negotiate a lower starting wage than in an unrestricted world without rigidities. In principle the firm offers an insurance against wage-cuts and the lower starting wage is the insurance premium. [37]

I was able to show that this wage rule leads to an increase in wage compression compared to a benchmark model with unrestricted Nash-bargaining. However, at the same time the rigid wage leads to higher turnover rates since firms are not allowed (or not willing) to lower wages. This is again in line with the empirical literature which suggests that managers prefer layoffs to wage cuts, because they fear the adverse effects on worker-morale.[38] Due to this increase in the probability of separations firm training is lower in the rigidity model.

This is especially interesting with regard to the empirical literature on the effect of minimum wages on firm training. As noted by Acemoglu and Pischke (2003) the results of these empirical studies are rather mixed. Applying the endogeneity of separations to a model of minimum wages will likely lead to similar results as pointed out in this paper.

---

[36] See for instance Lynch (1991).

[37] This interpretation should be taken with care: An insurance in this context does not really make sense since workers are assumed to be risk-neutral.

[38] See Bewley (2002).

This might be an interesting field for further research.

From a welfare point of view it is clear that in this model wage rigidity is a bad thing. Not only does it lead to a higher rate of turnover and thus to higher unemployment, but it also depresses firms' training investments. Thus we have higher unemployment and lower human capital in the rigidity-model. It might be argued that in a model with risk-averse agents, welfare might still increase with rigidity, because the variation in wages is diminished. However, this increase in welfare is counteracted by the increased risk of being unemployed. Thus it seems very unlikely that risk-aversion might change this result.

8   References

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

Acemoglu, Daron and Pischke, Jörn-Steffen (1998): *Why do Firms train? Theory and evidence*, Quarterly Journal of Economics 113, 79-119

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

Acemoglu, Daron and Pischke, Jörn-Steffen (1999b): *Beyond Becker: Training in Imperfect Labour Markets*, The Economic Journal 109, 112-142

Acemoglu, Daron and Pischke, Jörn-Steffen (2003): *Minimum Wages and On-the-Job Training*, Research in Labor Economics 22, 159-202

Acemoglu, Daron and Shimer, Robert (1999): *Holdups and Efficiency with Search Frictions*, International Economic Review 40, 827-849

Agell, Jonas and Bennmarker, Helge (2003): *Endogenous Wage Rigidity*, CESifo WP 1081

Agell, Jonas and Lundborg, Per (1995): *Theories of Pay and Unemployment: Survey Evidence from Swedish Manufacturing Firms*, Scandinavian Economic Journal 97, 295-307

Altonji, Joseph and Devereux, Paul (2000): *The Extent and Consequences of downward Nominal Wage Rigidity*, Research in Labor Economics 19, 383-431

Baker George, Gibbs Michael and Holmstrom Bengt (1994): *The Wage Policy of a Firm*, Quarterly Journal of Economics 109, 921-955

Barron John, Berger Mark and Black Dan (1999): *Do Workers pay for On-the-Job Training*, Journal of Human resources 35, 235-252

Bassanini, Andrea and Brunello, Giorgio (2003): *Is Training More Frequent When Wage Compression is Higher? Evidence form the European Community Household Panel*, IZA-DP 839

Becker, Gary S. (1962): *Investment in Human Capital: A Theoretical Analysis*, Journal of Political Economy 70, 9-49

Beissinger, Thomas and Knoppik, Christoph (2001): *Downward Nominal Rigidity in West-German Earnings 1975-1995*, German Economic Review 2, 385-418

Bewley, Truman (1999): Why wages don't fall during a recession, Harvard University Press, Cambridge

Bewley, Truman (2002): *Fairness, Reciprocity and Wage Rigidity*, Cowles Foundation DP 1383

Booth, Alison and Bryan, Mark (2002): *Who pays for General Training? New Evidence for British Men and Women*, IZA-DP 486

Camerer, Colin and Thaler, Richard (1995): *Ultimatums, Dictators and Manners*, Journal of Economic Perspectives 9, 209-219

Card, David and Hyslop, Dean (1997): *Does Inflation grease the Wheels of the Labor Market?*, in Romer and Romer, eds.: Reducing inflation: motivation and strategy, University of Chicago Press, Chicago, 71-114

