Kiel

Working Papers

Kiel Institute
for the World Economy

Firm Training and Wage Rigidity

by Wolfgang Lechthaler

No. 1452 | October 2008

Kiel Working Paper No. 1452 | October 2008

Firm Training and Wage Rigidity*

Wolfgang Lechthaler

Abstract:

Although wage rigidity is among the most prominent subjects in modern economics, its effects on wage compression and firm training have thus far not been considered.

This paper is trying to bridge this gap by using a simple two period model which can still by analyzed analytically. I am able to show that wage rigidity increases wage compression. However, contrary to previous work this is not sufficient to increase firms' training investments. The reason lies in the endogeneity of separations, which become more frequent.

Keywords: Human Capital, Wage Rigidity, Training

JEL classification: J24, J31, M53

Kiel Institute for the World Economy
Duesternbrooker Weg 120
24105 Kiel, Germany
Telephone: +49-8814-272
wolfgang.lechthaler@ifw-kiel.de

* I would like to thank Christian Keuschnigg, Dennis Snower and Thomas Hintermaier for their valuable comments.

1 Introduction

Wage-rigidity is receiving more and more attention, especially in the business cycle and the empirical literature. As noted by Shimer (2004), the standard matching model with flexible prices is not able to replicate the large movements in unemployment and vacancies over the business cycle. Especially in the the New-Keynesian literature it has been tried to remedy this shortcoming by introducing various sorts of rigid wages. In the empirical literature the evidence for wage-rigidity comes from three different sources: Experiments, econometric studies and surveys. Broadly speaking, all of them agree that there is a considerable amount of wage-rigidity, especially when it comes to downwards movements.[1]

Given the obvious importance of wage rigidity in the literature and in real life it is kind of surprising that - to my knowledge - its effect on wage compression has thus far not been thoroughly analyzed. Acemoglu (1997) sees wage compression as the main force driving firm's investments in worker training. Acemoglu and Pischke (1999a) discuss various sources of wage compression - among them search on the labor market, asymmetric information, specific human capital, efficiency wages, minimum wages and unions - but wage rigidity is not mentioned at all.

The purpose of this paper is to bring together these different strands of the literature and analyze the effects of wage rigidity on wage compression and firm training. The first - not so surprising - result is, that wage rigidity increases wage compression. The second, very surprising result is, that this is not sufficient to increase firm training. Indeed, through the endogeneity of separations wage rigidity unambiguously decreases firm training. This result is accordance with a companion paper (Lechthaler and Snower (2006)), in which it is shown that a minimum wage has two distinct effects on training: On the one hand the wage floor increases incentives to train because all the rents go to the firm. On the other hand turnover goes down which tends to decrease training. Depending on the distribution of an idiosyncratic shock both effects can dominate.

[1]For more details see further below.

The results in this paper are much stronger. First of all, while minimum wages only affect a small group of workers and especially those with the lowest productivity, wage rigidity is much more wide-spread, as workers in any skill-class might be affected. Secondly, while the assumption of a uniform distribution in Lechthaler and Snower (2006) leads to a canceling out of both effects, in the present paper even in such a setup the turnover effect dominates and training goes down.

The work is built on the idea that wage-rigidity has important effects on the wage-structure of an economy, potentially creating or increasing wage-compression.[2] According to the influential work of Acemoglu[3] 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[4] 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.

In many aspects, my model is 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: Firstly, the way in which wages are determined and secondly 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

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

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

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

on separations. As confirmed by the empirical evidence these become more frequent as rigidity is introduced. In the models of Acemoglu and Pischke separations typically take place at an exogenous rate and therefore this effect is ruled out by assumption. [5]

The relationship between firm training and wages 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.

Thus, most of the existing empirical literature is clearly supporting the concept of Acemoglu (1997) in favor of the one by Becker (1962). However, there is one source of wage compression for which the evidence is not so clear, or rather contradicting the idea of Acemoglu. As discussed in Acemoglu and Pischke (2003), the effects of minimum wages are found to be either insignificant or even negative. Although, to my knowledge, there are no studies directly relating wage rigidity and firm training, this evidence on minimum wages clearly points towards some problems of the Acemoglu model when it comes to wage compression caused by wage floors.

