

Efficiency Wages,
Bargaining Power,
and Wage Gaps

© 2002 Ralph Olthoff

All rights reserved. No part of this publication may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the author.

ISBN 90-9016-543-6

Printed and bound in the Netherlands by Datawyse, Maastricht

Efficiency Wages, Bargaining Power, and Wage Gaps

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. Dr. A. C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 17 januari 2003 om 16.00 uur

door

Ralph Olthoff

Promotor:

Prof. dr. P. P. de Gijzel

Beoordelingscommissie:

Prof. dr. J. Muysken (voorzitter)

Prof. dr. dr. F. Haslinger (Universität Hannover)

Dr. A. P. van Veen

For Christine

Contents

Preface	xv
Acknowledgements	xv
1 Introduction	1
1.1 Motivation	1
1.2 Outline	4
2 Wage gaps and wage drift in the Netherlands and Germany	7
2.1 Introduction	7
2.2 Definitions and measurement problems	9
2.3 The situation in the Netherlands	11
2.3.1 Labour-market institutions	11
2.3.2 Wage drift	13
2.4 The situation in Germany	17
2.4.1 German labour relations	17
2.4.2 Wage drift	20
2.4.3 Wage gaps	23
2.5 Conclusions	26
2.A Appendix: List of abbreviations	29
2.A.1 Netherlands	29
2.A.2 Germany	29
3 Theories of the wage gap and wage drift: a survey	31
3.1 Introduction	31
3.2 A short overview	31
3.3 Market determined wages and wage gaps	35
3.3.1 The Phillips-curve theory—wage drift on a neoclassical labour market	35
3.3.2 The contract wage as an information variable on the labour market	38
3.3.3 Adaptation versus mark-up hypotheses of wage drift	44
3.4 Efficiency-wage theories and wage drift	45
3.4.1 Models of efficiency wages and the influence of the contract wage	45
3.4.2 Primary and secondary drift and the stability of the relative wage structure	48

3.5	Models of local and central wage bargaining	50
3.5.1	The basic framework	50
3.5.2	Efficiency wages in models of local and central bargaining . .	53
3.5.3	Threat points and bargaining power in local wage bargaining	59
3.6	Conclusions	60
I	Bargaining in efficiency wage models	65
4	Threat points in local wage bargaining	69
4.1	Introduction	69
4.2	The payoff function of the firm and the workers	70
4.2.1	The no-shirking condition	70
4.2.2	The problem of the firm	72
4.3	Local wage bargaining	73
4.3.1	The bargaining structure	73
4.3.2	Scenario 1	76
4.3.3	Scenario 2	80
4.3.4	Scenario 3	82
4.3.5	The decision about bargaining	83
4.4	Conclusions	84
4.A	Appendix	87
5	Bargaining over the effort function	89
5.1	Introduction	89
5.2	The effort function	93
5.3	The firm	96
5.4	The contract, wages, and employment	98
5.4.1	The contract wage	98
5.4.2	The supported effort level	101
5.4.3	The sanction probability	102
5.5	Unions and the labour market	103
5.5.1	No unions	103
5.5.2	The insider union	104
5.5.3	The local union	105
5.5.4	The central union	107
5.6	Conclusions	108
5.A	Analysis of a parameterized effort function	111
II	Surplus sharing and the wage gap	113
6	Wage gaps, labour contracts and surplus sharing	117
6.1	Introduction	117
6.2	Wage contracts and wage flexibility	118

6.3	The model	120
6.3.1	Time structure of the model	120
6.3.2	The renegotiation game	121
6.4	The outcomes of renegotiations	124
6.4.1	Complete information	124
6.4.2	Asymmetric information	128
6.5	The wage gap and wage contracts	129
6.6	Conclusions	131
6.A	Derivation of the outcome of the renegotiation game	135
6.A.1	Complete information	135
6.A.2	Asymmetric information	140
7	Wage bargaining under asymmetric information	143
7.1	Introduction	143
7.2	Assumptions	146
7.3	Nash bargaining	149
7.3.1	The effort allocation of the worker	149
7.3.2	The belief of the firm	152
7.3.3	Nash bargaining with a minimum wage	155
7.4	Fixed wages and renegotiations	159
7.5	Local wage bargaining	164
7.6	Conclusions	172
7.A	Appendix	177
7.A.1	The expected outside option and marginal product	177
7.A.2	General characteristics of the expected payoff	178
7.A.3	The expected additional payoff in the Nash bargaining scenario	186
7.A.4	The expected additional payoff in the minimum wage scenario	187
8	Conclusions	189
8.1	Summary	189
8.2	Recommendations for economic policy	191
8.3	Suggestions for further research	193
	Bibliography	195
	List of symbols	205
	Samenvatting (Dutch summary)	207
	Zusammenfassung (German summary)	211
	Curriculum Vitae	215

List of Figures

2.1	Contract wages and gross wage drift in the Netherlands, 1948–1999	14
2.2	Contract wages, and gross and net wage drift in the manufacturing and service sector in the Netherlands, 1973–1983	16
2.3	Contract wages and gross wage drift in Germany, 1960–2000	20
3.1	The scope of contract and effective wage models	32
3.2	The labour market in the neoclassical model	36
3.3	Wage-productivity curve and cost per effective unit of labour in efficiency-wage models	45
3.4	Profit-utility pairs in the Holden model	51
3.5	Profit-utility pairs in the models of Muysken and van Veen, and de Gijssel	55
3.6	Employment and effective wage in the models of Muysken and van Veen and de Gijssel	58
4.1	The profit function π and its approximation $\tilde{\pi}$	73
4.2	The time structure of the bargaining game	75
4.3	Scenario 1	76
4.4	Scenario 2	81
4.5	Scenario 3	83
5.1	The effort function $e(w)$	96
5.2	Effective effort and wage as functions of the supported effort level and the contract wage	100
6.1	Time structure of the model	121
6.2	The stages of the renegotiation game in some subperiod t	123
6.3	Contract and effective wage	126
6.4	The effective wage if outside options change over time	128
6.5	Outcomes of the renegotiation game in the $(a_W, m - a_F)$ space	130
7.1	The time structure of the model	147
7.2	The worker's effort allocation as a function of the firm's bargaining power	152
7.3	The effort allocation with the opportunity to quit	155
7.4	The worker's effort allocation in the minimum wage scenario	158

List of Tables

2.1	Contract-wage increases, gross wage drift, and their correlation coefficient in selected periods, the Netherlands	14
2.2	Contract-wage increases, gross wage drift, and their correlation coefficient in selected periods, Germany	21
2.3	Coverage of collective agreements and relative wage gaps in West Germany	24
3.1	Characteristics of theories on wage drift	63
5.1	Effects of the contract wage, the supported effort level, and the sanction probability	104
7.1	Characteristics of the different wage-setting schemes	176

PREFACE

Chapter 4 is based on Peter de Gijzel and Ralph Olthoff, "Local Wage Bargaining, Efficiency Wages, and the Wage Gap. A Game Theoretic Approach", in "Aspects of the Distribution of Income", ed. by Franz Haslinger and Oliver Stöninger-Venkatarama, Metropolis-Verlag, Marburg, 1998.

ACKNOWLEDGMENTS

Writing a thesis may be a one man's responsibility, but it can never be a one man's work alone. When working on this thesis I have benefited greatly from many people. Their support and advice have certainly improved the result a lot. Yet, they are of course not responsible for any shortcomings.

First of all, I would like to thank my supervisor, Peter de Gijzel, for his intense support and the inspiring discussions. When I started working on wage gaps I was a freshman in labour market economics. Peter gave me the room for developing my own ideas, and during the discussions we had he helped sharpening my view on the topic. Moreover, he backed me when times were grim, and he showed patience and impatience with me when appropriate. At the very time, I was never happy about his impatience, but in retrospect I have to admit that he seemed to know me better than I do.

This thesis has also benefited greatly from the expertise of Joan Muysken, Franz Haslinger, Tom van Veen, Erik de Regt, Thomas Zwick, Marco Haan and many more. Joan Muysken, Tom van Veen, and Franz Haslinger have read the entire manuscript and made valuable suggestions for improvements. When I did not incorporate all their comments it was due to a lack of time and not because I have refused their suggestions. I am indebted to Erik de Regt, Tom van Veen, and Thomas Zwick for reading and commenting on earlier versions of various chapters.

The colleagues at the department of economics in Maastricht created a friendly and pleasant atmosphere and provided a lot of assistance. I owe much to Sylvia, who kindly helped me with the red tape, and who patiently gave me lessons on Dutch written language, removing again and again the same Germanisms from my texts. Thomas taught me that in writing, providing a good intuition is indispensable and that no formula is self-explanatory. From Marco I have learned how to overcome a writer's block. I am afraid that I have not been a good pupil to all three of them.

Next to my research, my other task at Maastricht University was to give courses in economics. When studying economics at the University of Hanover in Germany, I was much impressed by the lectures of Franz Haslinger. He is a great and inspiring teacher, who has shaped my ideal of how classes should be given in general, and of a teacher's attitude in particular.

For the much needed diversion I am grateful to Arjan and Hetti, Silke and Matthias, Maren and Harald, Olli and Nicol, Frank, Christoph, to my godchildren Ralph Friedrich and Wiebke, whom I saw much too infrequently in the past years, and to many more. I am especially grateful to Rita and my father, who have lovingly

supported me in all these years.

Last but not least, I would like to thank my wife Christine for sharing my life and for the beautiful time with her. Christine has patiently assisted me by taking responsibilities from me, by backing and motivating me, and by standing my moods. Therefore, I dedicate this thesis to her.

Maastricht, November 2002

CHAPTER 1

Introduction

“Well,” says payroll-administration service ADP¹, “those 3.2 per cent [bargained wage increase] in this year tell only half of the story.” According to ADP, the effective wage increase is about twice as much, because by incidental wage increases, [...] each employer continuously strives at encouraging his personnel to stay.

Robert Giebels, “Adieu, Poldermodel”, in NRC Handelsblad, 4 June 1998. Translation by the author.

1.1 MOTIVATION

After years of moderate wage increases, Dutch unions and employers' federations seem to have lost control of wages. Even the unions' representatives express regret about incidental wage raises, yet a whole lot of employers are forced to raise wages by far more than the collectively bargained rates in order to recruit, hold, and motivate their personnel. The result is a wage gap: the wage that is effectively paid is greater than the wage level agreed upon in collective bargaining.

At first glance, this wage gap does not raise a problem. Effective-wage increases, it could be argued, are higher than was bargained just because the labour market situation and incentive considerations of employers demand it. Obviously, collective bargaining just failed to settle an efficient wage increase. This view, however, falls short of an adequate explanation of wage gaps in general. Effective wages being higher than agreed upon in collective bargaining are not only observed in times of labour scarcity, but occur in all stages of the business cycle.² Moreover, if effective wages are really only determined on the labour market or by single employers and employees on their own, this view cannot explain why collective bargaining is often held so fiercely—why it is costly and time consuming, and why strikes and lock-outs occur—other but by assuming that unions and employer federations are just promoting an exhibition fight.

¹ADP (Automatic Data Processing) is a company specialized in processing the payroll administration of (Dutch) companies.

²See chapter 2.

Yet, wage gaps matter. Recent data for Germany³ are provided by Addison, Kraft, and Wagner (1993), Bellmann and Kohaut (1995), Meyer (1994, 1995) and Schnabel (1995, 1997). Using different samples, their estimates of the share of firms that are covered by collective bargaining, but that pay an effective wage higher than the contract wage, range between 73.0% and 90.2%. The average wage gap paid by these firms is around 10% of the contract wage, but this relative gap varies widely over the different industries and skill groups.⁴

Wage gaps and their development over time, the so called wage drift,⁵ are a challenge for both economic policy and theory. Though the mentioned figures are far from being representative, they show that the wage gap is a phenomenon that can hardly be ignored in the discussion about wage formation.

The two components of the effective wage—contract wage and wage gap—can obviously be regarded as the outcomes of two different processes in wage formation at two different stages. At one stage the contract wage is set. Commonly, it is considered to be the outcome of collective bargaining between unions and employers' federations on nation wide, industry wide or just on the individual firm's level. If collective bargaining matters, then its immediate purpose is obviously to influence the *effective* remuneration, mostly for a predetermined time period.

While the determination of contract wages is highly institutionalized in European economies, this is far less the case for the second stage of wage formation, at which the effective wage is determined. It is reasonable to assume, that the factors that make effective wages exceed the contract wage, are found on external and internal labour markets. Wage gaps may be due to deliberate decisions of individual firms, but it is possible that also issues of bargaining power between management and employees play a role.

The mere existence of the two stages suggests that none can be neglected. At each stage of the wage-determination process, the actors should be aware of the mechanisms and possible outcomes of the other. However, the literature on wage formation usually ignores this interdependency. Labour-market theory provides a whole lot of models of wage formation. But apparently, there is a gap between models focusing on effective wages and those focussing on contract wages.

Among the effective wage models, the textbook neoclassical theory of a competitive labour market is obviously at pain explaining wage gaps. In a world of complete information, individual rationality and the invisible hand, there is simply no room for collective bargaining, unions and employers' federations. Many contemporary economists find this framework too restrictive, because it overlooks apparent market imperfections, as for example bargaining power and incomplete information. The recent years have seen several new theories that abandon—with the exception

³Unfortunately, corresponding data is not available for the Netherlands.

⁴Blum (1983, pp. 139–40), looking at firms in the metal industry of a South German bargaining district in the 1970s, even reports that firms in the sample on average paid a mark-up between 20% and 50%. The individual figures range from 5.9% to as much as 82.9%. For a more detailed survey of the empirical evidence see chapter 2.

⁵In most studies, wage drift is defined as the relative change of the wage gap over time, i. e., the difference between the growth rates of effective wages and contract wages.

of rational behaviour—one or more of the basic assumptions of neoclassical theory, giving room for a more realistic view on the labour market. Focusing on the effects of rent-seeking behaviour (e.g. insider-outsider theories), incomplete information (e.g. efficiency-wage theories) or implicit contracts, these theories underline that bargaining power matters in employment relations. However, they remain theories of the effective wage.

Contract-wage models, on the other hand, take the existence of bargaining power for granted, and focus on behaviour of especially the trade unions—under the restrictions of employers' preferences, and input and output markets. Yet, most of the related theories assume at least implicitly, that the outcome of bargaining, the contract wage, will become the effective wage.

Theoretical studies that try to fill the gap between effective- and contract-wage models are scarce. Moreover, they approach the wage-gap problem from various angles, such that a coherent picture of the phenomenon is still missing.⁶ The intention behind this thesis is to provide a better understanding of the factors that lead to the occurrence of wage gaps. In order to provide a more coherent account for the phenomenon, it will especially focus on the behaviour and the underlying motivations of the agents that are involved in the process of wage formation at the level of individual firms. Therefore, the thesis is centred around two themes, efficiency wages and bargaining power, as the main—and strongly related—explanatory determinants of wage gaps.

Bargaining power matters on all stages of the wage-formation process. With respect to collective bargaining this is obvious. But also on the level of individual firms bargaining power becomes important if rents are created, for example due to transaction costs or moral hazard. Then, the need to ensure that employees keep their part of the contract, forces firms into a, at least implicit, bargain with their workforce.

The choice for efficiency wages is mainly motivated by their incentive aspect.⁷ Efficiency-wage theory postulates that it can be rational for firms to pay higher than market clearing wages. The underlying idea is that there is a positive relation between a worker's performance and her wage, which firms can exploit in order to minimize the cost of labour input. This argument alone can provide for an explanation, why firms (voluntarily) pay a mark-up on the contract wage.

Yet, the firms' ability to manipulate labour supply has even further-reaching effects. The incentive considerations can imply that wages become largely independent from short-term developments on the external labour market. Employees are then partly shielded from the discipline of the market, and can acquire bargaining power vis-à-vis the firm. In this way, workers—and unions—can get a stake in the wage-setting behaviour of firms, forcing wages up even further.

⁶See the survey in chapter 3.

⁷Originally, efficiency-wage theory was hailed for its putative ability to explain involuntary unemployment without having to rely on exogenously imposed wage rigidities or irrational behaviour. However, this view has been criticized sharply (see e.g. the discussion in Franz (1994, pp. 302–304) and Lazear (1995, pp. 70–72)).

1.2 OUTLINE

Why do firms pay a mark-up on the contract wage, and what are the consequences for the different stages of wage formation? This thesis approaches the problem from different directions, focussing on issues of bargaining power and incentive pay.

The chapters 2 and 3 discuss the empirical and theoretical literature on supplementary payments. The purpose of chapter 2 is to provide an empirical background for the theoretical analyses in the subsequent chapters. It describes the labour relations in the Netherlands and Germany, and discusses empirical studies on the size and the determinants of wage gaps and their development over time in these two countries. The chapter pays special attention to the question whether these studies support our hypothesis that efficiency wages and bargaining power are important explanatory factors for wage gaps. As will turn out, no definite proof can be given, but the empirical evidence is at least consistent with our hypothesis.

Chapter 3 discusses the theoretical literature on wage gaps. It shows that the results of different theories with respect to the nature of wage gaps strongly depend on the assumed labour-market structure. The assumption of a neoclassical labour market supports the view that, apart from imposing a restriction by their minimum-wage characteristic, contract wages do not matter in the formation of effective wages, at least not in the long run. In this case, the wage gap is not a mark-up on the contract wage, but just a buffer between effective and contract wage. This approach is contrasted with models that ascribe wage gaps to labour-market imperfections. Although the kind of imperfection differs from model to model, it is shown that their results are all in favour of a mark-up hypothesis of the wage gap. The discussion in chapter 3 therefore pays special attention to efficiency-wage models and models of local wage bargaining, because especially the first part of the following chapters builds forth on these models.