Danthine, Jean-Pierre and Kurmann, Andre (2004a): *Fair Wages in a New Keynesian model of the Business Cycle*, Review of Economic Dynamics 7, 107-142

Danthine, Jean-Pierre and Kurmann, Andre (2004b): *Efficiency Wages Revisited: The Internal Reference Perspective*, Universite de Lausanne, Ecole des HEC, DEEP, Cahiers de Recherches Economiques du Departement d'Econometrie et d'Economie politique (DEEP)

Falk Armin, Gächter Simon and Kovacs Judith (1998): *Intrinsic Motivation and Extrinsic Incentives in a Repeated Game with Incomplete Contracts*, Journal of Economic Psychology 20, 251-284

Fehr, Ernst and Falk, Armin (1999): *Wage rigidity in a competitive incomplete contract market*, Journal of Political Economy 107, 106-134

Fehr, Ernst and Götte, Lorenz (2000): The Robustness and Real Consequences of Nominal Wage Rigidity, CESifo WP 335

Gerfin, Michael (2003a): *Work-related Training and Wages, an Empirical Analysis for Male Workers in Switzerland*, University of Bern, Institute for Economics, Diskussionsschriften 03-16

Gerfin, Michael (2003b): *Firm-sponsored Work-related Training in Frictional Labor Markets*, University of Bern, Institute for Economics, Diskussionsschriften 03-17

Hall, Robert (2003): *Wage determination and Employment* Fluctuations, NBER WP 9967

Holden, Steinar (2002): *Downward nominal wage rigidity - contracts or fairness considerations*, working paper, Oslo

Howitt, Peter (2002): *Looking inside the labor market: a review article*, Journal of Economic Literature 40, 125-138

Katsimi, Margarita (2003): *Training, Job Security and Incentive Wages*, CESifo WP 955

Krause, Michael and Lubik, Thomas (2003): *The (Ir)relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions*, Center of Economic Research, Discussion Paper 113

Ljungqvist, Lars and Sargent, Thomas (1998): *The European unemployment dilemma*, Journal of Political Economy 106, 514-550

Ljungqvist, Lars and Sargent, Thomas (2002): *The European employment experience*, CEPR discussion paper 3543

Loewenstein, Mark and Spletzer, James (1998): *Dividing the Costs and Returns to General Training*, Journal of Labor Economics 16, 142-171

Loewenstein, Mark and Spletzer, James (1999): *General and Specific Training: Evidence and Implications*, Journal of Human Resources 34, 710-733

Lynch, Lisa (1991): *The Role of Off-the-Job vs. On-the-Job Training for the Mobility of Women Workers*, American Economic Review 81, 151-156

Malcomson, James (1999): *Individual Employment Contracts*, in Ashenfelter and Card, eds.: Handbook of Labor Economics 3, 2291-2372

Parking, Michael (2001): *What Have We Learned about Price Stability?*, in Price Stability and the Long-Run Target for Monetary Policy, Ottawa, Bank of Canada 223–259

Pissarides, Christopher (2000): *Equilibrium Unemployment Theory*, 2nd edition, MIT Press, Cambridge and London

Shaked, Avner and Sutton, John (1984): *Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model*, Econometrica 52, 1351-1364

Shimer, Robert (2004): The Consequences of Rigid Wages in Search Models, NBER working paper 10326

Strand, Jon (2003): *Wage Bargaining Versus Efficiency Wages: A Synthesis*, Bulletin of Economic Research 55, 1-20

Taylor, John (1979): *Staggered wage setting in a macro model*, American Economic Review 69, 108-113

9   Appendix A: wage compression in the benchmark

Wages would not be compressed in the benchmark only if U were equal to $\frac{y}{1-\rho}$. In this extreme case the threat-point of the worker is so high that she always gets a wage equal to her output. The derivative of the wage with respect to productivity as given in equation 9 would then be:

$\frac{\partial w_t^b}{\partial h} = \beta \frac{\partial y_t}{\partial h} + (1-\beta)\frac{\partial y_t}{\partial h} = \frac{\partial y_t}{\partial h}$

Wages react one to one to changes in productivity and thus the wage structure is not compressed.