The remainder of the paper is organized as follows. Since the wage setting mechanism is the main innovation of the paper it will be discussed in more detail in the following section. In section three I present the general framework of the model before I outline a benchmark model in which wages are determined without any restriction. Section five discusses the model with rigid wages while section six concludes.

[5] Acemoglu and Pischke (1998) being the only exception. However, with this model is not possible to analyze wage-rigidity since all workers receive the same wage.

2 Wage-setting

This section begins with a detailed review of the empirical evidence on wage rigidity which stems from three different sources. I will start with the econometric evidence and then proceed with surveys and experiments. Following the discussion of the empirical literature I will then introduce my own approach to model wage rigidity and compare it to others used in the theoretical literature and the empirical evidence.

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. They report that nominal wage-cuts are extremely rare: Overall 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.

Fehr and Goette (2000) do a similar exercise for two Swiss firms (large and medium-sized) in the service industry. 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 Goette 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.

Card and Hyslop (1997) use the US Panel Study of Income Dynamics. They as well find a 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. 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.[6]

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

The wage setting rule I use is closely related to the approaches introduced by Strand (2003) and Holden (2002). Holden discusses different ways of modeling wage rigidity that were used in the literature and compares them to empirical studies like the one by Fehr and Goette (2000) just described. Holden argues that so far no wage rule could explain the numerical results satisfactorily and suggests two alternatives that fit better. In both of his approaches he assumes that a drop in the wage will imply a negative productivity shock. The shock is motivated by the adverse effects of bad morale found by Fehr and Goette but also Bewley (2002), among others.

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 lead to a wage, which is too

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

low to assure that workers do not shirk, then the firm would 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 my model I assume that principally wages are determined according to Nash-bargaining [7] and thereby dependent on the output y of the worker, which can vary from one period to the other. However, a decrease in the wage from one period to the other will induce a negative productivity shock c due to the adverse effects of the wage cut on the morale of the worker. Thus, if the output of the worker goes down a bit, the firm would rather keep the wage constant in order to avoid the productivity cost c. Furthermore, I assume that c is so large that in case of very low output, the worker and the firm prefer to separate than to continue to work with a lower wage. This is were my approach deviates from the one by Holden (2002) who implicitly assumes that both parties always stay together.

Both cases are illustrated in figure 1. The variable w_0 denotes the wage of the previous period and c the productivity cost of a wage cut. The wage stays constant as long as the output of the worker y does not drop below $w_0 - c$. If the output is even lower than $w_0 - c$, then either the wage will be cut (in the Holden-model) or the parties will separate because the alternatives are better (in my model). [8].

I believe that I am on the safe side with my assumption, given the empirical fact that wage cuts are extremely rare and the finding by Bewley (1999) that managers rather fire workers than keeping them with a lower wage. It is even further confirmed by Gielen and van Ours (2006) who find that many matches are resolved even though it would be efficient to continue with a lower wage. From a modeling perspective the advantage of my assumption is, that I am still able to analytically determine the consequences of wage

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

[8] For illustrational purposes I have used unemployment benefits b as the relevant alternative of the worker although in the model to follow it is the profit constraint of the firm which becomes binding

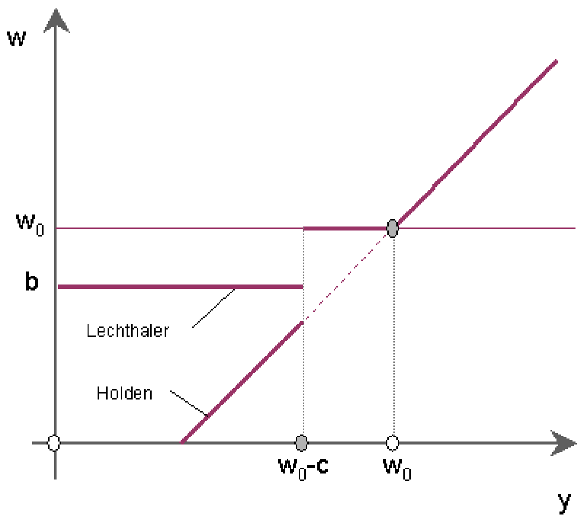

Figure 1: Wage Rigidity

rigidity for wage compression and firm training.