The remaining chapters can be divided into two parts. Part I (chapters 4 and 5) focuses on the occurrence and characteristics of wage gaps when effective wages are determined by both efficiency wages and bargaining. Unlike many other theoretical studies on wage gaps, the models of this part avoid *ad hoc* assumptions about the motivations and the behaviour of the parties involved in the wage-formation process. Instead, special attention is paid to a stringent microeconomic motivation of wage gaps.

Chapter 4 addresses the problem of bargaining power in local wage bargaining when firms are bound by a no-shirking condition. Different from collective bargaining, the parties to local negotiations on the level of an individual firm are bound by a peace clause, which prohibits industrial actions like strikes and lockouts. This raises the problem of identifying meaningful threat points of especially the workers. In chapter 4, the process of local bargaining between employees and firm is modelled explicitly. The model of this chapter compares the outcomes of a noncooperative bargaining game under different constellations of threat points, in order to qualify the diverging results of earlier models of local wage bargaining.

Chapter 5 relaxes the restrictive assumption that unions only bargain over wages,

and introduces effort as a second object of negotiations. In contrast to other studies of bargaining over effort, it is argued that, if firms pay efficiency wages, unions will in general hesitate to bargain over effective wage and effort levels directly, in order to avoid getting involved in the moral-hazard problem between firms and workers. Rather, unions will try to establish reference levels of wages and effort, which determine the actions a firm can take against workers, while leaving the moral-hazard problem to solve for the firm. The model analyses the resulting consequences for the wage gap and the behaviour of unions. Moreover, it is argued that the reference-effort level can help solve the problem of the workers' threat point in local wage bargaining.

Part II (chapters 6 and 7) focuses on a problem that has not been raised yet in the literature about wage gaps—What are the efficiency aspects of wage-formation systems that can lead to wage gaps? It would be presumptuous to assume that different wage-formation processes have evolved for mainly economic reasons, while neglecting social and political factors. Yet, economic factors might be important for predictions about whether a certain wage-formation system will survive, or whether it will be outperformed by alternative systems. On this background, the purpose of this part is to show, that wage-setting systems that provide for a base wage rate (the contract wage) combined with the possibility to alter this wage (the wage gap), contribute to the efficiency of employment relations, and that the parties involved in wage formation might prefer them to alternative wage-setting systems.

Chapter 6 focuses on the employment relation between a firm and a single employee. It is shown that if both parties have an interest in a fixed wage rate, e.g. in order to protect their returns from specific investments, it can nevertheless be efficient to allow for a certain degree of wage flexibility. The underlying reason is that although a fixed wage avoids hold-up in the employment relation, the parties must avoid foregoing all returns of the relation if it is threatened by changing exogenous circumstances. In the chapter a (re)negotiation game is analysed that is able to reconcile the conflicting needs for wage stability on the one hand, and wage flexibility on the other hand.

A similar framework is used in chapter 7, but now the focus lies on a different type of investment than in the preceding chapter, and on the impact of asymmetric information between firm and worker. In the model of chapter 7, the worker's outside option is endogenized. The worker has the choice between investing effort in her outside option, or in her work performance in the firm. Comparing the worker's optimal investment decision under different bargaining environments, it is concluded that only a system that combines a base wage with a performance-dependent component can achieve both an efficient investment decision *and* the protection of the employment relation from inefficient separation. With respect to the bargaining environment, this system turns out to be very similar to the two-tier wage-bargaining systems that are discussed in the chapters 3.5 and 4 of this thesis.

The conclusions (chapter 8) compare the results of different models of this thesis and provide some suggestions for further research.

CHAPTER 2

Wage gaps and wage drift in the Netherlands and Germany

2.1 INTRODUCTION

The purpose of this chapter is to provide a background for the theoretical analyses in the subsequent chapters. It will show that wage gaps indeed matter in the process of wage formation; for a lot of employees they make up for a considerable part of their earnings. Second, this chapter surveys empirical studies on the determinants of wage drift and wage gaps in the Netherlands and Germany in order to show that the basic assumption of this thesis is justified, namely that both efficiency wages and bargaining power are important for understanding why firms raise wages above the level that is determined by collective bargaining.

The first part of the chapter (section 2.2) provides some definitions that are useful for the empiric quantification of wage gaps and wage drift. Furthermore, it discusses common problems in measuring these variables, which can lead to significant misrepresentation of their true values in empiric research.

The sections 2.3 and 2.4 then discuss the situations in the Netherlands and Germany. First, the institutional settings of the labour markets in the two countries are briefly described in order to clarify the structures of labour relations. The second parts of these sections then discuss the development and possible determinants of wage drift, i. e., the development of supplementary payments over time, in the two countries. The section about Germany also discusses empirical results about wage gaps, which—different from the wage drift figures—also provide an insight in the size of supplementary payments.

The empirical part focuses on two problems. The first is the relation between collectively bargained (or, contract) wages and effective wages, and the nature of wage gaps. If effective and contract wages are determined by two independent processes or if contract wages only follow the development of effective wages, wage gaps should merely serve as a buffer between the two. In this case, any theory of wage determination could simply focus on effective wages and consider contract wages only as minimum wages that sometimes can become binding. If instead the effective wages are persistently influenced by the contract wages—and not just by their minimum wage characteristic—wage gaps become an interesting theoretical problem, as

they will be an independent force in the determination of effective wages.¹

The second problem is to identify the factors that determine wage gaps. In the literature about the phenomenon, roughly three different approaches can be distinguished.² According to the market or "Phillips-curve" approach, effective wages are mainly market determined. In this view, firms are forced to pay wages above contract if the contract wage is too low to clear the labour market. This idea was originally introduced by Hansen and Rehn (1956), and it was the common explanation for wage drift in the 1950s and 1960s. In this view, contract wages would indeed only play a minor part in the determination of effective wages.

The bargaining-power approach assumes that wage gaps are due to negotiations on the level of individual firms. These negotiations can have an explicit character in the form of bargaining between the management and workers' representatives, or they can be implicit. In the latter case, either workers' representatives could bargain about other topics than wages in order to force earnings up in an indirect way,³ or no explicit negotiations take place at all, but the management raises wages just for fear of a dispute with its employees.

Finally, the efficiency-wage approach attributes wage gaps to the incentive effect of wages on labour productivity. The argument is that firms can drive up labour productivity by paying wages above the market-clearing level. As was mentioned in the first chapter, efficiency wages are an attractive explanation because incentive pay is a central theme in the management literature about compensation.

In the discussion of the empirical evidence on wage gaps and wage drift it will become clear that it is virtually impossible to discriminate definitely between the three approaches. The reason is that data that allow for a direct test of the market-power and the efficiency-wage approach are missing or even impossible to measure. It is impossible to measure bargaining power directly, because it is a theoretical *ad hoc* construct, and a direct test on efficiency wages would require individual data on the wage and the performance of workers, which are not available. Empirical research therefore has to rely on proxies. But in most cases these cannot be clearly related to one or the other approach. For instance, as we will see in section 2.4, some studies try to relate bargaining power to the structure of employment in a firm. The argument here is that especially women, part-timers and low-skilled blue-collar workers should have low bargaining power in comparison to high-skilled workers. Yet, this idea is based on the observation that the former groups are more easily replaceable. If empirical studies show that the share of these groups has indeed an effect on the size of wage gaps, it is therefore not clear whether this is not just due to market forces. With efficiency wages we face a similar problem. A greater wage gap for skilled workers could be due to the fact that the performance of employees fulfilling more complex tasks cannot be easily monitored, and that firms therefore have to use wages in order to provide incentives. It could, however, also mean that firms try to

¹These two views on the role of the wage gap are sometimes referred to as the "adaptation hypothesis" (of contract wages) and the mark-up hypothesis (of the wage gap).

²These three approaches will be discussed in detail in the following chapter 3.

³See chapter 5 of this thesis for such a model.

reduce turnover costs because it is costly to replace these workers. Finally, the reason could just be that skilled workers are relatively scarce, and that their wage gap just increases their earnings to the market-clearing level. If empirical studies show that the employment structure and labour market variables are important determinants of wage gaps and wage drift, this can consequently not be interpreted as support for the market approach, but such a finding is consistent with all the three approaches.

The reader should take these difficulties into account when the results of empirical studies are discussed in sections 2.3 and 2.4. Yet, as will be shown, there is sufficient evidence that shows that bargaining power and especially efficiency wages are important determinants for wage gaps.

2.2 DEFINITIONS AND MEASUREMENT PROBLEMS

For analysing the divergence between contract wages (usually the wage settled by collective agreements) and effective wages two concepts, the wage gap and the wage drift, are usually used. The *wage gap* is a static variable measuring the deviation of the effective wage from the contract wage at a certain point in time. The *absolute wage gap* g^a is the difference between effective wage w and the contract wage \bar{w} , measured in units of money,

$$g^a = w - \bar{w},$$

while the *relative wage gap* g^r indicates by how many per cent the effective wage exceeds the contract wage:

$$g^r = \frac{w - \bar{w}}{\bar{w}} \cdot 100$$

Wage drift describes the deviation of changes in the effective wage from the changes in the contract wage over time. Gerfin (1969, p. 474 f.) describes two measurements, d^a and d^r ,

$$d^a = \left(\frac{w_t - w_{t-1}}{w_{t-1}} - \frac{\bar{w}_t - \bar{w}_{t-1}}{\bar{w}_{t-1}} \right) \cdot 100, \quad (2.1)$$

$$d^r = \frac{d^a}{\bar{w}_t/\bar{w}_{t-1}} = \left(\frac{w_t/w_{t-1}}{\bar{w}_t/\bar{w}_{t-1}} - 1 \right) \cdot 100, \quad (2.2)$$

of which the first is more commonly used. Both measurements have in common that the wage drift is equal to zero if effective and contract wage grow by the same rates. Consequently, a negative wage drift does not imply that the wage gap is decreasing, but that it is increasing by less than the effective wage. Gerfin (1969, p. 475 f.) defends this characteristic arguing that if negative drift rates prevail over many periods, the share of above-contract payments in the wage will become more and more insignificant, even if the absolute wage gaps do not decrease.⁴

The empirical research on the wage gap and wage drift has to struggle with a number of shortcomings of the official wage statistics, especially in the determination

⁴In order to avoid the problem of negative drift rates while the absolute wage gap is still increasing, Marquand (1967, p. 8) proposes other drift measurements that describe the absolute or

of the contract wage. The reason is that the contract-wage statistics often come in the form of index numbers based on the average contractual earnings per hour or per piece. As a consequence, the official contract-wage statistics omit some relevant components of the contract wage, which, as these components do well appear in the effective-wage statistics, commonly leads to an overestimation of the wage gap and to a bias in measuring the wage drift. Gerfin (1969, p. 477 f.) summarizes the responsible factors as follows.

1. Variations in factor input, e.g. overhours and short time:

Contract-wage statistics are usually calculated from base wages per hour or month, given the normal working hours. Moreover, they ignore overtime allowances determined by collective agreements. As the effects of overhours and short time are well included in the effective-wage statistics, the measured wage gap can (wrongly) show a cyclical component, if employers vary labour input not by changing the number of employees but by changing the number of hours worked. As a consequence, wage drift might be underestimated in recessions and overestimated in booms.

2. Changes in productivity of piece-workers:

Contract-wage statistics usually convert piece rates into hourly wages by using a norm output per hour. Changes of productivity are therefore not accounted for in the contract-wage statistics, but they are captured in the statistics of the effective wage.

3. Supplementary payments determined by collective agreements:

As the contract-wage statistics are based on normal hourly earnings, they do not recognize one-off payments as holiday pay, Christmas bonuses, child benefits, etc., or other supplementary payments as for nightwork, etc.

4. Changes in the composition of the workforce:

Collective agreements usually differ from industry to industry, and each of them determines the wages of different qualification groups of workers. Because in the official statistics the average contract wage is usually calculated by using fixed weights for the different groups and industries, they do not account for workers moving from one industry or qualification group to another.

Due to these factors the contract wages reported by the official statistics are biased downwards.⁵ Because the effective-wage statistics capture the effects of these

relative increase of the absolute wage gap over time, e.g.,

$$w^d = w_t^{ga} - w_{t-1}^{ga}, \quad \text{or} \quad w^d = \frac{w_t^{ga} - w_{t-1}^{ga}}{w_{t-1}^{ga}}.$$

These measurements have in common that the wage drift is negative only if the absolute wage gap is decreasing over time. However, Marquand's proposals have not gained general acceptance.

⁵Strictly speaking, the factors 2 and 4 could lead to a positive or a negative bias. In practice however, the bias should be positive, because a downward development of productivity or downgrading of workers are rare.

factors, the wage gaps calculated from official figures are usually too high. Whether the wage-drift figures are also affected depends on how the relative changes of the reported contract wages over time compare to the changes of the real contract wages. Only if both move in line, the calculated wage drift is equal to the real drift, even if the calculated wage gap is biased.

These measurement problems lead to the distinction of a gross and a net concept of the wage gap and wage drift. The *net wage gap/drift* reflects the theoretical concept of wage gaps/drift and accounts for the four factors mentioned above. The *gross wage gap/drift* is not adjusted for these factors. The difference between the gross and the net figures, which is due to the effects of the four factors, is called the *measurement or method gap/drift*.⁶ It should be noted that most of the wage drift figures that are discussed in the subsequent sections are not corrected for the measurement drift.

2.3 THE SITUATION IN THE NETHERLANDS

2.3.1 Labour-market institutions

Dutch unions are organized along industries and *weltanschauung*, i. e., for most industries like manufacture, nutrition, transport, etc., there are two unions, one non-denominational and one (Protestant) Christian, a distinction that is reflected by the existence of two different federations of trade unions, the FNV (Federation of Dutch Unions) and the CNV (Christian National Unions' Federation).⁷ In 1997, the proportion in membership between CNV and FNV was about 1:3.5. In international comparison union membership is relatively low: in 1993 only 26.2% of the active workforce were organized.⁸ In order to compensate for the accordingly low membership contributions, Dutch unions have to rely on contributions from employers.⁹ Moreover, in the last twenty years unions have tried to increase their strength by concentration. In the past two decades, the formerly independent socialist and the Catholic federations of trade unions have merged to form the FNV in 1981. Also within the federations mergers have taken place, the last being the merger of the

⁶The Dutch Central Office for Statistics (CBS) emphasizes factor 4 and uses the term "structural effect".

⁷The Protestant "Christian" unions were originally founded as countermovement to the socialist labour movement. The non-denominational unions nowadays encompass unions that have their origins in the socialist labour movement as well as in the Catholic social movement (both groups have merged in 1981). Both groups of unions are organized in different union federations, but their attitude has become more and more pragmatic rather than ideological. Next to these two federations of trade unions there is a third, which organizes senior white-collar workers. See Albeda, Dercksen, and Tros (1998, chapter 2) for an overview of the different organisations.

⁸The figures about union membership are taken from Albeda, Dercksen, and Tros (1998, p. 56-58).

⁹As a part of the collective agreements these contributions are compulsory. According to Teulings and Hartog (1998, p. 285), "for some unions these contributions make up a sizeable proportion of the revenues." The same authors report that employers see these contributions as a means to keep the smaller Christian unions alive in order to prevent a uniting of the union movement (*loc. cit.*, p. 293).

unions of the service, the manufacturing, the transportation and food sectors to FNV Bondgenoten in 1998, uniting about 40 % of the members of FNV unions. The federations of trade unions do not participate in collective bargaining, but they attempt to coordinate and support the bargaining strategies of their member unions. Nevertheless, the division of the union movement still weakens its bargaining position. Employers can, and frequently have, played off one union against the other, typically a more moderate Christian union against its counterpart of the FNV.

Apart from collective bargaining, labour representatives are also granted a say in matters of remuneration on the level of individual firms. By law, works councils have the right to veto a number of management decisions concerning the compensation of employees, to mention decisions about the pension scheme, profit-sharing arrangements, the employer's contributions to savings schemes, and about wage schemes. However, the consent of the works council is not needed if the point in question is already determined by a collective agreement (Albeda, Dercksen, and Tros 1998, p. 164). Moreover, works councils have the right to make proposals about labour relations in the firm, on which the management must comment.

The employer federations in the Netherlands, originally organized along similar lines as the unions, are nowadays much less divided. The employers of a certain branch are mostly represented by one branch organization. Most of these organizations are in turn members of one of the three central employer organizations, LTO Nederland (agriculture)¹⁰, MKB Nederland (medium and small sized companies), and VNO-NCW (for all types of companies). So in contrast to the union movement, there is no direct competition between the different organizations. In collective bargaining, Dutch employers are represented by their branch organizations. Only a few very large companies negotiate directly with the unions. The task of the central organizations is in general to provide coordination and support to its members.

The third important player in the wage formation process is the government. It provides the legal framework of collective bargaining, and it can intervene in wage formation by ordinance. During the interbellum the legislator has created means to extend collective agreements to companies and employees that are not represented by the bargaining bodies: firms were obliged to extend a collective agreement in which they participated to all their personnel, and the Minister for Social Affairs got the right to declare agreements in an industry binding for all firms and employees of the very industry.