Values of U higher than $\frac{y}{1-\rho}$ will make no sense, since in that case, the alternatives of the worker are unambiguously better, the pair would separate immediately at the beginning of the first period and no training would take place. Of course in practice the threat-point of the worker will be much lower. It is implausible that the worker can get a wage above productivity in another firm (which would be implied by $U = \frac{y}{1-\rho}$ since the denominator is smaller than zero). Additionally, the worker will not find a new job with certainty but will stay unemployed with a certain probability. This decreases the value of unemployment even further.

All these factors diminish the sensitivity of the worker's threat-point to productivity and thus create wage compression.

10   Appendix B: wage-rule of the rigidity model

I use the same rule to determine the wage as in the benchmark, equation 7:

$W - U = \beta(W + J - U)$

By plugging in the value functions for J and W as defined in equations 15 and 16 we get:[39]

$$w_t + \rho \int_{y_b}^{\max} W^u(y_{t+1})f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} W^r(w_t)f(y_{t+1})dy_{t+1} +$$

$$\rho \int_{\min}^{y_f} U(h)f(y_{t+1})dy_{t+1} - U =$$

$$\beta[y_t(h) + \rho \int_{y_b}^{\max} J^u(y_{t+1})f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} J^r(y_{t+1}, w_t)f(y_{t+1})dy_{t+1} +$$

$$\rho \int_{y_b}^{\max} W^u(y_{t+1})f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} W^r(w_t)f(y_{t+1})dy_{t+1}$$

$$+\rho \int_{\min}^{y_f} U(h)f(y_{t+1})dy_{t+1} - U] =$$

$$\beta[y_t(h) + \rho \int_{y_b}^{\max} [J^u(y_{t+1}) + W^u(y_{t+1})]f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} [J^r(y_{t+1}, w_t)$$

$$+W^r(w_t)]f(y_{t+1})dy_{t+1} + \rho \int_{\min}^{y_f} U(h)f(y_{t+1})dy_{t+1} - U]$$

Using the fact that the wage-rule equation 7 is valid in the second period as well, the unrestricted value-functions cancels out (with the exception of the value of unemployment). Plugging in the equations 12 and 14 for the remaining value-functions of the second period we get:

$$w_t + \rho \int_{y_b}^{\max} U(h)f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} [w_t + \rho U(h)]f(y_{t+1})dy_{t+1} +$$

$$\rho \int_{\min}^{y_f} U(h)f(y_{t+1})dy_{t+1} - U =$$

$$\beta[y_t(h) + \rho \int_{y_b}^{\max} U(h)f(y_{t+1})dy_{t+1} + \rho \int_{y_f}^{y_b} [y_{t+1} + \rho U(h)]f(y_{t+1})dy_{t+1} +$$

$$\rho \int_{\min}^{y_f} U(h)f(y_{t+1})dy_{t+1} - U]$$

By merging the terms with U the equation simplifies to:

$$w_t + \rho \int_{y_f}^{y_b} [w_t + (\rho - 1)U(h)]f(y_{t+1})dy_{t+1} + (\rho - 1)U =$$

$$\beta[y_t(h) + \rho \int_{y_f}^{y_b} [y_{t+1} + (\rho - 1)U(h)]f(y_{t+1})dy_{t+1} + (\rho - 1)U]$$

[39]Since the training cost is already sunk, it does not appear in the wage negotiations.

Now use the definition of $y_{t+1} = y_t + \pi$:

$$w_t + \rho \int_{y_f}^{y_b} [w_t + (\rho - 1)U(h)]f(y_{t+1})dy_{t+1} + (\rho - 1)U =$$

$$\beta[y_t(h) + \rho \int_{y_f}^{y_b} [y_t + \pi + (\rho - 1)U(h)]f(\pi)dy_{t+1} + (\rho - 1)U]$$

The only term in this equation that is random is the π on the right-hand side. All the other terms are constant and therefore can be taken out of the integral:

$$w_t + (\rho - 1)U + \rho[F(y_b) - F(y_f)][w_t + (\rho - 1)U(h)] + (\rho - 1)U =$$

$$\beta[y_t(h) + (\rho - 1)U + \rho[F(y_b) - F(y_f)][y_t + (\rho - 1)U(h)] + \rho \int_{y_f}^{y_b} \pi f(\pi)dy_{t+1}]$$