Especially in the business cycle 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.[9] 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.

[9] It seems that the literature is very much concerned about inefficient separations and eagerly trying to avoid them which is a bit surprising given the evidence that in practice many matches are resolved inefficiently (see for instance Bewley (1999) and Gielen and van Ours(2006)).

Danthine and Kurmann (2004a) try to incorporate rigid wages in an efficiency-wage 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[10] 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). Bewley 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 work-force and therefore they are avoided. The modeling approach I use 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

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

comparisons between my model and the related literature. Furthermore, it seems quite arbitrary to model training in an efficiency-wage model. How should this training affect the effort of the worker? To my knowledge the only work directly incorporating firm training and efficiency wages is Katsimi (2003), which creates a counterfactual negative relationship between training and wages.

3 General structure of the Model

This section describes the general structure of the model which is the same for both the benchmark model and the model including rigid wages. The structure is very similar to Acemoglu (1997) or Acemoglu and Pischke (1999a). 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, which is assumed to be frictional.[11]

At the beginning of the first period, the firm has the opportunity to train the worker, which improves her productivity $y(h)$ instantaneously. Wages are negotiated after the training decision but still at the beginning of the first period, i.e. before production takes place. This timing of events is necessary, in order to give a meaningful interpretation to the idea of wage rigidity and distinguish it from a simple minimum wage.[12]

At the end of period one the pair is hit by an idiosyncratic shock π which is assumed to be uniformly distributed with density $f(\pi) = 1/(\max - \min)$. The shock is joined additively to the productivity of the worker.[13] Due to the shock the output of period two equals:

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

[11]Assuming that the worker dies after the second period would yield the same results but imply more tedious wage functions.

[12]The output of the worker is assumed to be high enough to ensure positive wages.

[13]This is a standard way to introduce endogenous separations, see for instance Pissarides (2000).

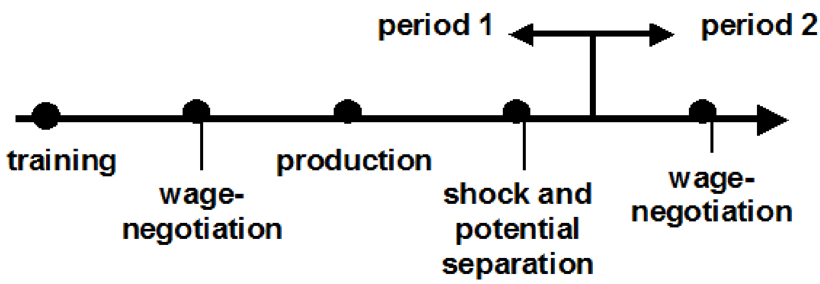

Figure 2: Time path

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 perfectly 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. The timing of important events is illustrated in figure (2).

4 Benchmark

4.1 Value functions

In this section I 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.

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. Since there are frictions on the labor market it will take some time for the worker to find another job. During search the worker receives unemployment benefits which are lower than wages. These assumptions imply that the value of unemployment is lower than the value of having a job. As in Acemoglu (1997) it is this kind of friction that creates wage compression and induces firm training.

For simplicity I assume that the value of unemployment does not depend on firm training as it would if it were truly general. However, since using general training does not affect the results but only makes the analytics more tedious, I stick to this assumption.[14]

The value of a filled job at the beginning of the first period is:[15]

$$J(y_t) = y(h) - w_t^b(h) - c(h) + \rho \int_{\Psi}^{\max} J(y_{t+1}) f(\pi) d\pi + \rho \int_{\min + y(h)}^{\Psi} V f(\pi) d\pi \qquad (2)$$

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. I will do here the latter and assume that productivity rises linearly with human capital.

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 firm's value of the second period depends on the shock π through its effect on output y_{t+1}. Ψ is the threshold-productivity: If the idiosyncratic component 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 Ψ, the match 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-value is implicitly defined by:

$$J(y(h) + \Psi) = V \qquad (3)$$

or alternatively by:

$$W(y(h) + \Psi) = U$$

[14] See the working paper version of this paper which can be provided by the author upon request.