After the war, the Dutch government implemented a complex policy of guided wages. Advised by the Social Economic Council (SER)¹¹, the government formulated guidelines for the development of wages, which unions and employer federations had to take into account when negotiating collective agreements at the industry level. The agreement needed approval by the Board of National Arbitrators (College van

¹⁰In contrast to the other organizations, individual companies or farmers can be member of LTO Nederland. Moreover, LTO Nederland is the only central organization that participates in collective bargaining.

¹¹The members of the SER are appointed representatives of employees, employers, and the government. As an advisory board, the SER took the place of the Foundation of Labour (STAR), a body of employers' and trade union federations, in 1950.

Rijksbemiddelaars). The board decided upon approval, again advised by the SER, on the grounds of the government's guidelines.¹²

From 1959 on, the Dutch government loosened its grip on wage setting, and in 1964 it finally left the control over collective agreements to the unions and employer federations. The Wage Law (*Loonwet*) of 1970 legalized this situation, but it still gave the government far-reaching opportunities to intervene in wage formation: the government can freeze wages (article 10), and—under certain conditions—it can declare a collective agreement non-binding (article 8). The government used its right to freeze wage several times until 1981. Moreover, it stimulated national agreements between unions and employer federations in order to limit the scope of negotiations on the industry level. In general, however, the aim of moderating wage increases in order to counteract the economic downturn was not achieved. Finally, in 1982, the government relinquished its right to intervene directly in the process of wage formation, a step that was legally established in 1986. Yet, the government has reserved the competence of deciding about the extension of collective agreements to the whole industry, and it has repeatedly used this competence to put pressure on the bargaining parties.¹³

After the government has withdrawn from an active role in wage formation, the federations of trade unions and employers kept coordinating the negotiations on collective agreements through the 1980s. It was not before the last decade that the bargaining partners in the different industries started to pursue different agendas, allowing for agreements tailored to the circumstances in the industries. However, collective bargaining for each industry still takes place on the national level, although the number of company agreements is increasing.

2.3.2 *Wage drift*

Despite of the long Dutch tradition of government interventions in collective bargaining, empiric research on wage gaps and wage drift is surprisingly scarce. To the knowledge of the author, microeconomic data on wage gaps is not available. Aggregated data on the development of contract and gross wages, from which the gross wage drift can be calculated, are published on a regular base by the Dutch Central Bureau of Statistics (CBS). Moreover, (net) drift data are available for selected time periods.

Figure 2.1 depicts the development of contract wages and the gross wage drift in the Netherlands from 1948 to 1999, and table 2.1 summarizes the development of contract wages, the wage drift and their correlation for selected periods. The relatively low drift figures up to the mid-1950s suggest that the guided wage policy

¹²See also Albeda, Dercksen, and Tros (1998, p. 74 ff.) and van Veen (1997, p. 18 f.) for more comprehensive description of the guided wage policy.

¹³In general, employers might be more interested in an extension of a collective agreement than unions. If agreements are extended, employees are not forced to join the union in order to be covered. Moreover, by extending an agreement to firms which are not member of an employers' federation, the member firms can gain control over the labour costs of their competitors (Haucap, Pauly, and Wey 2000).

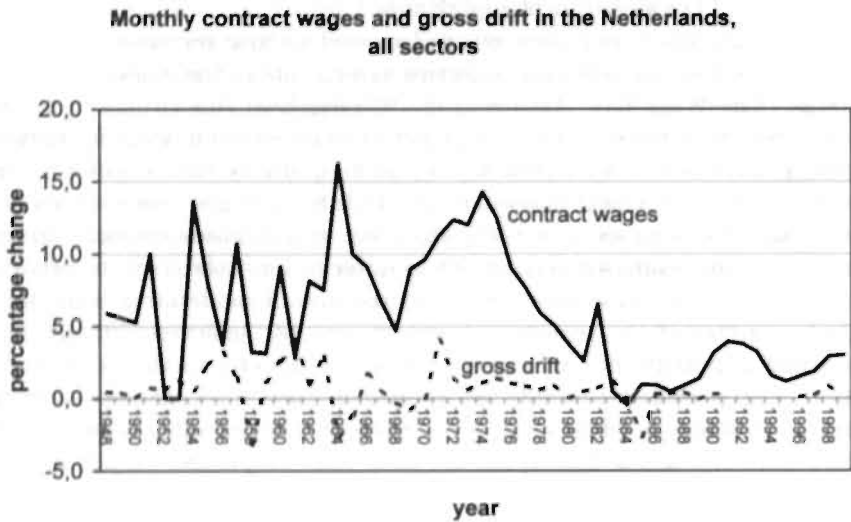


Figure 2.1: The development of contract wages and the gross wage drift in the Netherlands, 1948–1999. Sources: Centraal Bureau voor de Statistiek (1975), van der Zwan (1995), Centraal Bureau voor de Statistiek (2000), Schulte Nordholt (2001).

Period	average contract-wage increases	average gross drift	correlation coefficient
1948–1954	5.8 %	0.6 %	-0.27
1955–1963	6.2 %	1.7 %	0.14
1964–1973	10.1 %	0.4 %	-0.15
1974–1981	7.6 %	0.8 %	0.76
1982–1989	1.4 %	1.7 %	0.13
1990–1999	2.6 %	0.5 %	0.68

Table 2.1: Average yearly contract-wage increases, gross wage drift, and their correlation coefficient in selected periods, the Netherlands. Sources: Centraal Bureau voor de Statistiek (1975), van der Zwan (1995), Centraal Bureau voor de Statistiek (2000), Schulte Nordholt (2001), own calculations.

initially succeeded in controlling not only the contract wages, but also the effective wages—the development of the two do not deviate much. From 1955 to the end of the guided wage policy in 1964, however, wage drift was in average much higher. Despite of the increasing demand for labour the government attempted to keep wages low, and although the government tried to make the system more flexible by using the development of productivity in certain sectors as a criterion for wage increase, firms, especially in the construction industry, began to pay “black wages” (Albeda, Dercksen, and Tros 1998, p. 76 f.).

During the following period of strong growth, 1964–1973, the government rarely intervened in collective bargaining. Around 1969/73 the Dutch labour movement became more radical. Due to the tight labour market, unions gained more influence, and in addition individual employees started putting pressure on their employers, even by wildcat strikes (Albeda, Dercksen, and Tros 1998, p. 79), which is reflected in both high contract-wage increases and considerable drift rates. Over the whole period, however, wage drift is relatively moderate, which is partly due to the fact that right after 1964 the “black wages” were incorporated in the contract wage, which led to a great increase in the contract wage, but to a negative drift (van Veen 1997, p. 34). In view of the economic crises in the 1970s and early 1980s, both contract-wage increases and gross drift moderated over time. In 1982, unions, employers’ federations and the government agreed on pursuing moderate collective agreements in the future, in order to alleviate the unemployment problem. With no pressures from the labour market, gross wage drift was low, yet positive, too.

As was argued in section 2.2, net wage drift is typically lower than its gross counterpart. Especially if the gross figures are negative but close to zero, it could very well be that the net drift is negative. Net drift figures of the Netherlands are available for the manufacturing and service sector in the period 1973–1983.¹⁴ Figure 2.2 shows net wage drift and contrasts it with the gross patterns.¹⁵ Not only does net drift turn out to be considerably lower than the gross drift, but also their relation to the contract-wage increases is different. While the correlation coefficient between contract-wage increases and gross drift is (slightly) positive with $\rho = 0.0345$, it is negative for contract-wage increases and net drift, $\rho = -0.0981$. Some authors have concluded from this negative relation that wage drift in this period, rather than being driven by contract-wage increases, partially corrected undesired developments of contract wages.¹⁶ The diverging figures also show that statements about the relation between wage drift and contract wages that are based on gross drift figures should therefore be taken with some care.

¹⁴The Dutch Central Bureau of Statistics has also published another time-series of the net drift for the years 1978–1985 (van der Werf 1987). Unfortunately, its figures are calculated using different wage data, such that they are not comparable to the previous series in van Reijn (1982) and van Reijn (1984).

¹⁵The term “net drift” is not entirely correct. The reported net drift figures are derived from the gross drift (“incidental wage increases”), by correcting for changes in the structure of the workforce and excluding overtime payments. For the exact calculation methods see for example van Reijn (1984).

¹⁶See also Driehuis (1975), Lever (1993, chapter 5), van Veen (1997, p. 35).



Figure 2.2: The development of contract wages, and the gross and net wage drift in the manufacturing and service sector in the Netherlands, 1973–1983. Sources: van Reijn (1982), van Reijn (1984), Centraal Bureau voor de Statistiek (2000)

In the description of the development of wages in the Netherlands, we have mentioned only labour-market pressures as a determinant of wage drift. However, the core hypothesis of this thesis is that bargaining power and efficiency wages play a prominent role in the determination of mark-ups on the contract wage. To the knowledge of the author, the only study that attributes the Dutch drift patterns to efficiency wages is Lever (1993, chapter 5). His model is based on the fair wage-hypothesis of Akerlof (1982),¹⁷ but it can easily be interpreted in line with alternative efficiency-wage theories, as for example the shirking approach, that is used in this thesis (see chapters 4 and 5). According to the fair wage-hypothesis workers are willing to provide more effort the better the wage they earn compares to some “fair” wage. Under the restriction that the contract wage must not be undercut, firms determine the profit-maximizing levels of the effective wage and labour demand, taking the relation between effective wage and effort into account. The contract wage is set by a monopoly union, that maximizes its utility depending on the contract wage and employment. From the profit-maximizing behaviour of the firm an equation for wage drift is derived, while the union’s problem yields an equation for contract-wage increases. Lever assumes that the fair wage depends on a weighed average of some alternative wage and unemployment benefits, with the weights being the unemployment rate. In the regression equations the (increase of the) alternative wage

¹⁷See also chapter 3.4 of this thesis, in which a similar model of Schlicht (1992) is discussed.

is replaced by the consumer-price index, labour productivity and a trend variable.

Lever used, next to other aggregated data of the Dutch economy, the described net drift data of the manufacturing and service sector to estimate the parameters of the equations for the contract wage and wage drift. The effect of the labour-market situation is as one would expect: both the contract wage-increases and wage drift depend negatively on the change in the rate of unemployment and positively on the change of the vacancy rate. The result that seems to support—at least weakly—the efficiency-wage hypothesis concerns the effects of contract-wage increases and wage drift on each other: while wage drift is fully anticipated in contract-wage increases—its coefficient is not significantly different from -1 —, an increase in the contract wage is only partially offset by a correspondingly lower wage drift. According to the estimation results, a one per cent increase in the contract wage only leads to a 0.7 per cent decrease in wage drift. Although Lever has found a negative correlation between contract-wage increases and wage drift in the period under observation, wage drift was obviously not only a buffer used by firms to absorb undesired contract-wage changes.

2.4 THE SITUATION IN GERMANY

2.4.1 *German labour relations*

In contrast to the intricate history of Dutch post-war labour relations, the situation in (West) Germany was much steadier. In order to avoid the factionalism which had weakened the union movement during the Republic of Weimar, German labour representatives strived at the formation of a strong, unified union movement, organized along industries. In 1949, the DGB (German Trade Union Federation) was founded as an umbrella organization for the different industry unions, which had formed up after 1945. However, a comprehensive unification could not be achieved due to dissents about the political orientation and the particular interests of white-collar workers and civil servants. In the 1950s, some unionists became dissatisfied by the DGB's partisanship for the Social Democratic Party, and ultimately founded Christian unions, which—together with some other smaller unions—formed up the CBG (Christian Trade Union Movement of Germany) in 1959.¹⁸ In contrast to the Dutch situation, the Christian unions are relatively insignificant, and German employers can seldom play off DGB and CBG unions against each other. In 1992, the proportion in membership between CBG and DGB unions was about 1:35 (Franz 1994, p. 229).¹⁹ However, the existence of two large professional trade unions outside the DGB, the DAG (German Trade Union of Salaried Employees) for white-collar

¹⁸In contrast to the DGB, which strictly supports the principle of industrial unions, the CBG also accepts professional associations as members.

¹⁹The relation between DGB and CBG unions ranges from cooperation over ignorance to antagonism. CBG unionists are often not represented in collective negotiations, sometimes they negotiate together with their counterparts from the DGB union, and in some cases they are the exclusive representatives of the employees in collective bargaining. The latter, however, occurs only in niches, where the representation of DGB unions is low (e.g. handicraft in some rural regions, or negotiations with small and medium sized firms which are not members of an employers' federation).

workers and the DBB (German Federation of Civil Servants) for civil servants, has contributed to frequent disagreements in the union movement.

Although the degree of organization is still greater than in the Netherlands, German unions have experienced a decline in membership, too. In 1990, the gross degree of organization was about 35 % (Franz 1994, p. 231). As a consequence, some DGB member-unions have merged in recent years, reducing the number of member unions from 16 in 1988 to only eight in 2001. The last merger, in 2001, finally achieved the integration of the DAG in the DGB: the DAG merged with four DGB unions to form VERDI (Unified Service Trade Union).

The position of the DGB as the dominant federation of trade unions is much weaker than that of its Dutch counterparts. The DGB represents its member unions at the government and in the public, and it coordinates the policies of its member organization. However, especially the larger member unions resist against attempts of the DGB to intervene in matters of collective bargaining.

In general, German works councils have similar rights and duties as their Dutch counterparts. The Law on Regulations Governing Industrial Relations (*Betriebsverfassungsgesetz*), § 77.3, however, prohibits works councils from negotiating wages with the management, unless this is explicitly allowed for by a collective agreement. Yet, according to Teschner (1977, p. 9), this ban is occasionally flouted, especially when the works councils are in a favourable bargaining position. Moreover, § 87.1 of the same law extends codetermination of works councils to the wage structure and principles of compensation, and the introduction, change and application of compensation methods. In addition, works councils have far-reaching competences in matters of personnel and employment.

The organization structure of German employers is two-dimensional. First, employers are represented by industry-specific organizations, and second, they are represented by local organizations that represent the employers of different industries in a certain region or *land*. Both types of organization are in turn member of the central employers' federation BDA (Federation of German Employers). Individual firms cannot be members. For historical and legal reasons, the BDA does not cover the employers of the public service and the coal and steel industry. The functions of the BDA are mainly representative. In matters of collective bargaining its member organizations act more or less independently.

The degree of centralization in collective bargaining in Germany is—by its formal aspects—considerably lower than in the Netherlands. The negotiations for a certain industry usually do not take place on the national level, but on the level of “bargaining regions”, which sometimes correspond to the territory of a *land*, but can also be considerably smaller. Moreover, there are special company agreements for very large firms. Yet in practice, there is a considerable degree of coordination between the different regions and industries. Usually, a year's bargaining round is opened by negotiations between the unions and employers of the metal-processing and electrical industry in a certain region.²⁰ The resulting agreement is then used

²⁰Often, the forerunner is the “bargaining region” Nordwürttemberg-Nordbaden in South-West Germany.

as a benchmark for negotiations in other regions and industries, and their outcomes will in general not differ much. Accordingly, the German metal-workers' union (IG Metall), the dominating union in the DGB, does not base its wage claims on the development of productivity in the metal-processing industry, but on its development in the whole economy.

Same as its Dutch counterpart, the German government provides the legal framework for collective bargaining, but it has almost no opportunities to intervene directly in the process of wage formation. The legal bases of collective bargaining are the German constitution of 1949 and the Law on Collective Agreements (Tarifvertragsgesetz) of 1949 and 1969. The constitution, article 9.3, guarantees the autonomy of collective bargaining by allowing employees and employers to form coalitions, and to negotiate without interference of the government. The Law on Collective Agreements, however, gives the government a role in the extension of collective agreements to non-members of the negotiating union and employer federation in a certain industry. The Federal Minister for Social Affairs can extend an agreement, if the employers bound by the agreement employ more than half of the employees in the area and industry, for which the agreement has been negotiated, and if a council consisting of representatives of employers and employees (the so called *Tarifausschuß*), agrees. The significance of these extensions seems to be low. Franz (1994, p. 225) reports that out of the about 36,000 collective agreements that were valid in 1992, only 525 were extended. In 1989, about 4.5 million employees in West Germany were covered by an extended agreement, but only one million of these actually benefited from the extension because their employer was not directly bound by collective agreements. Consequently, the government could not use its competence to extend agreements to put pressure on the bargaining parties. In spite of the relatively low numbers of extensions, most workers benefit from collective agreements because firms themselves usually extend the agreements to employees who are not member of a union, and because many labour contracts explicitly refer to collective agreements.

Though the government has to respect the autonomy of collective bargaining, it has repeatedly tried to tie unions and employers into (not binding) agreements about wage policy and employment. With the Law on Economic Growth and Stability (*Stabilitätsgesetz*) of 1967 the government institutionalized a consultative body of unions, employers' federations and government, the Concerted Action. However, especially the unions refused to accept any wage guidelines, and in 1977 the Concerted Action was finally dissolved. Two similar initiatives in 1995/96 and in 1999, the "pacts for labour", had respectively seem to have, the same fate. In all three cases, the unions initially agreed to moderate their wage claims, but reversed this policy after a year or two.

Despite its apparently weak position to intervene directly in the process of wage formation, the German government, as any other government, can still exert an indirect influence on collective bargaining by shaping the socio-economic framework. For instance, the generosity of the welfare state and monetary policy will determine how far especially the unions can shift the burden of "exaggerated" wage increases

on to the whole society.