By joining the terms and bringing all the U to the right-hand side we get:

$$w_t[1 + \rho[F(y_b) - F(y_f)]] =$$

$$(1 - \rho)U[1 + \rho[F(y_b) - F(y_f)]] + \beta[(y_t(h) - (1 - \rho)U)[1 + \rho[F(y_b) - F(y_f)]]$$

$$+\rho \int_{y_f}^{y_b} \pi f(\pi)dy_{t+1}]$$

Finally, we arrive at the wage given in equation 17 by dividing through the term in square brackets on the left-hand side:

$$w_t = (1 - \rho)U + \beta[(y_t(h) - (1 - \rho)U)] + \frac{\beta\rho \int_{y_f}^{y_b} \pi f(\pi)dy_{t+1}}{1 + \rho[F(y_b) - F(y_f)]}$$

11 Appendix C: wage compression in the rigidity model

To see whether wage-compression is higher in the benchmark or in the rigidity model, it is sufficient to look at the extra term in equation 17 giving the wage of the rigidity model, since the remaining terms (output and the value of unemployment) react equally in both models. The wage structure of the rigidity model is more compressed if this term is decreasing with output and vice versa. First of all I define the extra-term in the rigidity wage as Λ and use the assumption of uniformly distributed productivity shocks to get:

$$\Lambda = \frac{\rho \int_{y_f - y_t}^{y_b - y_t} \pi_t f(\pi_t)d\pi_t}{1 + \rho F(y_b) - \rho F(y_f)} = \frac{\rho \int_{y_f - y_t}^{y_b - y_t} \pi_t \frac{1}{\max - \min} d\pi_t}{1 + \frac{\rho(y_b - y_f)}{\max - \min}} = \frac{\rho \frac{(y_b - y_t)^2 - (y_f - y_t)^2}{2(\max - \min)}}{\frac{\max - \min + \rho(y_b - y_f)}{\max - \min}}$$

$$= \frac{\rho(y_b-y_t)^2-(y_f-y_t)^2}{2[\max-\min+\rho(y_b-y_f)]} = \rho\frac{y_b{}^2-y_f{}^2-2y(y_b-y_f)}{2[\max-\min+\rho(y_b-y_f)]}$$

By noting that $y_b{}^2 - y_f{}^2 = (y_b + y_f)(y_b - y_f)$ this equation simplifies to:

$$\Lambda = \rho\frac{(y_b+y_f-2y)(y_b-y_f)}{2[\max-\min+\rho(y_b-y_f)]}$$

Now the derivative of Λ with respect to productivity can be written as:

$$\frac{\partial\Lambda}{\partial y} = \rho\frac{\frac{\partial(y_b+y_f-2y)}{\partial y}(y_b-y_f)2[\max-\min+\rho(y_b-y_f)]}{4[\max-\min+\rho(y_b-y_f)]^2}+$$

$$+\rho\frac{\frac{\partial(y_b-y_f)}{\partial y}(y_b+y_f-2y)2[\max-\min+\rho(y_b-y_f)]-\frac{\partial(y_b-y_f)}{\partial y}2\rho(y_b+y_f-2y)(y_b-y_f)}{4[\max-\min+\rho(y_b-y_f)]^2}=$$

$$\rho\frac{\frac{\partial(y_b+y_f-2y)}{\partial y}(y_b-y_f)2[\max-\min+\rho(y_b-y_f)]}{4[\max-\min+\rho(y_b-y_f)]^2}+$$

$$+\rho\frac{\frac{\partial(y_b-y_f)}{\partial y}(y_b+y_f-2y)2[\max-\min]}{4[\max-\min+\rho(y_b-y_f)]^2} \tag{23}$$

It turns out to be useful to not further split up the derivatives of the sum and the difference of the thresholds. For convenience let me repeat the definitions of these thresholds as given in equations 20 and 21:

$$y_b = y_t + \frac{\rho\int_{y_f-y_t}^{y_b-y_t}\pi_t f(\pi_t)d\pi_t}{1+\rho F(y_b)-\rho F(y_f)}$$

$$y_f = \beta y_t + (1-\beta)(1-\rho)U + \beta\frac{\rho\int_{y_f-y_t}^{y_b-y_t}\pi_t f(\pi_t)d\pi_t}{1+\rho F(y_b)-\rho F(y_f)}$$