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

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. For any shock lower than Ψ, 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:

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

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 worker in the second period. If the output lies above Ψ 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.

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:

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

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

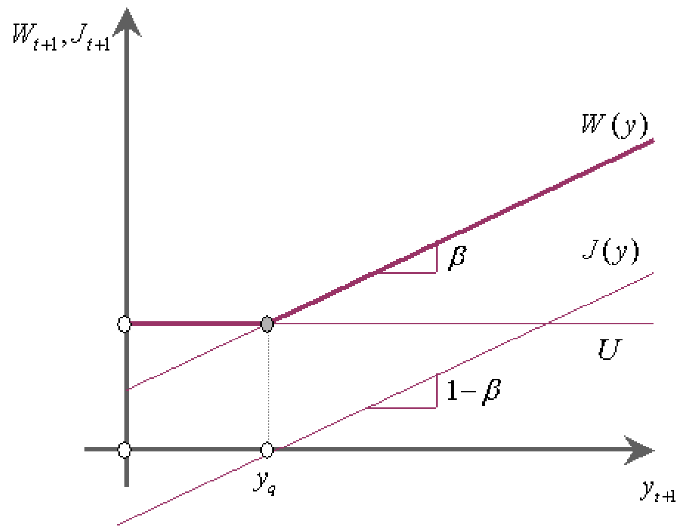

Figure 3: Value functions of the second Period - Flexible Wages

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[16] because the threat-point of both parties is independent of the idiosyncratic shock.[17] 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 Ψ. 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

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

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

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 Ψ 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.

4.2 Wages

As was mentioned above, wages are determined by Nash-bargaining.[18] Assuming that the bargaining power of workers is given by β, the negotiated wage maximizes the Nash-product:

which implies that the surplus of the match over the threat-points of both parties is shared according to their respective bargaining powers. From the perspective of the worker this means that her surplus $(W - U)$ is a share β of the rent of the whole match $(W + J - U - V)$:

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

For both periods this results in the following wage formula:[19]

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

This is a standard result: The worker gets at least the value according to her threat-point[20] plus a share β of the surplus over that threat-point.

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

[19] 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 $(1 - \rho)\, U$ would have to be replaced by $U_t - \rho U_{t+1}$.

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

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):

$$\Psi = (1 - \rho)\, U - y(h) \tag{9}$$

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.[21] 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.

4.3 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:

$$\frac{\partial w^b}{\partial h} = \beta \frac{\partial y}{\partial h} < \frac{\partial y}{\partial h} \tag{10}$$

The equation illustrates how the wage is reacting to changes in productivity. Since the worker always gets a share β of the value of production, she also gets a share β of the training's value. Thus the wage of the worker increases with her productivity but not one-to-one. In other words, the wage structure is compressed. According to Acemoglu and Pischke (1999a) this is sufficient and necessary to induce firm-sponsored training.

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

4.4 Training

Before wages are negotiated the firm decides privately about the amount of training. Using the wage equation the value function becomes:

$$J = (1-\beta)\,y_t - (1-\beta)\,(1-\rho)\,U - c(h) + \rho \int_{\Psi}^{\max} \left[(1-\beta)\,y_{t+1} - (1-\beta)\,(1-\rho)\,U\right] d\pi \quad (11)$$

The optimal decision is found by taking the derivative of this equation with respect to training and setting it equal to zero:

$$\begin{aligned} \frac{\partial J}{\partial h} &= (1-\beta)\,\frac{\partial y}{\partial h} + (1-\beta)\,\rho \int_{\Psi}^{\max} \frac{\partial y}{\partial h} d\pi - \frac{\partial \Psi}{\partial h}\left[(1-\beta)\,(y(h)+\Psi) - (1-\beta)\,(1-\rho)\,U\right] - \frac{\partial c}{\partial h} = \\ &= (1-\beta)\,\frac{\partial y}{\partial h} + (1-\beta)\,\rho \int_{\Psi}^{\max} \frac{\partial y}{\partial h} d\pi - \frac{\partial c}{\partial h} = 0 \end{aligned} \quad (12)$$

where the last step follows from the definition of the quitting threshold (see equation (9)). The third term is the marginal cost of additional training, whereas the first and the second terms illustrate the effect on profits in the current and the following period, respectively. The output increases but at the same time the wage increases and therefore the firm can only accrue a share $(1-\beta)$ of the returns to training.