2.4.2 Wage drift

The development of contract-wage changes and gross wage drift in Germany, 1961–2000, is depicted in figure 2.3. In general, both the increase of contract wages as well as the wage drift show a cyclical pattern; during recessions, both seem to be lower than in other years. Gross drift has been positive in West Germany throughout the 1960s, while in the following three decades it has been fluctuating around zero. The figure shows no clear relation between contract-wage changes and wage drift, although it seems that from the mid-1970s on, and especially in the 1980s, there has been a negative correlation between the two variables. As table 2.2 shows, there was a positive, yet weak correlation between the contract-wage increases and gross drift between 1961 and 1970, while for the subsequent years the correlation coefficient was indeed negative. The constantly negative gross drift for the united Germany in the 1990s can partly be explained by a negative measurement drift. After the reunification boom had ended in 1992, the number of overhours decreased. One reason for this development was a lower demand for labour, the other, that more and more collective agreements in the 1990s provided for more flexible working hours, which allows employers increasing the hours worked without having to pay overtime allowances (Deutsche Bundesbank 2000). Another reason is the particular development in East Germany, where many employers left their employers' federation in order to avoid to be bound by collective agreements.

Monthly contract and effective earnings in Germany, all sectors

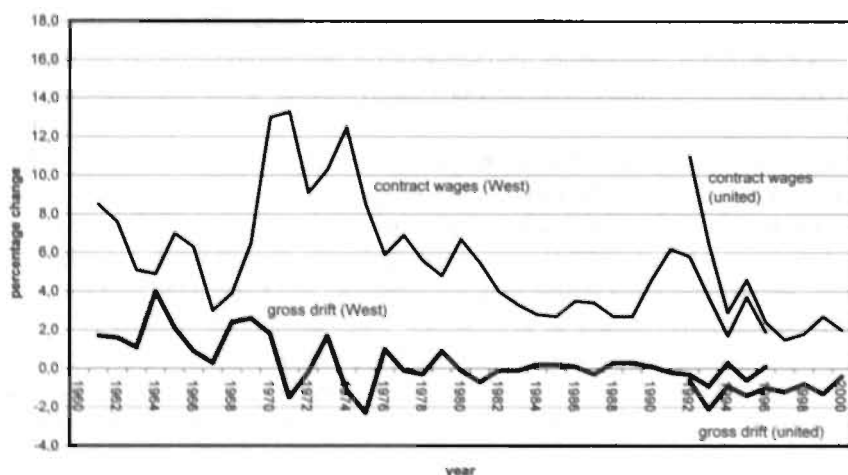


Figure 2.3: The development of contract wages and the gross wage drift in Germany, 1960–2000. Source: Deutsche Bundesbank (various volumes), own calculations.

Period	average contract-wage increases	average gross drift	correlation coefficient
1961–1970	6.6 %	1.9 %	0.04
1970–1979	8.5 %	–0.2 %	–0.43
1980–1989	3.7 %	0.0 %	–0.65
1990–1996 (West)	3.9 %	–0.2 %	–0.33
1992–2000 (united)	3.9 %	–1.1 %	–0.09

Table 2.2: Average yearly contract-wage increases, gross wage drift, and their correlation coefficient in selected periods, Germany. Source: Deutsche Bundesbank (various volumes), own calculations.

The relation between contract and effective wages in Germany was examined in more detail by Möller (1990), Kleierl (1993), and Schnabel (1995, 1997). Kleierl, using aggregated quarterly data of contract- and effective-wage increases in West Germany in the period 1961–1992, finds a structural change in their relationship in the second half of 1970. Before and after, cointegration analyses of both variables show stable, yet different long-term relations: up to 1970, the contract-wage increases are influenced by the increases in the effective wage, while in the time after 1970 the increases in both wages are interdependent (Kleierl 1993, p. 97 ff.). Möller (1990) and Schnabel (1995, 1997), using different data sets, report similar results for the latter period.

The influence of effective wages on the contract wage and the positive drift rates in the 1960s suggest that in this period the collective agreements were merely following the development of the effective wages.²¹ This behaviour is not yet fully understood. In the 1960s, it was widely believed that effective wages were driven by excess demand for labour (Hansen and Rehn 1956). Indeed, Gerfin (1969, p. 488 ff.) as well as Kleierl (1993, p. 120 f.) find a negative relation between wage drift and (changes of) the unemployment rate, while their findings about the influence of productivity increases on the wage drift are mixed. In any case, it is unclear why the unions should have failed to incorporate the effective-wage increases in the contract wages. Anecdotal evidence illustrates this problem. Gerfin (1969, p. 514 f.), for example, reports that in the period 1957–1966 the inter-industry differentials of the effective wages were decreasing by more than the differentials of the contract wages. Apparently, the German unions were ineffective even in one of their *affaires de cœur*.²²

²¹ See also chapter 3.3, which discusses this “adaptation hypothesis” on theoretical grounds.

²² An additional explanation for the perpetuated positive wage drift, that is given in the literature of the 1960s, is the distinction between primary and secondary drift (Phelps Brown 1962). Although this hypothesis was never empirically tested it is of some interest because it can be interpreted as an early efficiency-wage argument. According to this theory, an initial (primary) wage increase, that disturbs the relative wage structure, evokes a secondary drift in order to restore the old structure. (For a more detailed discussion of this hypothesis see chapter 3.4.2 of this thesis.) Secondary drift

One reason for the different behaviour of wage drift after 1970 might be the events in the late 1960s: in the aftermath of the recession 1966/67 the government tied the unions in the Concerted Action and committed them to moderating their wage claims. However, after the economy had recovered the unions failed to adapt their claims, which—assisted by the political unrest of those days—led to the wildcat-“September” strikes in 1969. Frightened to lose the support of their members, the unions abandoned the policy of wage moderation and began to pursue a more aggressive wage policy. In the early 1970s, contract wage increased sometimes by more than 10%. That these increases coincide with the first negative drift rates might hint at an incorporation of wage gaps in the contract wages.²³ Although the increases were considerably lower after the mid-1970s and throughout the subsequent two decades, the negative correlation between contract-wage increases and wage drift suggests that the unions have changed their behaviour permanently.²⁴

The described relation between contract and effective wages in Germany are in principle consistent with the efficiency-wage hypothesis. If firms pay efficiency wages, the effective wages will—if at all—depend only indirectly on the contract wages, unless the contract wages exceed the efficiency wage. If in the 1960s the contract-wage increases lagged behind the effective-wage increases, the unions might indeed not have been able to raise the contract wage above the efficiency wage. After 1970, the unions showed a greater militancy, and the drift patterns might indicate that the contract wages exceeded the efficiency wages in some years, such that employers were forced to reduce supplementary payments. This might explain why the drift figures fluctuate around zero and why the causal relation between contract and efficiency wages becomes indeterminate.

It should be noted that consistency with the efficiency-wage hypothesis does not mean that efficiency wages are indeed the reason for the described wage drift patterns. Without knowledge about the variable “workers’ effort”, however, any attempt to test directly for efficiency wages is on principle futile. Indirect tests, however, are possible. Gahlen and Ramser (1987), for instance, studied the determinants of labour productivity in four different industries in Germany from 1960 to 1982, and found a positive influence of the wage drift and the unemployment rate. Yet, such results are open to other interpretations. At best they can support, but they can never be a proof of the efficiency-wage hypothesis. The following section, discussing studies on the wage gap in Germany, will provide more indications for efficiency wages and their significance for above-contract payments.

might explain why the contract wages did not catch up with the effective wages. However, it cannot explain why the contract-wage increases were influenced by the increases of the effective wage, and not the other way around.

²³Negative drift rates just mean that the wage gap increases by less than the contract wage. However, according to Kleierl (1993, p. 99) the absolute wage gap decreased in the early 1970s.

²⁴Rather than being due to wage moderation, the lower contract-wage increases after 1974 seem to be due to the changed monetary policy of the German central bank, which—unwilling to correct “exaggerated” wage increases by a higher rate of inflation—began to announce strict targets for money supply.

2.4.3 Wage gaps

Most of the recent theoretical studies on above-contract payments, including the models discussed in this thesis, deal with wage gaps rather than wage drift. However, most of the empirical evidence is highly aggregated data about wage drift. Evidence on wage gaps is scarce, because meaningful data can only be ascertained on a disaggregated, in the ideal case individual level. Most evidence is therefore anecdotal, and only in the past decade serious attempts have been made to fill the gap.

The German Federal Agency for Statistics (Statistisches Bundesamt) stopped publishing data on earnings structures that allowed a comparison between contract and effective earnings by 1962. For the lack of alternative material with a similarly broad coverage, its last data set (Statistisches Bundesamt 1962) was used for the subsequent three decades in several studies about wage gaps, as in von der Decken (1964), Robak (1978), Kleinhückelskoten and Spaetling (1980), and Schnabel (1995). According to the latter, in 1962 the effective earnings exceeded the contract wages (including all contractual supplementary payments) in average by 14.0% (men) and 10.5% (women). He reports furthermore that the statistics show that the relative wage gaps did not differ much between earnings groups within the industries, but that they varied a lot between industries: for male workers from 0.9% in the brown-coal mining industry to 54.3% in the catering trade. He suggests that these differences were not related to excess demand for labour, the wage share or labour productivity. On the other hand, both von der Decken (1964) and Robak (1978, p. 91 ff.) find a negative relation between the average wage gap in the different industries and the levels of the collectively bargained wages. Robak (*loc. cit.*, p. 98) moreover reports that the wage gaps were positively related to the share of piece-workers and to the share of employees in large companies.

Teschner (1977), analysing a sample of earnings data from 17 companies in four different industries in 1970, also finds that the wage gaps vary considerably between the industries. The average relative wage gaps range from 0% in the tobacco industry, over 16.5% (textile industry) and 19.6% (chemical industry) to 24.2% in the metal-processing industry. The average gap paid by firms is negatively related to the regional unemployment rate. They do not seem to be influenced by the employment structure of the firms, yet skilled and piece-workers enjoy above-average supplement payments.

The latter result is confirmed by Blum (1983) and Schnabel (1995).²⁵ Blum analyses official earnings statistics and data provided by the employers' federation of the metal-processing industry in the *land* Baden-Württemberg from 1960 to 1979. For the period 1972–1979 he reports relative wage gaps of 8.9–32.5% for workers paid hourly rates and 27.7–51.5% for piece-workers (Blum 1983, p. 241). According to the data he presents, the rank order of earnings groups with respect to their wage gap is surprisingly stable over time; at least for the period 1968–1979 the wage gaps are positively related to the skill level. Schnabel (1995) comes to the same result

²⁵ Again in Schnabel (1997, chapter 6).

by analysing earnings data from 1993, which were provided by an anonymous West-German employers' federation. His study also confirms that wage gaps are greater in large firms, and he did not find a significant influence of the regional unemployment rate on the size of wage gaps.

In the last decade, two attempts have been made to ascertain disaggregated earnings data on a more representative basis: the 1989/90 company panel "Producing in Baden-Württemberg and Lower Saxony", in 1991 extended to the *land* of North Rhine-Westphalia, (henceforth "company panel") and the 1993 "IAB-Company Panel" (henceforth "IAB panel"), covering all West Germany. Both panels rely on interviews of companies. Next to the usual questions about the employment and earnings structures, the questionnaires also asked about relative wage gaps and for the motivations behind the wages policies of companies.²⁶ A drawback of both panels is that they do not allow for a direct comparison between contract wages and wage gaps, as the interviewed firms were only asked to estimate the average relative wage gap in their company.

Panel Sample	Company panel "Producing ..."		IAB-Company Panel		
	Lower Saxony and Baden-Württemberg	North Rhine-Westphalia	Lower Saxony and Baden-Württemberg	North Rhine-Westphalia	West Germany
Year	1989	1991	1993	1993	1993
Number of firms	103	175	465	378	4 265
Firms covered by collective agreements	76.7 %	88.0 %	82.4 %	87.1 %	72.0 %
Out of these pay above contract	83.6 %	90.2 %	73.0 %	73.2 %	57.0 %
Relative wage gap (a)	11.7 %	11.0 %	9.3 %	8.8 %	7.6 %
Relative wage gap (b)	<i>13.8 %</i>	<i>12.0 %</i>	<i>12.7 %</i>	<i>12.0 %</i>	13.4 %

Table 2.3: Coverage of collective agreements and relative wage gaps in West Germany. Source: Meyer (1995), Bellmann and Kohaut (1995), own calculations. The relative wage gap (a) is an unweighted average over all firms covered by a collective agreement, the gap (b) includes only firms that reported to pay above contract. The figures in italics are calculated by the author on the basis of the other figures in the table.

The relevant figures for the relative wage gaps in the two panels are given in ta-

²⁶For a description of the company panel see Schmidt (1991, 1992), for the IAB-panel Dostal, Hadler, Kühl, Lahner, Ulrich, Wendler, and Wolfsteiner (1994). The former panel included the questions whether the management believed in the incentive effect of above-contract wages and wage differentiation (Meyer 1995), while the latter asked, whether the firms paid seniority wages (Bellmann and Kohaut 1995).

ble 2.3. Apparently, the reported average gaps are similar to those of the last official statistics about earnings structures from 1962 (see above). Meyer (1995) moreover concludes on the basis of the company panel, that also the inter-industry differentials in wage gaps have not changed much since that time.²⁷ Another remarkable result of the panels is the relationship between the size of a firm and the wage gap. While the company panel does not allow to draw clear conclusions (Meyer 1995), the figures of the IAB panel suggest that, contrary to the common believe, the relative wage gap is decreasing with the firms' size for all but the largest firms (Bellmann and Kohaut 1995).

Both panels have been used to examine the determinants of wage gaps: the company panel by Addison, Kraft, and Wagner (1993) and Meyer (1994, 1995), and the IAB-panel by Bellmann and Kohaut (1995). Both Meyer, and Bellmann and Kohaut try to discriminate between the possible theoretical explanations for wage gaps—market forces, efficiency wages and bargaining power—, while Addison, Kraft, and Wagner are especially interested in the influence of the companies' employment structure as an indicator of the bargaining power of employees vis-à-vis the management. Although all studies use different (sub)samples,²⁸ their results are similar.

Addison, Kraft, and Wagner find that relative wage gaps are negatively related to the share of women and blue-collar workers in the workforce, but positively related to the share of skilled workers. Wage gaps are not related to a rent-sharing variable, and the existence of a works council even has a significantly negative effect. These results are in essence confirmed by the study of Bellmann and Kohaut. They did not test for the effect of the share of skilled workers, but found an additional negative effect of the share of part-time workers among the employees. Meyer (1995) gets mixed results for the influence of works councils, and also in his study, rent sharing does not play a significant role.

In the two studies of Meyer and in the study of Bellmann and Kohaut wage gaps are positively related to the number of vacancies per employee. These authors also test for the significance of efficiency wages. Meyer (1994, 1995) includes a dummy in his regression for companies that reported they believe that the payment above the collectively bargained wage and/or the increase in earnings differentials have a positive impact on workers' performance. He indeed finds a significantly positive influence of the dummy on the wage gaps. Bellmann and Kohaut instead use an indirect approach and include a dummy for companies that increase wages by seniority, which is commonly seen as an incentive mechanism that is alternative to efficiency wages.²⁹ If efficiency wages are a determinant for wage gaps, they argue, than seniority wages should have a negative impact on the supplement payments.

²⁷ Although this suggests a high degree of stability of wage gaps, it could also be coincidence. At the end, almost nothing is known about the developments of wage gaps in the thirty years between the two observations.

²⁸ Addison, Kraft, and Wagner (1993) use a subsample of the 1989 company panel, Meyer (1994) the whole 1989 sample, Meyer (1995) the 1989 sample plus the 1991 sample from North Rhine-Westphalia, and Bellmann and Kohaut (1995) the IAB data for West Germany.

²⁹ See e.g. Lazear (1979, 1981).

Their findings about the influence of the related dummy are, however, mixed and in any case insignificant.

Meyer's (1994) result indicates that efficiency wages play a role in the determination of wage gaps. The results of the other studies with respect to the influence of the employment structure and the number of vacancies are at least consistent with the efficiency-wage hypothesis, but as mentioned in the introduction, they can also be interpreted as evidence for the alternative hypotheses. Schnabel's (1995) finding, that wage gaps are not significantly influenced by the regional unemployment rate (see page 24), might support the efficiency-wage and the bargaining approach against the market approach. Unfortunately, it is not clear, whether his result is representative.

The reported negative impact of the presence of a works council in a firm on the wage gap is of course most peculiar, as it contradicts any intuition. Meyer (1995) points out that in the company panel only 13.4% of the firms let the works council participate actively in the wage-formation process, but this does not explain its negative impact.³⁰ One possible explanation might be related to Bellmann and Kohaut's (1995) finding that wage gaps are higher both in very small and very large firms. The reason for their negative impact could then be that in smaller firms works councils are less frequent, and that they are even absent in very small firms. Another possible reason is that works councils are more frequent in industries where unions are strongly represented and where the contract wages are comparably high. In this respect, it is to regret that both panels do not allow to relate the wage gaps to collectively bargained wages.