Consequently the difference between the two thresholds is:

$$y_b - y_f = (1-\beta)(y_t + \frac{\rho\int_{y_f-y_t}^{y_b-y_t}\pi_t f(\pi_t)d\pi_t}{1+\rho F(y_b)-\rho F(y_f)} - (1-\rho)U) \tag{24}$$

while the sum of the two thresholds is given by:

$$y_b + y_f = (1+\beta)(y_t + \frac{\rho\int_{y_f-y_t}^{y_b-y_t}\pi_t f(\pi_t)d\pi_t}{1+\rho F(y_b)-\rho F(y_f)}) + (1-\beta)(1-\rho)U \tag{25}$$

By taking the derivatives of equations 24 and 25 with respect to productivity y, plugging them both into equation 23 and bringing all terms with Λ to the left-hand side we get:

$$\frac{\partial \Lambda}{\partial y}(1 - \frac{(1+\beta)\rho(y_b - y_f)}{2[\max - \min + \rho(y_b - y_f)]} + \rho\frac{-(1-\beta)(y_b + y_f - 2y)2[\max - \min]}{4[\max - \min + \rho(y_b - y_f)]^2})$$

$$= \rho\frac{[1 + \beta + (1-\beta)(1-\rho)\frac{\partial U}{\partial y} - 2](y_b - y_f)2[\max - \min + \rho(y_b - y_f)]}{4[\max - \min + \rho(y_b - y_f)]^2} +$$

$$+ \rho\frac{(1-\beta)(1 - (1-\rho)\frac{\partial U}{\partial y})(y_b + y_f - 2y)2[\max - \min]}{4[\max - \min + \rho(y_b - y_f)]^2} \tag{26}$$

This equation looks rather complicated. However, the only thing of relevance are the signs of the numerators. The term in brackets on the left-hand side of the equation is clearly positive since the second term inside the brackets is smaller than one while the third term is positive (since $y_b + y_f - 2y < 0$ as can be seen from equation 25). Consequently, Λ will have the same sign as the right-hand side of the equation above. Thus the problem boils down to the determination of the signs of $[1 + \beta + (1-\beta)(1-\rho)\frac{\partial U}{\partial y} - 2]$ (taken out of the first term) and of $(1 - (1-\rho)\frac{\partial U}{\partial y})$ (taken out of the second term) - the signs of all the other terms are obvious.

To do so we need to look more closely at the threat-point of the worker: As discussed in Appendix A U is most likely lower than $\frac{y}{1-\rho}$. Nevertheless it is useful to discuss this extreme case to clarify the relationship of the thresholds.

Given that $U = \frac{y}{1-\rho}$, the difference of the thresholds simplifies to:

$$y_b - y_f = (1-\beta)\frac{\rho \int_{y_f - y_t}^{y_b - y_t} \pi_t f(\pi_t) d\pi_t}{1 + \rho F(y_b) - \rho F(y_f)}$$

The only solution to this equation is obviously $y_b = y_f = y_t$, since a positive difference between the two thresholds would imply a negative value for the integral - which is a contradiction - and vice versa. It follows that the wages in the benchmark and the rigidity model are exactly equal to each other. Moreover, all the thresholds are equal as well, so that the worker will quit, whenever her output decreases due to a negative shock - in that case she can earn more on the labor market no matter whether wages are principally downwards rigid. Since the wage rigidity can never be binding there is no need to compensate the firm and so wages in both models are equal.

However, as discussed in Appendix A, whenever U is smaller than $\frac{y}{1-\rho}$, the threat point of the worker is less sensitive to changes in productivity than output. Then it is obvious that the following is true:

$1 + \beta + (1-\beta)(1-\rho)\frac{\partial U}{\partial y} - 2 < 0$

$1 - (1-\rho)\frac{\partial U}{\partial y} > 0$

It follows that the right-hand side of equation 26 is unambiguously negative. Thus the extra term in the wage of the rigidity model Λ becomes more and more negative as productivity increases and the difference between wages in the benchmark and the rigidity model becomes larger. In other words, the wage-structure will be more compressed when wages are downwards rigid - whenever $U < \frac{y}{1-\rho}$.