Using the fact that (due to the additivity of the shock) the output y can be taken out of the integral and rearranging, we arrive at the following - more meaningful - equation, implicitly defining h_b^*, the opitmal amount of training in benchmark model:

$$\frac{\partial c}{\partial h_b^*} = (1-\beta)\,\frac{\partial y}{\partial h_b^*}\left[1 + \rho(1 - F(\Psi))\right] \quad (13)$$

where $F(\pi)$ is the cumulative distribution function (CDF) of $f(\pi)$. Now we have marginal costs on the left-hand side and marginal revenues on the right-hand side. The first two terms show the marginal revenue per period while the term in 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 Ψ. While $F(\Psi)$ is the probability of separation, the probability of staying together is $1 - F(\Psi)$.

5 Rigidity model

5.1 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 a wage-cut would imply a negative productivity shock and is therefore never realized. 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 we 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:

$$J^u(y_{t+1}) = y_{t+1} - w_{t+1} \tag{14}$$

$$J^r(y_{t+1}, w_t) = y - w_t \tag{15}$$

Again the firm-value of the second period is straight forward, it is just the value of 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:

$$W^u(y_{t+1}) = w_{t+1} + \rho U \quad (16)$$

$$W^r(w_t) = w_t + \rho U \quad (17)$$

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: [22]

$$\begin{aligned} J(y_t) &= y(h) - w_t - c(h) + \rho \int_{\Omega}^{\max} J^u(y_{t+1}) f(\pi) d\pi + \\ & \rho \int_{\Phi}^{\Omega} J^r(y_{t+1}, w_t) f(\pi) d\pi - \rho \int_{\min}^{\Phi} V f(\pi) d\pi \qquad (18) \\ W(y_t) &= w_t + \rho \int_{\Omega}^{\max} W^u(y_{t+1}) f(\pi) d\pi + \\ & \rho \int_{\Phi}^{\Omega} W^r(w_t) f(\pi) d\pi - \rho \int_{\min}^{\Phi} U f(\pi) d\pi \qquad (19) \end{aligned}$$

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 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 Ω, 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 the shock is below Ω, 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

[22] 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.

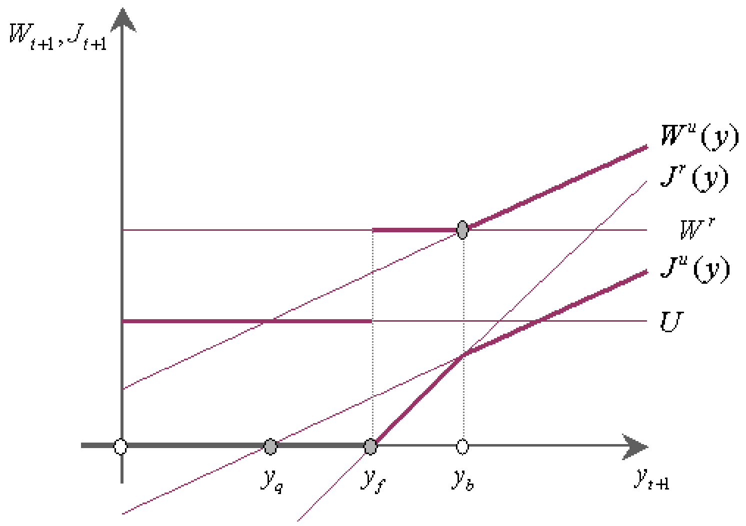

Figure 4: Value functions of the second Period - Rigidity-model

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 (here Φ), 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. 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.[23] 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 a one to one increase in firm value.

[23]To keep the graph simple, I have not drawn the cost of bad morale c which is not paid anyway.