Lacking information about contract wages is a general deficit of the two panels. In order to compensate, the studies include different regional and industry dummies in their regressions. In the study of Meyer, the regional dummies show to be significant, while Bellmann and Kohaut report significant influences of the industry dummies. These differences are most likely due to different delimitations of the regions and industries and the much smaller sample size of the company panel used by Meyer. Although these dummies will capture the effect of contract wages, they do not allow identifying the effect of contract wages on the wage gap. Moreover, these dummies will also conceal the effects of other variables, e.g. the regional unemployment rate, which are possible determinants of wage gaps.

2.5 CONCLUSIONS

The previous sections have discussed the labour relations and empirical evidence on wage gaps and wage drift in the Netherlands and Germany. In the institutional aspects of wage formation both countries nowadays show remarkable correspondences, especially as far as the legal competences of the players in the labour market—unions, employers, government and works councils—are concerned. In practice, however, there are still differences in the way these players act and in their relations.

³⁰The reported 13,4% is already a surprisingly high number if one takes into account that German works councils officially have no say in wage formation.

In both countries, the law now guarantees the autonomy of unions and employers' federations in collective bargaining. But what is a tradition in Germany has evolved in the Netherlands only in the past two or three decades. Before, the Dutch government intervened frequently in collective bargaining, and nowadays it sometimes still uses the right to decide about the extension of a collective agreement to the whole industry as a threat to the bargaining parties. In Germany, the extension of agreements is far less important, and the government, if it wants to influence the outcome of negotiations, can only rely on moral suasion—in general with only little success.

When comparing the union movement in both countries the institutions look similar, but their significance again is different. In contrast to the German union movement, which is dominated by the block of DGB unions, the Dutch union movement is still split along ideological backgrounds. Moreover, union membership is much lower in the Netherlands, and Dutch unions are financially not that well equipped as their German counterparts. Due to these factors, Dutch unions have a weaker bargaining position, which should *ceteris paribus* be reflected by lower contract-wage increases.³¹

While the different labour relations might have an impact on collective bargaining, there is no reason to believe that the factors determining wage formation on the level of individual firms should be much different. On the formal level, the competences of Dutch and German works councils are similar, and informal aspects as implicit bargaining power, the labour market and efficiency wages should play a similar role in both countries. Unfortunately, there is no data about wage gaps in the Netherlands that would allow a comparison with the German situation.

The two problems that had to be answered by the empirical part of this chapter were whether wage gaps really matter in wage formation, and what the determinants of wage gaps are. The survey on studies of wage gaps in Germany perhaps provides the best insights in these problems. It has shown that wage gaps, depending on the considered industry and earnings group, can make up for a considerable share in the earnings of employees. Moreover, wage gaps are no temporary phenomenon; it seems that the size of and the differentials in wage gaps have been stable over long periods of time. As to the determinants of wage gaps, the results are not that clear. At least the results of Meyer (1994, 1995) suggest that efficiency wages are an important determinant for wage gaps. The effects of other variables, except of the negative influence of works councils, are at least consistent with the efficiency-wage and the bargaining hypothesis.

The position of the efficiency-wage and the bargaining hypothesis would be strengthened if it could be shown that, as Schnabel (1995) suggests, the regional unemployment rate has indeed no influence on the size of wage gaps. That the discussion of wage drift has shown that in both the Netherlands and Germany wage drift

³¹It is the *ceteris-paribus* condition which makes a comparison of the aggregated Dutch and German figures (see figures 2.1 on page 14 and 2.3 on page 20) futile. Not only the inflation rates in both countries were different, but also the labour-market situation, the industry structure, and probably the agendas of the unions.

depends on *changes* in the unemployment rate is then no contradiction. The two results combined—the level of unemployment has no effect but its change has—could instead provide support against the market hypothesis. Such a result would, for example, be consistent with the insider-outsider theory, which attributes bargaining power of the employees of a firm vis-à-vis the management to the cost a firm incurs when replacing an employee.³² The bargaining position of the employees is then only threatened by a decrease in labour demand, while the level of unemployment has no influence.

Apart from that, the discussion of studies on wage drift provided only few additional insights in the nature of wage gaps. For the Netherlands, the study of Lever must be emphasized, because it shows that the Dutch drift patterns are consistent with an efficiency-wage hypothesis of wage gaps. The German studies on the relationship between contract and effective wages have shown, that different from the situation in the 1960s, the development of the effective wage did not have a causal effect on the contract wages in the subsequent two decades. That wage drift has been fluctuating around zero in both countries in the past three decades is in line with the finding that the size and the structure of wage gaps did not change much over this time. Obviously, there are persistent forces that drive up the effective wage above the contract wage.

³²Cf. Solow (1985), Blanchard and Summers (1986), and Lindbeck and Snower (1988).

2.A APPENDIX: LIST OF ABBREVIATIONS

2.A.1 Netherlands

- CBS** Centraal Bureau voor de Statistiek (Central Bureau of Statistics)
- CNV** Christelijk Nationaal Vakverbond (Christian National Trade Union Federation)
- FNV** Federatie Nederlandse Vakbeweging (Federation of Dutch Trade Unions)
- LTO-Nederland** Land- en Tuinbouw Nederland (Dutch Employers' Federation of Agriculture and Market Gardening)
- MKB-Nederland** Midden- en Kleinbedrijf Nederland (Dutch Employers' Federation of Medium and Small Companies)
- SER** Sociaal-Economische Raad (Social Economic Council)
- STAR** Stichting van de Arbeid (Foundation of Labour)
- VNO-NCW** Verbond van Nederlandse Ondernemingen-Nederlands Christelijk Werkgeversbond (Federation of Dutch Companies-Dutch Christian Employers' Federation)

2.A.2 Germany

- BDA** Bundesvereinigung der Deutschen Arbeitgeber (Federation of German Employers)
- CBG** Christlicher Gewerkschaftsbund Deutschlands (Christian Trade Union Movement of Germany)
- DAG** Deutsche Angestelltengewerkschaft (German Trade Union for Salaried Employees)
- DBB** Deutscher Beamtenbund (German Federation of Civil Servants)
- DGB** Deutscher Gewerkschaftsbund (German Trade Union Federation)
- IAB** Institut für Arbeitsmarkt- und Berufsforschung (Research Institute of the German Labour Agency)
- IG Metall** Industriegewerkschaft Metall (Union of the Metal-processing Industry)
- VERDI** Vereinigte Dienstleistungsgewerkschaft (Unified Service Trade Union)

CHAPTER 3

Theories of the wage gap and wage drift: A survey

3.1 INTRODUCTION

Unlike other fields of economic research, there is no standard textbook model of the wage gap and wage drift, from which a discussion could depart. The (scarce) theoretical literature in this field provides the most diverse explanations for wage gaps, some of them complementary, others competing, but virtually never referring to each other.

This chapter has two purposes. First, it attempts to structure the approaches to explain wage gaps/drift by providing a meaningful framework in which the diverse ideas can be categorized. Second, it shall motivate the choice to focus on efficiency-wage arguments and (local) wage bargaining as important explanations for the occurrence of wage gaps.

The chapter first provides a short overview of the scopes of different theories in order to place them into context. Then a more detailed survey follows, that is structured along the underlying mechanisms, institutions, and agents that are assumed to be responsible for the occurrence of wage gaps. The final section summarizes the implications of the different theories and relates them to the models provided in this thesis.

3.2 A SHORT OVERVIEW

Most models of wage formation can be roughly divided along their focus into contract- and effective-wage models. The scopes of these two strands of theory are depicted in figure 3.1. Contract-wage models focus on collective bargaining between unions and employers' federations, while effective-wage models focus on the relation between a firm and its employees, searching for the factors that determine the (effective) wage paid by the firm mainly on the labour market and in the power structure of employees and management in the firm. Usually, the two spheres are kept strictly separate. Models that focus on the contract wage implicitly assume that the contract wage will automatically become the effective wage, and contract-wage models search for the determinants of the effective wage everywhere but in collective bargaining. If the

models account for a difference between contract and effective wage, they attempt to explain the occurrence of wage gaps mainly from their point of view.

Adherents of contract-wage models consider the wage gap to be a mark-up on the contract wage. They bypass the effective-wage model, assuming that the mark-up is mainly determined institutionally.¹ This view was open to attack because empirical studies showed that the influence of economic variables on the wage gap is significant.²

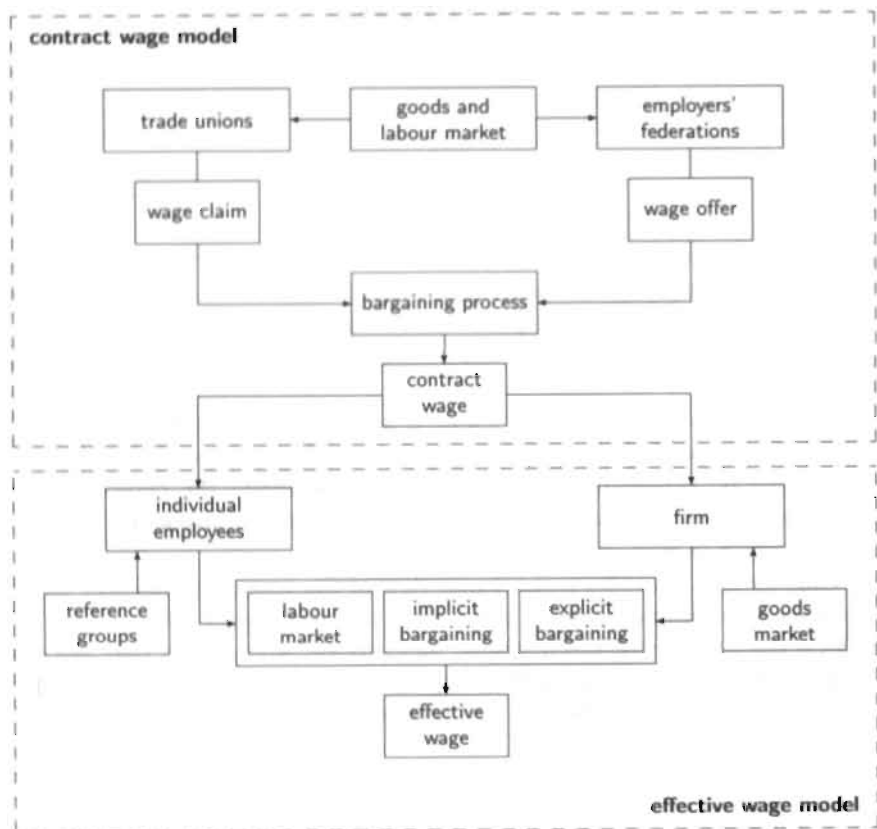


Figure 3.1: The scope of contract and effective-wage models.
After Kleinhüekelskoten and Spaetling (1980).

In effective-wage models emphasis is put on the formation of effective wages,

¹For example by the degree of centralization and the frequency of bargaining or by effective wage clauses. See Kleinhüekelskoten and Spaetling (1980, pp. 182 and 184.).

²In most empirical studies, as for example by Gerfin (1969), Gould (1967), Isachsen (1977), Söderström and Uddén-Jondal (1982), Holden (1989), and Holden (1998), the wage gap is typically assigned to—amongst other determinants—variables as indicators of labour demand, labour productivity, output prices and, closely related to the latter, profits.

whereas the contract wage commonly enters these models as an exogenous variable. Beginning with the first serious analysis of wage drift by Hansen and Rehn (1956), effective-wage models dominated the discussion. Hansen and Rehn's seminal work immediately opened the door for further research, but though at that time a remarkably high number of different possible explanations for wage gaps and wage drift were given—some of them taken up just recently—the view took shape that wage gaps just bring effective (relative) wages into a line with market determined wages.³ According to this view, the outcome of central wage bargaining is irrelevant for the effective wage structure: if contract wages increase by more than effective wages, the result will be a reduction of the wage gap.⁴ As Pehkonen and Viskari (1994, p. 397), state, wage drift thus nullifies the decisions taken at the central level of bargaining. Therefore, any income policy which is directed at stopping the rise of labour costs just by influencing collective wage bargaining is *a priori* condemned to be senseless.⁵

The common assumption about the role of collective bargaining was that the bargaining parties were just adopting the situation on the labour market, following the development of the effective wage as good as possible when setting the contract wage. According to this adaptation hypothesis, contract wages can never be adjusted perfectly to effective wages.⁶ A wage gap is therefore inevitable.

It is not surprising that this so called Phillips-curve approach was prevailing in the 1950s and 1960s, when the labour markets in most Western European countries were characterized by excess demand for labour. Consequently, this strand of research has lost much of its attractiveness since the 1970s, when these economies first faced persistent, high-level unemployment, which against all expectations did not eliminate the wage gaps.⁷ Moreover, the observed persistence of the wage gaps was in conflict with the theoretical prediction that the gaps should be a short-term phenomenon, which tend to eliminate their own cause.

Although these arguments are partially refuted by the observation that especially high skilled workers receive supplement payments, and that institutional obstacles might prevent the contract wage from fully catching up to the effective wage level, Phillips-curve like models certainly failed to explain why unions and employers' federations generally struggle hard in central wage bargaining. Unless this behaviour is beyond economic reasoning, the contract wage must have an influence in the determination of the effective wage. Even more, there is no reason why unions and employers should not influence *all* levels of wage formation. However, if this is the case there should be a mutual dependency between the setting of the contract wage

³Note that "market determined wages" does not necessarily mean competitive wages, though just this is frequently assumed in the early literature about wage drift.

⁴See Phelps Brown (1962, pp. 339 and 352) and Pehkonen and Viskari (1994, p. 397–98).

⁵Cf. Phelps Brown (1962, p. 339), Gerfin (1969).

⁶The reasons are that contract wages are bargained on a rather central level and cannot take the situations of individual firms into account, that the contracts are not differentiated enough, and that during their period in force there might be pressure to adapt effective payments (interbargain drift).

⁷See also chapter 2.4.3 of this thesis.

on the one hand and the effective wage on the other hand.

These arguments have been partially taken into account by a number of studies in the last twenty years. Incorporating new theories of wage setting, such as efficiency wages, search theories and models of implicit or explicit bargaining, these models still focus on the formation of effective wages. Nevertheless, they consider the contract wage as an important explanatory variable.⁸ In bargaining theories of the wage gap, which have arisen in the late 80s, the effective wage is determined by a two-tier wage bargaining process. First, central wage bargaining on the contract wage takes place between unions and employers federations. Second, on a local level, unions and employers negotiate on a mark-up on the contract wage. Since unions act on both levels of wage bargaining, more importance is attached to central wage bargaining. Then, from an extreme view, "wage drift is just an institutional and fully-anticipated element of a centralized wage setting system." (Pehkonen and Viskari 1994, p. 397)

Half way between these theories and the Phillips-curve approach there is a number of models which attempt to extend or combine existing wage formation theories. Many authors who consider excess demand for labour as the main reason for the occurrence of wage gaps also mention efficiency-wage arguments—even before this theory was formalized in the 1970s and 1980s. Furthermore, also advocates of the Phillips-curve theory introduced aspects of bargaining in their theories.⁹ A theory which adequately takes the mutual dependency between contract and effective wage into account is still to be missed, however.

The following sections present the different theories in more detail. Instead of taking a historic perspective, an approach is chosen, which avoids the usual simplification that contrasts the "oldish" Phillips-curve theories with "modern" theories and mostly keeps quiet about various alternative explanations that have already been given in older works. Instead, the theories are ordered by underlying mechanisms, institutions and agents which are assumed to be responsible for the occurrence of wage gaps. It will be shown that the simplistic view that effective-wage models are generally in favour of the adaptation hypothesis whereas contract-wage models support the mark-up hypothesis of wage gaps, is not true at all. Doing so, it is chosen for a rather eclectic approach. In order to avoid to skim superficially over as much models as possible, selected models which appear to be typical for their class will be discussed in detail in order to make the underlying mechanisms, which lead to the occurrence of wage gaps, clear.

Section 3.3 discusses wage-drift models in which effective wages are market determined on a more or less competitive labour market. First, the basic idea of the Phillips-curve approach is presented. The model of Holmlund (1986) then shows which effects this wage setting behaviour might have on the formation of the contract wage. Kleinhückelskoten and Spaetling (1980) finally abolish the strong assumption of a deterministic neoclassical labour market and look at the effects of a contract

⁸In these models the contract wage has an impact in different ways. For example in the model of Kleinhückelskoten and Spaetling (1980) it influences the agents' expectations on a stochastic labour market, whereas Schlicht (1992) assumes that it influences workers' fairness perception.

⁹See for example the paper of Lerner and Marquand (1962) on workshop bargaining.

wage providing information about the distribution of effective wages. With its proximity to labour-turnover models, their model also provides a link to section 3.4, which deals with the implications of efficiency wages for the wage gap. After giving a concise overview of different efficiency-wage arguments the model of Schlicht (1992) is described, where the contract wage has an impact on efforts by influencing workers' perception whether the effective wage is "fair." Section 3.5 provides an overview of models which ascribe the wage gap to explicit negotiations on a local level between employers and workers' representatives.

3.3 MARKET DETERMINED WAGES AND WAGE GAPS

3.3.1 *The Phillips-curve theory—wage drift on a neoclassical labour market*

A striking feature of most of the literature that attributes wage drift mainly to an excess demand for labour is the complete lack of theoretic modelling. In contrast, emphasis was put on empirical studies using wage drift equations which have been formulated *ad hoc*. Pehkonen and Viskari (1994, pp. 398–99) describe the common proceeding as follows:

"First, wage drift is related to the disequilibrium of the labour market, usually proxied by the aggregate unemployment rate or the vacancy rate. This relationship, which is occasionally modelled explicitly, forms the basis of the empirical analysis. Second, the relationship between wage drift and the disequilibrium in the labour market is augmented by variables which are assumed to capture other influences on wage drift, including measures of productivity growth. Third, wage drift and contract-wage increases, the latter determined by collective agreements, are treated as independent processes: in general, no explicit relationship between these two components of wages is postulated, although it is, however, often argued that the relationship should be an inverse one."