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 the shock lies below the threshold Ω 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 Φ, which lies clearly above Ψ. 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 Φ and Ψ 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 the ensuing drop in productivity would also imply losses for the firm. The problem from the perspective of the worker is that she cannot credibly commit to provide sufficient effort to stay employed.

5.2 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: [24]

$$w_t = (1-\rho)\,U + \beta\,(y(h) - (1-\rho)\,U) + \frac{\beta\rho\int_{\Phi_t}^{\Omega_t}\pi f(\pi)d\pi}{1+\rho\,(F(\Omega)-F(\Phi))} \tag{20}$$

[24]See Appendix B for a proof.

Compared to the wage of the benchmark model (see equation (8)) there is one additional term on the right-hand side.[25] 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 $y(h) + \pi$, while the output of the first period is just $y(h)$ - thus π 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.

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: [26]

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

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

$$(1+\rho F(\Omega)-\rho F(\Phi))w_t = \beta\left[(1+\rho F(\Omega)-\rho F(\Phi))y(h)+\rho\int_{\Phi}^{\Omega}\pi f(\pi)d\pi\right]$$

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:

$$w_{t+1} = Max[(1-\rho)\,U + \beta\,(y_{t+1} - (1-\rho)\,U)\,,w_t] \tag{21}$$

Using the assumption of uniformly distributed productivity shocks it can be shown that in the rigidity model the wage structure of the first period is more compressed than in the benchmark:[27]

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

[27]For a proof see Appendix C.

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.

5.3 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:

$$\begin{aligned} W^u(y+\Omega) &= W^r(w_t) \qquad (22) \\ J^u(y+\Omega) &= J^r(y+\Omega, w_t) \\ w_{t+1}(y+\Omega) &= w_t \end{aligned}$$

where $w_{t+1}(y+\Omega)$ denotes the freely bargained wage for productivity $y+\Omega$. 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 Φ is not equal to the separation threshold of the benchmark model Ψ due to the inflexibility of wages. I will call it firing threshold because 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 Ψ. 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:

$$J^r(y + \Phi, w_t) = V \tag{23}$$

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 Φ, 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 Φ and Ψ. 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.[28] Due to this higher wage, the firm is less reluctant to fire the worker and consequently separations are more frequent. Summarizing, we can state the following about the order of thresholds:

$$\Omega > \Phi > \Psi \Longrightarrow F(\Omega) > F(\Phi) > F(\Psi)$$

Consequently, we can distinguish four different intervals for the second period. From max to Ω, wages and values are the same in both models. From Ω to Φ the wage-rigidity becomes relevant so that the wage in the benchmark is lower. From Φ to Ψ a worker is

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

fired in the rigidity model but not in the benchmark model. Below Ψ 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:

$$\Omega = \frac{\rho \int_{\Phi}^{\Omega_t} \pi g(\pi) d\pi}{1 + \rho F(\Omega) - \rho F(\Phi)} \tag{24}$$

$$\Phi = w_t - y(h) \tag{25}$$

These equations can be interpreted as follows. Because the wage adjustment[29] 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 Ω, 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.

5.4 Training

The optimal amount of training can be found by taking the derivative of equation (18), the first period value function of the firm, an setting it equal to zero. However, to be better able to compare it with the benchmark model it is useful to replace the value functions with their respective definitions (equations 15 and 14):

$$J = y_t - w_t - c(h) + \rho \int_{\Omega}^{\max} (y_{t+1} - w_{t+1})\, dy_\pi + \rho \int_{\Phi}^{\Omega} (y_{t+1} - w_t)\, d\pi \tag{26}$$

$$J = y_t - w_t(1 + \rho F(\Omega) - \rho F(\Phi)) - c(h) + \rho \int_{\Omega}^{\max} (y_{t+1} - w_{t+1})\, d\pi + \rho \int_{\Phi}^{\Omega} y_{t+1} d\pi \tag{27}$$

[29]The additional term in equation (20).

Replacing both wages with equations 21 and 20 this becomes:[30]

$$
\begin{aligned}
J &= y_t - (1 + \rho F(\Omega) - \rho F(\Phi)) (\beta y(h) - (1-\beta)(1-\rho) U) + \\
&+ \rho\beta \int_{\Phi}^{\Omega} \pi d\pi + \rho \int_{\Omega}^{\max} (y_{t+1} - w_{t+1})\, d\pi + \rho \int_{\Phi}^{\Omega} y_{t+1} d\pi - c(h) \\
J &= (1-\beta)(y_t - (1-\rho) U) - \beta\rho \int_{\Phi}^{\Omega} y_{t+1} d\pi + \rho \int_{\Phi}^{\Omega} y_{t+1} d\pi + \\
&+ \rho \int_{\Omega}^{\max} (1-\beta)\, y_{t+1} d\pi - \rho(1 - F(\Phi))(1-\beta)(1-\rho) U - c(h) \\
J &= (1-\beta)(y_t - (1-\rho) U) + \rho \int_{\Phi}^{\max} [(1-\beta)(y_{t+1} - (1-\rho) U)]\, d\pi - c(h)
\end{aligned}
$$

which looks almost exactly the same as the corresponding equation (11) of the benchmark model. In fact, the only difference is the threshold which signifies separations. This result is due to the way wages are determined when assuming Nash-bargaining. It assures that the whole surplus of the match is shared between both parties according to their respective bargaining powers and so the firm gets $(1-\beta)$ of the total expected rent of the match.

Taking the derivative of this equation with respect to training and setting it equal to zero yields:

$$
\frac{\partial c}{\partial h} = (1-\beta) \frac{y_t}{\partial h} + \rho \int_{\Phi}^{\max} (1-\beta) \frac{y_{t+1}}{\partial h} dy_{t+1} - \frac{\partial \Phi}{\partial h} [(1-\beta)\Phi - (1-\beta)(1-\rho) U]\, f(\Phi)
$$

Note that - contrary to the benchmark model - the last term does not cancel out. It can be thought of as capturing the effect of the increase in wage compression. Taking y out of the integral and collecting terms yields the optimality condition for training in the benchmark model:

$$
\frac{\partial c}{\partial h} = (1-\beta) \frac{\partial y}{\partial h} [1 + (1 - F(\Phi))] - \frac{\partial \Phi}{\partial h} [(1-\beta)\Phi - (1-\beta)(1-\rho) U]\, f(\Phi) \qquad (28)
$$

which is different from the benchmark (equation **??**) with two respects: The separation threshold and the additional, last term of the equation. As already discussed above

[30] Note that $[F(\Omega) - F(\Phi)]\, y(h) + \int_{\Phi}^{\Omega} \pi d\pi = \int_{\Phi}^{\Omega} y_{t+1} d\pi$

the firing threshold in the rigidity model is higher than the separation threshold of the benchmark, which implies that separations are more frequent in the rigidity model. This clearly tends to decrease firm training. However, the last term - representing the effect of increased wage compression - is positive and therefore tends to increase training so, as expected both effects work in opposing directions. Nevertheless, it is possible to show analytically that the turnover effect dominates the compression effect and therefore training is unambiguously lower in the rigidity model.

This can be seen by using the definition of the firing threshold Φ in equation (25) and the assumption of uniformly distributed shocks:

$$\frac{\partial c}{\partial h} = (1-\beta)\frac{\partial y}{\partial h}\left[1+(1-\frac{\Phi-\min}{\max-\min}\right]-\left(\frac{\partial w}{\partial h}-\frac{\partial y}{\partial h}\right)\left[(1-\beta)\,\Phi-(1-\beta)\,(1-\rho)\,U\right]\frac{1}{\max-\min}$$

Collecting terms including the derivative of productivity yields:

$$\begin{aligned}\frac{\partial c}{\partial h} &= (1-\beta)\frac{\partial y}{\partial h}\left[1+(1-\frac{(1-\rho)\,U-\min}{\max-\min}\right]-\left(\frac{\partial w}{\partial h}\right)\left[(1-\beta)\,\Phi-(1-\beta)\,(1-\rho)\,U\right]\frac{1}{\max-\min}\\ \frac{\partial c}{\partial h_r^*} &= (1-\beta)\frac{\partial y}{\partial h_r^*}\left[1+(1-F(\Psi)\right]-\left(\frac{\partial w}{\partial h_r^*}\right)\left[(1-\beta)\,\Phi-(1-\beta)\,(1-\rho)\,U\right]\frac{1}{\max-\min} \qquad (29)\end{aligned}$$

The only remaining difference to the corresponding equation (13) in the benchmark model is the second term on the right hand side. As the wage clearly increases with firm training, the second term is positive, thereby lowering the marginal returns to human capital and decreasing the optimal amount of training:

$$h_r^* < h_b^*$$

The difference in results to Acemoglu and Pischke (1999a) is due to the endogeneity of separations. In Acemoglu and Pischke 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.