The underlying theory relating wage gaps to the state of the labour market is rather simple. It can be traced back to Hansen and Rehn (1956, pp. 92–94), who used a simple neoclassical model of a competitive labour market, depicted in figure 3.2. Labour demand is—in the short run—fixed at n_0 . Given the demand curve D_0 , the labour market is in equilibrium at point A with a wage rate of w_0 . Now an upward shift of the demand curve to D_1 (for example due to an increase in the product price or productivity) creates an excess demand Δn for labour at the given wage rate. In order to bring the market back to equilibrium, the wage rate has to rise to w_1 .

If the contract wage is fixed at the equilibrium wage w_0 at the beginning of a bargaining period, and the labour demand curve shifts upwards from D_0 to D_1 before a new contract wage is bargained, then a wage gap of Δw occurs.

Whether the resulting wage gap remains or even increases if the demand curve shifts further depends on how far the contract wage is aligned with the effective

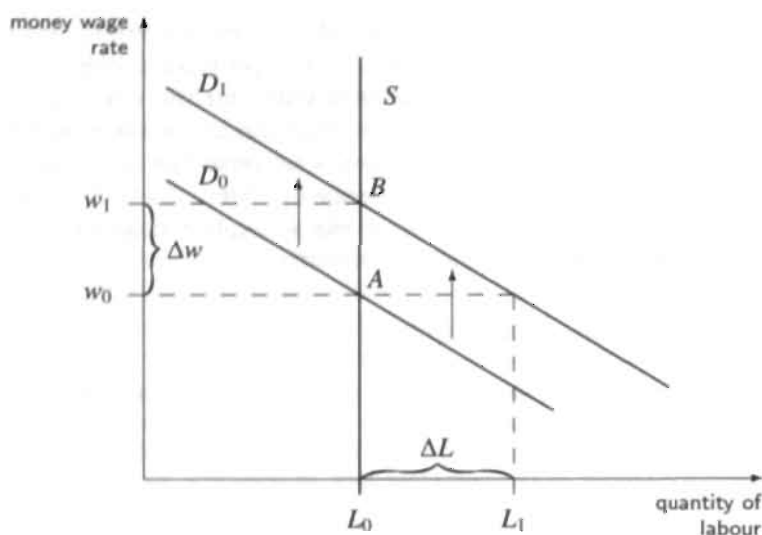


Figure 3.2: The labour market in the neoclassical model with an inelastic supply curve S and a normal-shaped demand curve D . Starting from the full employment equilibrium A an upward shift of the demand curve will result in a new equilibrium B characterized by a higher wage rate. After Hansen and Rehn (1956, p. 93).

wage by the next wage negotiations. If the new contract consolidates the increases of the effective wage totally, wage gaps are just a short-run phenomenon—they lead to the so called interbargain drift.¹⁰ Persistent wage gaps can only occur in case the contract wage does not catch up with the effective wage.

The assumption of interbargain drift gives rise to the question why the development of the effective wage is not anticipated when the contract wage is set. According to Holmlund (1986), even in the case of rational expectations a wage gap can occur due to uncertainty about the future state of the economy. In his model¹¹ an industry faces two possible states of a nature—the “good” state occurring with probability p and the “bad” state with $1 - p$. Only these probabilities and not the actual state of the economy are known by the union when determining the contract wage.

The industry’s labour demand is given by the linear function

$$L_i = \alpha_0 - \alpha_1 w_i + z_i, \quad i = 1, 2$$

where α_0 and α_1 are positive parameters and $z_1 > z_2$ represent the state of the economy. In the good state a wage gap occurs due to excess demand for labour, and

¹⁰See for example Gerfin (1969, p. 502).

¹¹At this place only a sketched version of the model is presented. In his work Holmlund assumes that workers who do not find employment in the industry are partially absorbed by the public sector.

the effective wage consists of the contract wage and an extra payment, $w_1 = \bar{w} + \Delta w$. For the bad state, however, it is assumed that no excess demand for labour can occur and that the effective wage just equals the contract wage, $w_2 = \bar{w}$.¹²

Labour supply is assumed to be fixed at N . In the good state the wage gap is determined such that a fraction $\lambda \in [0, 1]$ of excess demand at the contract wage \bar{w} is eliminated:¹³

$$\Delta w = \frac{\lambda}{\alpha_1} (L_1(\bar{w}) - N) \quad (3.1)$$

The (central) union is assumed to have full control over the contract wage. It maximizes the expected income of the workers,¹⁴

$$U = pN(\bar{w} + \Delta w) + (1-p)[L_2\bar{w} + (N - L_2)a], \quad (3.2)$$

where a is the unemployment benefit. The utility maximizing contract wage is

$$\bar{w} = \frac{1}{2} \left(a + \frac{\alpha_0 + z_2}{\alpha_1} + \frac{p}{1-p} \frac{N(1-\lambda)}{\alpha_1} \right). \quad (3.3)$$

The contract wage thus consists of a fixed base rate depending on the unemployment benefits and the parameter values of the bad state of the industry, and a second part, that is determined by the probability of the good state and the velocity at which excess demand is eliminated. Not surprisingly, the contract wage is set the higher the higher the probability of the good state is. If only little of the excess demand for labour is eliminated within one period, the contract wage is also set higher, because the union cannot expect a high extra payment by the drift process. On the other hand, if λ is close to one, the second term vanishes and the contract wage is set on the level of the base rate. In this case, the union can expect wage drift to raise the effective wage to the market clearing level.

By equations (3.1) and (3.3) the effective wage is given by

$$w = \begin{cases} \bar{w} + \lambda \left(\frac{\alpha_0 + z_1 - N}{\alpha_1} - \bar{w} \right) & \text{in the good state} \\ \bar{w} & \text{in the bad state} \end{cases} \quad (3.4)$$

As long as $\lambda < 1$, the effective wage is in both states increasing with the contract wage. If however in the good state excess demand for labour is perfectly eliminated, $\lambda = 1$, the effective wage does not depend on the contract wage. The wage gap, the second term in the equation for the good state, shows a negative relation to the contract wage.

Under the assumption that the economy is in the good state, an expression for wage drift can be derived from equation (3.4):

$$\hat{w}_1 - \hat{\bar{w}} = -\lambda \frac{\alpha_0 - N + z_1}{\alpha_1 w_1} \hat{\bar{w}}, \quad (3.5)$$

¹²This imposes a restriction on the contract wage. Since in the bad state the working force must be greater than or equal to labour demand, $N \geq L_2$, it must hold $\bar{w} \geq \frac{1}{\alpha_1} (\alpha_0 + z_2 - N)$.

¹³That is, effective labour demand is equal to $L_1(w) = (1-\lambda)L_1(\bar{w}) + \lambda N$.

¹⁴Holmlund maximizes the expected utility of the workers. At this place, we assume that $u(w) = w$.

where $\hat{x} = \frac{\dot{x}}{x}$ is the growth rate of variable x . The equation shows the expected inverse relationship between wage drift and an increase in the contract wage.¹⁵

It is worth mentioning that these results hinge on the assumption of a utilitarian utility function. If the central union did not care about employment,¹⁶ the union could set the contract wage above the equilibrium wage in the good state, thus accepting unemployment in that case. This possibility is not taken into consideration in the setting of the utility function, anyway.

Though in most empirical studies a correlation between variations in wage drift and labour shortage has been found,¹⁷ the Phillips-curve thesis has unattractive theoretical characteristics. It suggests that the relationship between wage drift and centrally negotiated wage increases should be inverse, such that in the long run total earnings are not affected by the setting of the contract wage. In Holmlund's model this argument even holds in the case that wage drift is anticipated in central negotiations. With the union caring about its possibly unemployed members, the influence of the contract wage on the effective wage is, at least in the good state, the smaller, the faster excess demand for labour is eliminated, i.e., the closer λ is to unity. In the extreme case of $\lambda = 1$, the effective wage is fully determined by the market forces. Consequently, accepting this view would mean to assign only a minor part to unions in the determination of total earnings.

Furthermore, the Phillips-curve theory implies that wage drift must be a short term phenomenon, because it tends to eliminate its own cause, excess demand for labour. Hence, in order to be able to explain the permanent drift observed in most European countries one would have to assume the labour market to be constantly in disequilibrium.

3.3.2 *The contract wage as an information variable on the labour market*

In the Phillips-curve models, that were described in the last section, the effective wage was—at least in the long run—purely determined by the market forces in a neoclassical labour market. The contract wage mainly served as a constraint on wage formation, and only in case it was set higher than the market equilibrium wage, it determined the effective wage. This relationship can be expressed by

$$w = \max \{w^*, \bar{w}\}, \quad (3.6)$$

where w^* is the unrestricted market equilibrium wage. Obviously, this relationship might be important for the wage formation when the assumption of fully informed

¹⁵The drift equation is just given as an illustration. Of course, in the static model that is presented here, neither the contract wage nor the effective wage can really grow.

¹⁶It is, for example, argued by Oswald (1985) and Weitzman (1987) that when employees are laid off by seniority, the union will not care about employment as long as its median member is not afraid to be affected.

¹⁷See for example Hansen and Rehn (1956), Gerfin (1969) and Isachsen (1977) and the discussion in chapter 2 of this thesis. It is interesting to note, however, that in a study of Pehkonen and Viskari (1994), who investigated the empirical performance of Phillips-curve and bargaining models of wage drift on data of the Finnish metal industry, the latter outperformed the first.

agents on the labour market is relaxed. Because the contract wage truncates the distribution of equilibrium wages, the expected value of the effective wage,

$$E[w|w \geq \bar{w}] = \frac{\int_{\bar{w}}^{\infty} w^* f(w^*) dw^*}{\int_{\bar{w}}^{\infty} f(w^*) dw^*},$$

depends positively on the contract wage:

$$\frac{dE[w|w \geq \bar{w}]}{d\bar{w}} = f(\bar{w}) \frac{E[w|w \geq \bar{w}] - \bar{w}}{\int_{\bar{w}}^{\infty} f(w^*) dw^*} \geq 0$$

where $f(w^*)$ is the probability density function of the unrestricted equilibrium wage. Just by this relationship the contract wage influences the average effective wage positively. Moreover, by influencing the agents' expectations, the contract wage might persistently affect the formation of effective wages.

Kleinhückelskoten and Spaetling (1980) present a model in which wages are determined by search processes on both sides of the labour market. Workers are searching for the job which provides the highest possible labour income, while employers on the one hand try to fill vacancies, and on the other hand want to prevent their employees from starting to search for a new job.

Consider first the problem of a worker who has decided to search for a new job. For each job offer that the worker receives she will compare the income from taking this job with the expected income from continuing the search, which depends on her income during search, the arrival rate and distribution of job offers, and on her search costs. Her problem is therefore to find an optimal rule for stopping the search; she has to find a reservation wage which maximizes her expected income.¹⁸

Since the expected income will depend on the distribution of wage offers and it is common knowledge that the contract wage from collective bargaining \bar{w} is the legal lower boundary of all wage offers, the level of \bar{w} will influence the reservation wage and the effective wage rate via the expectation formation process. Because of this relationship Kleinhückelskoten and Spaetling argue that the contract wage might also influence the distribution of wage offers itself, because firms have to take the searching behaviour of the workers into account when making a job offer.

The expected value of an acceptable wage offer is therefore given by

$$E[w_t|w_t \geq \bar{w} \geq \bar{w}] = \frac{\int_{\bar{w}}^{\infty} w f(w, \bar{w}) dw}{\int_{\bar{w}}^{\infty} f(w, \bar{w}) dw}, \quad (3.7)$$

where \bar{w} is the reservation wage, and $f(w, \bar{w})$ the probability density function of the effective wage given the contract wage, with $F(w) = \int_0^w f(w, \bar{w}) = 1$ and $\frac{dF(w)}{d\bar{w}} = \int_0^{\infty} \frac{\partial f(w, \bar{w})}{\partial \bar{w}} = 0$.¹⁹ Because of the minimum wage characteristic of the

¹⁸The described sequential search strategy goes back to McCall (1970) and Mortensen (1970).

¹⁹The assumption that workers already know that the firms will optimally adapt their wage offers to changes of the contract wage is in fact an implicit equilibrium assumption. Workers are aware of the optimal strategy of the firms, which will be analysed on pages 41 f.

contract wage we can assume that

$$\frac{\partial E[w_t | w_t \geq \bar{w}]}{\partial \bar{w}_t} \geq 0.$$

Assume that the worker receives one job offer in each period and that job duration is given by $z = \text{const.}$ ²⁰ Neglecting that workers might discount future income, the expected labour income when accepting a job offer is given by

$$E[y_t^a] = zE[w_t | w_t \geq \bar{w}]. \quad (3.8)$$

If the worker rejects the job offer, she can expect an income of

$$E[y_t^r] = (z - E[s])E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}] - cE[s] + vE[s], \quad (3.9)$$

where s is the expected number of further search periods, c are search costs per period, and v is present income per period during job search.

The expected income for a period t is given by

$$E[y_t] = \phi_t E[y_t^a] + (1 - \phi_t) E[y_t^r], \quad (3.10)$$

where ϕ_t is the probability that w_t is accepted in period t . Two cases have to be distinguished:

$$\phi_t = \begin{cases} \int_{\bar{w}_t}^{\infty} f(w, \bar{w}) dw & \text{if } \bar{w}_t > \bar{w} \\ \int_{\bar{w}_t}^{\infty} f(w, \bar{w}) dw & \text{if } \bar{w}_t \leq \bar{w} \end{cases} \quad (3.11)$$

Assume for simplicity that job offers in different periods are independently and identically distributed. Furthermore, let s be geometrically distributed with the probability density function $f(s) = \phi(1 - \phi)^{s-1}$ and expected value $E[s] = \frac{1}{\phi}$. If $\bar{w}_t \geq \bar{w}$, expected search duration decreases when the contract wage increases:

$$\left. \frac{\partial E[s]}{\partial \bar{w}_t} \right|_{\bar{w}_t > \bar{w}_t} = - \frac{\int_{\bar{w}_t}^{\infty} \frac{\partial f(w, \bar{w})}{\partial \bar{w}} dw}{\left[\int_{\bar{w}_t}^{\infty} f(w, \bar{w}_t) dw \right]^2} < 0 \quad (3.12)$$

If in contrast $\bar{w}_t \leq \bar{w}$, the effect is indeterminate:

$$\left. \frac{\partial E[s]}{\partial \bar{w}_t} \right|_{\bar{w}_t \leq \bar{w}_t} = \frac{f(w, \bar{w}_t) - \int_{\bar{w}_t}^{\infty} \frac{\partial f(w, \bar{w}_t)}{\partial \bar{w}_t} dw}{\left[\int_{\bar{w}_t}^{\infty} f(w, \bar{w}_t) dw \right]^2} \geq 0 \quad (3.13)$$

Kleinhüeckelskoten and Spaetling argue that these two cases might represent different phases of the business cycle. If in a beginning economic downturn relatively high contract wages are negotiated, they might exceed the reservation wage of searching workers. Furthermore, employers might not be willing to raise their wage offers.

²⁰The subsequent equation (3.9) however suggests that z is better interpreted as the worker's time horizon.

In this case, the distribution of effective wages is independent of the contract wage and the left-hand side of equation (3.13) is strictly positive—an increase in the contract wage will increase the duration of job search. If in contrast relatively low contract wages are negotiated in economic upturns, the distribution of job offers might change, and search time will decrease according to equation (3.12).

Confining to the case where the reservation wage is higher than the contract wage, equation (3.10) can be written as

$$\begin{aligned} E[y_t] &= z \int_{\bar{w}_t}^{\infty} w f(w, \bar{w}) dw \\ &+ [(z - E[s]) E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}] + (v - c) E[T^*]] \\ &\times \int_0^{\bar{w}} f(w, \bar{w}) dw. \end{aligned}$$

The optimal search strategy determines the reservation wage such that $E[y_t]$ is maximized. The first order condition of this maximization problem is

$$z \bar{w}_t = E[y_t^*]. \quad (3.14)$$

The income from accepting the job offer has to be equal to the expected income from rejecting the offer. Rearranging and expanding equation (3.14) yields

$$\bar{w}_t = \frac{E[s]}{z} (v - c) + \left(1 - \frac{E[s]}{z}\right) E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}_{t+E[s]}]$$

for the reservation wage. It depends positively on the contract wage:²¹

$$\begin{aligned} \frac{\partial \bar{w}_t}{\partial \bar{w}_t} &= \frac{v - c - E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}_{t+E[s]}]}{z} \frac{\partial E[s]}{\partial \bar{w}_t} \\ &+ \left(1 - \frac{E[s]}{z}\right) \frac{\partial E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}_{t+E[s]}]}{\partial \bar{w}_t} > 0 \end{aligned}$$

However, against the background of the two cases discussed above, this effect will be smaller in an economic downturn than in an upturn.