6 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.[31] 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 modeled 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. [32]

[31] See for instance Lynch (1991).

[32] This interpretation should be taken with care: An insurance in this context does not really make

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.[33] Due to this increase in the probability of separations firm training is lower in the rigidity model.

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.

7 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

sense since workers are assumed to be risk-neutral. Nevertheless, I believe it helps understanding the functioning of wage rigidity in the model.

[33] See Bewley (2002).

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

Gielen, Anne and van Ours, Jan (2006): Why do Worker-Firm Matches Dissolve, *IZA* DP 2165

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* DP 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* WP 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

Appendix to Firm Training and Wage Rigidity

A Appendix A: wage-rule of the rigidity model

The starting point is the same rule 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 18 and 19 we get:[34]

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

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

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

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

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

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

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

Using the fact that the wage-rule equation 7 is valid in the second period as well, the unrestricted value-functions cancel out (with the exception of the value of unemployment). Plugging in the equations 15 and 17 for the remaining value-functions of the second period we get:

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

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

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

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

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

By merging the terms with U the equation simplifies to:

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

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

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 20 by dividing through the term in square brackets on the left-hand side:

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

B Appendix B: 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 20 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 let me 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:

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

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

$$\begin{aligned}&\rho\frac{\frac{\partial(y_b+y_f-2y)}{\partial y}(y_b-y_f)2[\max-\min+\rho(y_b-y_f)]}{4[\max-\min+\rho(y_b-y_f)]^2}+\\&+\rho\frac{\frac{\partial(y_b-y_f)}{\partial y}(y_b+y_f-2y)2[\max-\min]}{4[\max-\min+\rho(y_b-y_f)]^2}\end{aligned} \tag{A1}$$

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 22 and 23:

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

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

Consequently the difference between the two thresholds is:

$$y_b - y_f = (1-\beta)(y_t + \frac{\rho\int_{y_f-y_t}^{y_b-y_t}\pi_t f(\pi_t)d\pi_t}{1+\rho F(y_b)-\rho F(y_f)} - (1-\rho)U) \tag{A2}$$

while the sum of the two thresholds is given by:

$$y_b + y_f = (1+\beta)(y_t + \frac{\rho \int_{y_f - y_t}^{y_b - y_t} \pi_t f(\pi_t) d\pi_t}{1 + \rho F(y_b) - \rho F(y_f)}) + (1-\beta)(1-\rho)U \qquad \text{(A3)}$$

By taking the derivatives of equations A2 and A3 with respect to productivity y, plugging them both into equation A1 and bringing all terms with Λ to the left-hand side we get:

$$\begin{aligned} & \frac{\partial \Lambda}{\partial y}(1 - \frac{(1+\beta)\rho(y_b - y_f)}{2[\max - \min + \rho(y_b - y_f)]} + \rho \frac{-(1-\beta)(y_b + y_f - 2y)2[\max - \min]}{4[\max - \min + \rho(y_b - y_f)]^2}) \\ = \; & \rho \frac{[1+\beta-2](y_b - y_f)2[\max - \min + \rho(y_b - y_f)]}{4[\max - \min + \rho(y_b - y_f)]^2} + \\ & + \rho \frac{(1-\beta)(y_b + y_f - 2y)2[\max - \min]}{4[\max - \min + \rho(y_b - y_f)]^2} \qquad \text{(A4)} \end{aligned}$$

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 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 A3). Consequently, Λ will have the same sign as the right-hand side of the equation above, which is clearly negative since $\beta < 1$. It follows that wage rigidity unambiguously increases wage compression or put formally:

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