Having analysed the behaviour of the labour supply side, we now turn to labour demand. In order to avoid costly labour turnover, firms will try to prevent their employees from quitting. Therefore an employer has to set the effective wage v for already employed such that it is not exceeded by the reservation wage, which is the

²¹Note that

$$\frac{\partial E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}_{t+E[s]}]}{\partial \bar{w}_t} = \frac{\partial E[w_{t+E[s]} | w_{t+E[s]} \geq \bar{w}_{t+E[s]}]}{\partial \bar{w}_{t+E[s]}} \frac{\partial \bar{w}_{t+E[s]}}{\partial \bar{w}_t}.$$

Consequently the second term of the equation above will only be positive if the contract wage in t influences future contract wages positively.

expected income from searching.²²

$$v_t \geq \tilde{w}_t = \frac{E[s]}{z} (v - c) + \left(1 - \frac{E[s]}{z}\right) E[w_{t+E[s]} | w_{t+E[s]} \geq \tilde{w}_{t+E[s]}]$$

Minimizing labour costs, the firm will optimally set the effective wage such that this relation is fulfilled with sign of equality. Then

$$v = E[w_{t+E[s]} | w_{t+E[s]} \geq \tilde{w}_{t+E[s]}] - c \frac{E[s]}{z - E[s]}.$$

Consequently, an increase of the contract wage has a positive effect on the effective wage for employees:

$$\frac{\partial v}{\partial \tilde{w}_t} = \frac{\partial E[w_{t+E[s]} | w_{t+E[s]} \geq \tilde{w}_{t+E[s]}]}{\partial \tilde{w}_t} - \frac{cz}{(z - E[s])^2} \frac{\partial E[s]}{\partial \tilde{w}_t} > 0$$

Furthermore, the firm has to fill vacancies. Because there is no Walrasian auctioneer in the labour market, a representative firm has to make a wage offer that optimally weighs the costs of a vacancy with the costs and benefits of a filled job. Given the searching behaviour of the workers and taking into account the distribution of reservation wages and the arrival rate of applicants, expected profits are given by

$$E[\pi] = pg(E[L_1] + E[L_2]) - w(E[L_1] + E[L_2]) - c^F,$$

where p is the product price, $g(\cdot)$ the short term revenue function, c^F fixed cost, and $E[L_1]$ and $E[L_2]$ the expected number of newly hired workers and incumbent workers. Furthermore it is assumed that incumbents and newly hired workers receive the same wage. Because of the search behaviour of the workers an employer knows that he can expect to fill a vacancy the sooner the higher the wage offer he makes. The optimal strategy of the firm will consist of a wage offer w , which maximizes expected profits. Hence,

$$w = pg' - \frac{E[L_1] + E[L_2]}{\frac{\partial E[L_1]}{\partial w} + \frac{\partial E[L_2]}{\partial w}}. \quad (3.15)$$

Because of the search behaviour of the workers we have²³

$$\frac{\partial E[L_i]}{\partial w} > 0, \quad \frac{\partial E[L_i]}{\partial w^C} < 0, \quad \frac{\partial^2 E[L_i]}{\partial w \partial w^C} > 0, \quad i = 1, 2.$$

²²For this argument it does not matter whether v is indeed the present wage or unemployment benefits.

²³Assume that the firm currently employs m workers. If the arrival rate of searching workers is Poisson distributed with expected value $\mu(\bar{w})$, $\mu'(\bar{w}) > 0$, and $\tau_1 := \int_{\bar{w}}^w f_1(\bar{w}, \bar{w}) d\bar{w}$, $\frac{\partial \tau_1}{\partial w} > 0$, $\frac{\partial \tau_1}{\partial \bar{w}} < 0$, where f^1 is the probability density function of the reservation wages over all not incumbent workers, and τ_1 is the probability that an applying worker accepts the job offer w , then the expected number of hirings is $E[L_1] = \mu\tau_1 = \mu(\bar{w}) \int_{\bar{w}}^w f_1(\bar{w}, \bar{w}) d\bar{w}$. Let the probability of acceptance be relatively more reactive to changes of the contract wage than the arrival rate,

Differentiating equation (3.15) yields

$$\frac{dw}{dw^C} = \frac{1}{\frac{d^2E[\pi]}{dw^2}} \left[(w - pg') \left(\frac{\partial^2 E[L_1]}{\partial w \partial w^C} + \frac{\partial^2 E[L_2]}{\partial w \partial w^C} \right) + \left(\frac{\partial E[L_1]}{\partial w^C} + \frac{\partial E[L_2]}{\partial w^C} \right) \right] > 0.$$

The firm will raise its wage offer when the contract wage increases. This reflects the fact that a firm which wants to attract or keep labour, must take into account not only the absolute level of its wage offer, but the relative level with respect to the contract wage. The reason is that firms know that workers expect wage offers to increase when the contract wage goes up.

From equation (3.15) a wage drift equation can be calculated, too. In order to keep the analysis manageable, the following simplifications are made: The revenue function be linear, $g(E[L_1] + E[L_2]) := \theta(E[L_1] + E[L_2])$, where θ is labour productivity, and the reservation wage of both groups of workers have the same rectangular distribution with probability density

$$f_1 = f_2 = f(\bar{w}, w^C) = \frac{1}{w^{max} - w^C}.$$

Then equation (3.15) becomes

$$w = p\theta - w + w^C \quad \Leftrightarrow \quad w = \frac{1}{2}(p\theta + w^C).$$

Taking logarithms, differentiating with respect to time and rearranging yields the wage drift equation

$$\hat{w} - \hat{w}^C = \frac{p\theta}{p\theta + w^C} (\hat{p} + \hat{\theta} - \hat{w}^C)$$

where $\hat{x} := \frac{\dot{x}}{x}$ is the relative growth rate of x over time, and $\frac{p\theta}{p\theta + w^C} \in [\frac{1}{2}, 1]$ because of the minimum wage character of the contract wage. Though the effective wage

$\frac{\partial \tau_1}{\partial \bar{w}} \frac{\bar{w}}{\tau_1} > -\frac{d\mu}{d\bar{w}} \frac{\bar{w}}{\mu}$, then

$$\begin{aligned} \frac{\partial E[L_1]}{\partial w} &= \mu(\bar{w}) f_1(w, \bar{w}) > 0, \\ \frac{\partial E[L_1]}{\partial \bar{w}} &= \tau_1 \frac{d\mu}{d\bar{w}} + \mu \frac{\partial \tau_1}{\partial \bar{w}} = \frac{\mu \tau_1}{\bar{w}} \left(\frac{d\mu}{d\bar{w}} \frac{\bar{w}}{\mu} + \frac{\partial \tau_1}{\partial \bar{w}} \frac{\bar{w}}{\tau_1} \right) < 0, \\ \frac{\partial^2 E[L_1]}{\partial w \partial \bar{w}} &= \frac{d\mu}{d\bar{w}} f_1(w, \bar{w}) + \mu(\bar{w}) \frac{\partial f_1(w, \bar{w})}{\partial \bar{w}} > 0. \end{aligned}$$

The expected number of incumbent workers who accept (i.e., they will not quit) the wage offer w can be calculated analogously by $E[L_2] = m\tau_2 = m \int_{\bar{w}}^w f_2(\bar{w}, \bar{w}) d\bar{w}$, where f_2 is the probability density function of reservation wages over the incumbent workers. Hence,

$$\begin{aligned} \frac{\partial E[L_2]}{\partial w} &= m f_2(w, \bar{w}) > 0, \\ \frac{\partial E[L_2]}{\partial \bar{w}} &= m \left(\int_{\bar{w}}^w \frac{\partial f_2(\bar{w}, \bar{w})}{\partial \bar{w}} d\bar{w} - f_2(\bar{w}) \right) < 0, \\ \frac{\partial^2 E[L_2]}{\partial w \partial \bar{w}} &= m \frac{\partial f_2(\bar{w}, \bar{w})}{\partial \bar{w}} > 0. \end{aligned}$$

depends positively on the contract wage, wage drift is negatively correlated with the contract wage. It increases however, when the product price or labour productivity go up.

3.3.3 *Adaptation versus mark-up hypotheses of wage drift*

Although the wage drift pattern the model of Kleinhüeckelskoten and Spaetling (1980) predicts is in line with the Phillips-curve models, the underlying relationship between contract and effective wage is different. Whereas the Phillips-curve theory predicts that both wages are independent in the long run, the contract wage *persistently* drives effective earnings here. This is remarkable because in both models effective wages are market determined. Obviously, the different results are based on the assumptions about how prices are found on the market. In both models, firms have to compete for workers in order to fill vacancies. However, as long as there is complete information about all relevant variables, the contract wage can influence effective wages only by its legal minimum wage characteristic. It neither affects the structural variables which determine labour demand and supply nor the behaviour of the agents. Unless the contract wage exceeds the equilibrium wage, the wage gap is just the difference between the two wage levels.

In contrast, in an environment of incomplete information, the contract wage will influence the market outcome if it provides information about the other side of the market. Though it leaves the structural variables of the market unaffected, it influences the agents' behaviour. Due to the expectation formation of the workers the firms have to compete for labour by paying a mark-up on the contract wage. Apparently, it is not even important whether the information the contract wage provides is true; if for example the contract wage was not the legal lower boundary of all effective earnings but the workers only believed it was, the outcome would still be the same in the model of Kleinhüeckelskoten and Spaetling. Assuming that effective wages are market determined consequently does not *per se* support the view that wage drift nullifies the decisions taken at the central level of wage bargaining, as the adherents of the Phillips-curve theory claim.

It has been argued that the empirical distinction between adaptational and mark-up patterns of wage drift coincides with the theoretical distinction between theories of market determined and institutionally determined effective wages.²⁴ The latter focuses on the contract wage. The wage gap is just seen as a function of institutionally rather than economically determined variables. In view of the models described in this and the following sections this strict identification of empiricism and theory proves to be untenable. Instead of focusing on either contract or effective wage formation and attempting to explain the differences between these two variables, it is more promising to look at the mutual dependency of both wage levels—whether and how they influence each other.

²⁴For example by Kleinhüeckelskoten and Spaetling (1980) themselves, Deutschmann and Schmiede (1983) and less explicitly by Pehkonen and Viskari (1994) and other authors.

3.4 EFFICIENCY-WAGE THEORIES AND WAGE DRIFT

3.4.1 Models of efficiency wages and the influence of the contract wage

Efficiency-wage theories are based on the idea that the quality of a good can depend on its price. Especially when there is uncertainty about the good's quality, prices can have either incentive or selection effects. In this case, however, economic principles as the "Law of Supply and Demand" or the "Law of the Single Price" do not necessarily hold anymore.²⁵ Demand curves might not be downward sloping, and market demand and supply might be interdependent.

The fact that productivity (the "quality") of labour may depend on its remuneration has long been recognized by economists.²⁶ Attempts to formalize the relationship are not older than forty years, though. These efficiency-wage theories establish a positive relationship between the wage rate and the productivity of labour, leading to a U-shaped cost function of effective labour input, which is minimized at the so called "efficiency wage". This situation is depicted in figure 3.3.

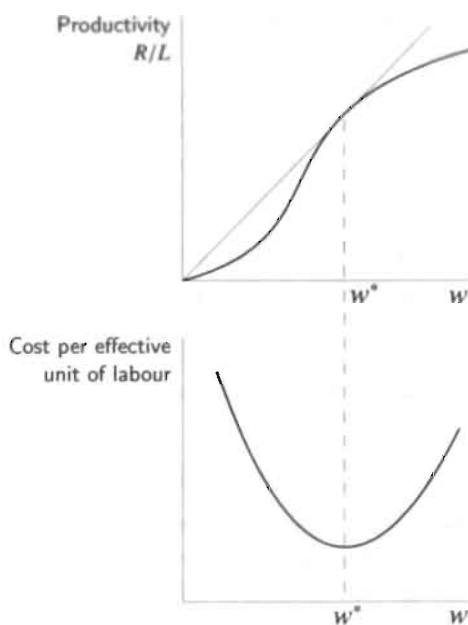


Figure 3.3: Wage-productivity curve and cost per effective unit of labour in efficiency-wage models. The efficiency wage is denoted by w^* .

²⁵See Stiglitz (1987) for a survey of this kind of models and their implications.

²⁶Alfred Marshall (1920), for example, noted, that "[...] highly paid labour is generally efficient and therefore not dear labour; [...]" And Adam Smith (1776) stated that "A plentiful subsistence increases the bodily strength of the labourer and the comfortable hope of bettering his condition and ending his days perhaps in ease and plenty animates him to exert that strength to the utmost."

The most important explanations for why the wage rate might affect labour productivity can be structured as follows.²⁷

1. *Nutritional arguments* (e. g. Mirrlees 1975, Stiglitz 1976):

This argument is put forward with respect to least developed countries. With subsistence wages labour productivity is low. A wage increase improves the workers' standard of living (and especially the nutritional and health situation) which in turn might increase productivity.

2. *Shirking* (e. g. Shapiro and Stiglitz 1984):

Shirking models base on the moral hazard problem that occurs when firms are unable to monitor effort perfectly. Firms can motivate their workers to provide a desired effort level by raising wages above the market clearing level, which creates queues for their jobs, and (at the same time) by threatening workers to fire them in case they are detected shirking. Both measures create a rent of being employed, which, if it is sufficiently high to outweigh the rent due to moral hazard, prevents workers from shirking. In this way, a positive relationship between effort and wage rate is established.

The shirking approach is taken up in the chapters 4 and 5 of this thesis. Chapter 4 combines the shirking model with local wage bargaining, while chapter 5 focuses on the outcome of a shirking model in which a union has say over the punishment of a shirker.

3. *Labour turnover* (e. g. Salop 1979, Schlicht 1978, Stiglitz 1974):

These models provide a similar explanation as the shirking models. If firms are not able to pass the costs of hiring and schooling on the workers, they will have an interest to reduce the quit rate of employees. Paying a higher wage than the outside option makes handing in their notice costly for employees, thus making them more reluctant to quit.

4. *Sociological arguments* (e. g. Akerlof (1982, 1984)):

The sociological models base on a notion of fairness. The underlying assumption is that if workers have the impression to be treated (more than) fairly, their job satisfaction is high and they are willing to exert. Whether a worker believes to be treated fairly might depend on the working rules, her wage rate, or her wage in relation to other comparable employees or to the firm's profits.

Especially the last three arguments have the potential to explain the occurrence of a wage gap. A wage gap will always occur if the contract wage is set below the efficiency wage, because in this case the contract wage is suboptimal for firms. This argument is perfectly in line with the adaptation hypothesis, as long as labour productivity does not depend on the contract wage. If however a link between productivity and

²⁷See also Katz (1986) and Stiglitz (1987).

the contract wage can be established, the contract wage can become a driving force behind the effective wage, which would support the mark-up hypothesis.

Of the models discussed so far, the model of Kleinhückelskotten and Spaetling (1980) can be interpreted in this direction, because the contract wage influences labour turnover.

In order to provide an example of how the link between contract and effective wage could look like in a "real" efficiency-wage model, we consider the model of Schlicht (1992), that adapts on the sociological argument. Schlicht's idea is that workers have a perception of a fair wage level w_f , which depends on the average wage w_0 and the contract wage fixed by collective bargaining \bar{w} . The average wage influences the notion of a fair wage since workers compare their earnings with those of other comparable workers. With respect to the contract wage Schlicht argues that it affects "fairness considerations too, since an increase in the standard wage [i.e., the contract or tariff wage; the author] will create an entitlement to wage increases, and such entitlements influence fairness perceptions (...)"²⁸ Through this mechanism the contract wage can have a strong impact on the effective wage within the efficiency-wage theory. This might explain the fact that collective wage bargaining generally is a struggle, even though employers normally pay a considerable mark-up on the contract wage.

The fair wage is assumed to be a concave function in its two arguments:

$$w_f := \phi(w_0, \bar{w}), \quad \frac{\partial \phi}{\partial w_0}, \frac{\partial \phi}{\partial \bar{w}} > 0, \quad \frac{\partial^2 \phi}{\partial w_0^2}, \frac{\partial^2 \phi}{\partial \bar{w}^2} < 0$$

Workers' effort depends on the degree of generosity of the firm, $\gamma := \frac{w}{w_f}$, which is defined as the ratio of the wage rate w effectively paid by the firm and the fair wage w_f :²⁹

$$e := e(\gamma), \quad e'(\gamma) > 0, \quad e''(\gamma) < 0$$

The firm's profit function is given by

$$\pi = R(e(\gamma)L) - wL,$$

where $R(\cdot)$ is the concave revenue function and L is employment. The firm maximizes its profits with respect to employment and the relative wage rate under the restriction that the effective wage rate must not be below the contract wage. The first order conditions are

$$\begin{aligned} eR' &= w, \\ e'R' &= w_f \vee w = \bar{w}. \end{aligned}$$

²⁸See Schlicht (1992, p. 440).

²⁹Schlicht (1992) makes even stricter assumptions about the fair wage and the effort function. In his paper, ϕ is a linear homogenous function in its arguments and $\lim_{\gamma \rightarrow \infty} e(\gamma) = 1$. However, for the argumentation here, these assumptions are not necessary at all. On the contrary it will turn out that assuming linear homogeneity of the fair wage will have a strong effect on the predicted wage drift pattern.

Consequently, the profit maximizing wage w^* is either equal to the efficiency wage given by the Solow condition,

$$\frac{e' \left(\frac{w^*}{w_f} \right)}{e \left(\frac{w^*}{w_f} \right)} \cdot \frac{w^*}{w_f} := \frac{e'(\gamma^*)}{e(\gamma^*)} \gamma^* = 1, \quad (3.16)$$

or, if the contract wage is greater, it is equal to the contract wage, $w^* = \bar{w}$.

Under the assumption that all firms and workers are identical, the optimal effective wages must be the same in all firms, $w_0^* = w^*$. Then the relative average wage gap can be calculated from equation (3.16) by

$$\frac{w_0^*}{\bar{w}} = \begin{cases} \frac{\phi(w_0^*, \bar{w})}{\bar{w}} \frac{e(\gamma^*)}{e'(\gamma^*)} & \text{if } \gamma^* > 1 \\ 1 & \text{else} \end{cases} \quad (3.17)$$

From the equations (3.17) and (3.16) a wage drift equation can be derived for the case that $\gamma^* > 1$ (i.e. $w_0^* > w_f$):

$$\hat{w}_0^* - \hat{\bar{w}} = \left(\frac{\frac{\partial w_f^*}{\partial \bar{w}} \frac{\bar{w}}{w_f^*}}{1 - \frac{\partial w_f^*}{\partial w_0^*} \frac{w_0^*}{w_f^*}} - 1 \right) \hat{\bar{w}},$$

where $\hat{x} := \frac{\dot{x}}{x}$ denotes the growth rate of x . The wage drift is positive if

$$\frac{\partial w_f^*}{\partial \bar{w}} \frac{\bar{w}}{w_f^*} + \frac{\partial w_f^*}{\partial w_0^*} \frac{w_0^*}{w_f^*} > 1 \wedge \frac{\partial w_f^*}{\partial w_0^*} \frac{w_0^*}{w_f^*} < 1,$$

that is if the elasticities of the fair wage with respect to the average effective wage and to the contract wage sum up to more than one, with the former elasticity being less than one.³⁰ If these conditions are fulfilled wage drift tends to be higher the more reactive the fair wage is with respect to its arguments. The higher the fair wage is, the higher must be the effective wage the firm sets in order to maintain the optimal degree of generosity. If, however, the elasticity of the fair wage is greater than one, firms can—in aggregate—not catch up with an overall increase in effective wages. Raising the effective wage as a reaction to an increase in the average effective wage increases average effective wages even more and would make the fair wage go up over-proportionally. The degree of generosity would become even lower, leading to upward pressure on the effective wage again.

3.4.2 Primary and secondary drift and the stability of the relative wage structure

As stated before, in the “older” literature the occurrence of a wage gap was above all attributed to an excess demand for labour. However, additional arguments have

³⁰Note that wage drift is equal to zero if Schlicht’s assumption holds that the fair wage is linear homogeneous in its arguments.

been put forward which can be interpreted in the view of efficiency-wage theories. This was on the one hand an attempt to explain the permanence of the phenomenon and its occurrence also in situations of increasing unemployment.³¹ On the other hand it gave grounds for the supposition that "drift proceeds from whatever level negotiations set [...], at a rate that is largely independent of that level."³²

In a number of studies a distinction is made between *primary* and *secondary wage drift*.³³ The argument is based on the observation that for different groups of employees different remuneration systems are implemented. Primary drift occurs when the effective wage of only a part of the workers is affected and thereby the relative wage structure to other groups of employees is altered. This might cause dissatisfaction of the other employees, if the new wage hierarchy is not considered to be "fair" anymore. As a consequence, pressure is put on the management to restore the former relative wage structure by increasing the effective wage also of the other groups of workers—secondary drift occurs.

Several explanations have been given for primary drift.³⁴ First, technical progress and an increasing capital intensity tend to raise labour productivity. From this especially employees in the manufacturing industry and above all piece-workers might benefit. Second, the management might consider the negotiated relative wage structure unsuitable for the special needs of the firm. Since the contract wage from collective bargaining cannot be legally undercut, a change in relative wages can only be achieved by increasing the effective wage of some groups of employees. This argument might be of some importance for countries with centralized wage bargaining systems. Finally, there might be excess demand for a special qualification of workers leading to a wage increase for this group.³⁵

The mechanisms behind secondary drift, however, are not made clear in the literature. Gerfin just mentions a "need for correction" of the disturbed wage hierarchy,³⁶ whereas other authors as Phelps Brown (1962) and Lerner and Marquand (1962) explain it as a result of negotiations on the firm level which are initiated by workers who are dissatisfied by the relative wage structure. Indeed, secondary drift can be explained by an "envy" argument in terms of efficiency wages. As was argued in the previous subsection, employees who have the impression to be treated fairly are willing to exert. In turn, firms must be afraid that workers lower their effort if they are discontented.

Gerfin and Phelps Brown report that secondary drift on the firm level can initiate

³¹ See for example Gerfin (1969, p. 502 f.).

³² Phelps Brown (1962, p. 339). Gerfin (1969), however, points out that there is a positive relation between wage drift and the degree of centralization in wage bargaining.

³³ See for example Phelps Brown (1962), Gerfin (1969), and—less explicit—Lerner and Marquand (1962).

³⁴ See for example Gerfin (1969, pp. 504–506).

³⁵ The reasoning behind this argument is based on the increased market value of the qualification. On the one hand firms then have to offer higher wages in order to recruit these workers. On the other hand also an efficiency-wage argument could play a role. With a rising market wage, the outside option of the employed workers increases. This might induce them to quit their job, unless their wage is adapted. So firms will increase the wage in order to reduce costly labour turnover.

³⁶ Gerfin (1969, p. 504).

further drift in other firms or even other branches if it disturbs the interfirm wage structure.³⁷ Furthermore, if central wage negotiations attempt to alter the "normal" wage hierarchy (within or between firms) the described mechanism will tend to restore the old structure.³⁸

3.5 MODELS OF LOCAL AND CENTRAL WAGE BARGAINING

3.5.1 *The basic framework*

In most countries where the contract wage is negotiated on a nation- or industry-wide level, a second round of bargaining takes place on the level of the individual firms after the contract wage has been set.³⁹ In these second level negotiations local unions or the local representatives of the central union bargain with the firm's management about supplements on the contract wage—thus on the wage gap.

Holden (1988) analyses such a wage setting system in three stages. First, the contract wage is set unilaterally by the union. Afterwards the firm determines employment, and in the last stage the local union bargains with the firm's management about the wage gap. The model is solved by backward induction.

Both the union and the firm are risk neutral. They are fully informed about the other party's objective function. The union's utility is equal to the effective wage, $u = w$, whereas the firm maximizes its profits, $\pi = R(L) - wL$, where $R(\cdot)$ is the (concave) revenue function of the firm, L is employment, and w the effective wage rate. Since industrial actions are not allowed during local wage bargaining, workers can threaten the firm only by work to rule. The firm on the other hand has to pay at least the contract wage. Its profits during local wage bargaining are then given by

$$\pi^{\text{dis}} = \alpha R(L) - \bar{w}L, \quad \alpha \in]0, 1[,$$

where α is the factor by which revenues decrease due to work to rule.

Local wage bargaining is modelled as an asymmetric Nash bargaining game. The effective wage w is found by

$$w = \arg \max (\pi - \pi^{\text{dis}})^{\theta} (w - \bar{w})^{1-\theta} \Big|_L ,$$

with θ and $1 - \theta$ being the relative bargaining power of the firm and the workers respectively. The profit-utility pairs corresponding to all feasible outcomes are depicted in figure 3.4. The first order condition implies

$$w = \bar{w} + (1 - \theta) (1 - \alpha) \frac{R(L)}{L}. \quad (3.18)$$

³⁷ See Gerfin (1969, p. 506), and Phelps Brown (1962, pp. 349–350).

³⁸ See for example Söderström and Uddén-Jondal (1982), who describe this effect for Sweden, where unions used to pursue an egalitarian wage policy.

³⁹ Such wage formation systems are, for example, referred to by Calmfors (1990, pp. 51–52) for the Scandinavian countries and by Franz (1994, p. 281). It should be noted, however, that explicit negotiations about the wage rate between management and works councils are illegal in other countries, e.g. Germany. On the other hand, this does not preclude negotiations about extra payments on an informal and discretionary level.

Hence, the firm has to pay a mark-up on the contract wage, which depends on the workers' bargaining power, the decrease in revenues in conflict and on employment.

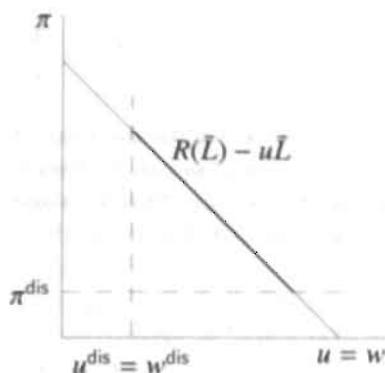


Figure 3.4: Profit-utility pairs corresponding to all feasible outcomes in the Holden model for a given employment level. The possible outcomes are indicated by the bold part of the line.

The firm anticipates the outcome of local wage bargaining when setting employment. Maximizing its profits after all negotiations are finished,

$$\pi = R(L) - wL = [1 - (1 - \theta)(1 - \alpha)] R(L) - \bar{w}L,$$

with respect to employment yields the first order condition

$$R'(L) = \frac{\bar{w}}{1 - (1 - \theta)(1 - \alpha)}. \quad (3.19)$$

Now $w > R'(L)$, and therefore

$$\frac{R(L)}{L} > R'(L) = \frac{\bar{w}}{1 - (1 - \theta)(1 - \alpha)},$$

because $R(L)$ is a concave function. Hence, the marginal cost of labour is smaller than the effective wage rate w . As a result of bargaining, revenue per worker is shared between the employer and the union.

Both employment and the bargained effective wage depend on the contract wage \bar{w} . The total differential of equation (3.19) is given by

$$\frac{dL}{d\bar{w}} = \frac{1}{[1 - (1 - \theta)(1 - \alpha)] R''(L)} < 0,$$

and differentiating equation (3.18) we get

$$\begin{aligned} \frac{dw}{d\bar{w}} &= 1 + (1 - \theta)(1 - \alpha) \frac{1}{L} \left(R'(L) - \frac{R(L)}{L} \right) \frac{dL}{d\bar{w}} \\ &= 1 + \frac{(1 - \theta)(1 - \alpha)}{1 - (1 - \theta)(1 - \alpha)} \cdot \frac{R'(L) - \frac{R(L)}{L}}{LR''(L)} > 1. \end{aligned}$$

The model thus implies a positive relationship between effective and contract wage. Moreover, the two equations above imply

$$\frac{dw}{dL} = [1 - (1 - \theta)(1 - \alpha)] R''(L) + (1 - \theta)(1 - \alpha) \frac{1}{L} \left(R'(L) - \frac{R(L)}{L} \right) < 0.$$

Furthermore, it is interesting to evaluate the effect that bargaining power has on the effective wage and employment. As can be seen from equation (3.18), if the firm has full bargaining power, $\theta = 1$, the effective wage equals the contract wage. In general, we get from the equations (3.18) and (3.19):

$$\begin{aligned} \frac{dL}{d\theta} &= -\frac{(1 - \alpha)\bar{w}}{[1 - (1 - \theta)(1 - \alpha)]^2 R''(L)} > 0 \\ \frac{dw}{d\theta} &= (1 - \theta)(1 - \alpha) \frac{1}{L} \left(R'(L) - \frac{R(L)}{L} \right) \frac{dL}{d\theta} - (1 - \alpha) \frac{R(L)}{L} < 0 \end{aligned}$$

Hence, a greater bargaining power of the firm tends to lower the effective wage and to increase employment.⁴⁰

Finally, wage drift can be calculated by

$$\begin{aligned} \hat{w} - \hat{w} &= \left[\frac{\bar{w}}{w} \left(\frac{\partial w}{\partial \bar{w}} + \frac{\partial w}{\partial L} \frac{dL}{d\bar{w}} \right) - 1 \right] \hat{w} \\ &= -\frac{(1 - \theta)(1 - \alpha)}{1 + (1 - \theta)(1 - \alpha) \left(\frac{\frac{R(L)}{L} - R'(L)}{R'(L)} \right)} \cdot \frac{1}{L} \left[\frac{\frac{R(L)}{L} + R'(L)}{R''(L)} - \frac{\frac{R(L)}{L}}{R'(L)} \right] \hat{w}. \end{aligned}$$

It is greater than zero if the total elasticity of the effective wage with respect to the contract wage is greater than one, or if

$$R'(L) < \frac{R(L)}{L} \left(1 + \frac{R''(L)}{R'(L)} \right).$$

Since $R'(L) < \frac{R(L)}{L}$ wage drift is more likely to be positive the less the curvature of the revenue function is at L . The size of drift furthermore depends negatively on the parameters α and θ .

⁴⁰In view of its interpretation as the inverse of workers' ability to work to rule, also α could be a function of bargaining power, $\alpha = \alpha(\theta)$, $\alpha' > 0$. Then the effects of θ on the effective wage and employment would be

$$\begin{aligned} \frac{dL}{d\theta} &= -\frac{((1 - \theta)\alpha' + (1 - \alpha))\bar{w}}{[1 - (1 - \theta)(1 - \alpha)]^2 R''(L)} > 0, \\ \frac{dw}{d\theta} &= (1 - \theta)(1 - \alpha) \frac{1}{L} \left(R'(L) - \frac{R(L)}{L} \right) \frac{dL}{d\theta} - ((1 - \theta)\alpha' + (1 - \alpha)) \frac{R(L)}{L} < 0. \end{aligned}$$

3.5.2 Efficiency wages in models of local and central bargaining

Muysken and van Veen (1996) and de Gijzel (1996) extended Holden's framework by including efficiency wages. In both models effort is an increasing function in the effective wage:

$$e = e(w), \quad e'(w) > 0$$

It enters production as a productivity factor, i.e., the firm's profits are given by $\pi = R(eL) - wL$. Workers utility is again equal to the effective wage, $u = w$.

If there is a delay in local wage bargaining ("disagreement") workers' effort is given by e^{dis} . The assumption about the disagreement effort is the main difference between the models of Muysken and van Veen on the one hand and de Gijzel on the other hand. While the former assume that the same effort function prevails in agreement as well as in disagreement, $e^{\text{dis}} = e(\bar{w})$, the latter assumes that the workers can deviate from this level as a threat in order to express their dissatisfaction. To which extent the workers can deviate depends on their bargaining power, $e^{\text{dis}} = \bar{e}(\theta)$, $\bar{e}' < 0$.⁴¹ The firm has to pay at least the contract wage in a delay, $w^{\text{dis}} = \bar{w}$. Its disagreement profit is then given by

$$\pi^{\text{dis}} = R(e^{\text{dis}}L) - \bar{w}L.$$

The effective wage is found by solving

$$w = \arg \max (R(e(w)L) - wL - \pi^{\text{dis}})^{\theta} (w - \bar{w})^{1-\theta} \Big|_L \quad \text{s.t. } w \geq w^*. \quad (3.20)$$

In figure 3.5 the profit-utility pairs of all feasible outcomes of this bargaining game are depicted. In view of the reverse U-shape of the related curve two cases have to be distinguished: first, if the contract wage is greater than the profit maximizing wage, $\bar{w} > w^*$, the contract wage is the lower boundary of the possible outcomes of bargaining. Second, if $w^* > \bar{w}$, this lower boundary is given by the profit maximizing wage.⁴² Because of the nonlinearities in equation (3.20) it is impossible to find a general solution. Therefore, only the extreme cases where either the union or the firm have full bargaining power are examined. Consider first the case where the firm has full bargaining power, $\theta = 1$. In this case, the firm can set the effective wage unilaterally under the restriction that the effective wage must not be below the contract wage. The maximization problem thus becomes

$$\max_w R(e(w)L) - wL - \pi^{\text{dis}} \quad \text{s.t. } w \geq \bar{w}.$$

From the Kuhn-Tucker conditions we get

$$e'(w_F) R'(e(w_F)L) = 1 \quad \forall \quad w_F = \bar{w}, \quad (3.21)$$

⁴¹ Another possible assumption about the disagreement effort is $e^{\text{dis}} = \text{const.}$ (see also chapter 4 of this thesis).

⁴² This distinction reflects Nash's axiom that the bargaining outcome has to be Pareto efficient. See Nash (1950). In order to make this argument clear, assume that the bargaining outcome is $w < w^*$. Then a wage rate \hat{w} can be found with $\hat{w} > w$, $\pi(\hat{w}) = \pi(w)$ and $u(\hat{w}) > u(w)$. Hence, w is not efficient.

where w_F denotes the outcome of local wage "bargaining" in this case, in which the firm has full bargaining power. Anticipating this outcome, the firm sets employment in order to maximize profits after all negotiations are finished. The related maximization problem is

$$\max_L R(e(w_F)L) - w_F L,$$

with the first order condition

$$e(w_F) R'(e(w_F)L_F) = w_F. \quad (3.22)$$

Combining the equations (3.21) and (3.22) yields

$$\frac{e'(w_F)}{e(w_F)} w_F = 1 \vee e(\bar{w}) R'(e(\bar{w})L_F) = \bar{w}.$$

Hence, either the Solow condition is fulfilled and the firm pays the efficiency wage, $w_F = w^*$, or the contract wage is greater than the efficiency wage and the firm just pays the contract wage, $w_F = \bar{w}$. The related employment levels can be calculated using equation (3.22). The outcome can be summarized by

$$w_F = \begin{cases} w^* & \text{for } \bar{w} < w^* \\ \bar{w} & \text{for } \bar{w} \geq w^* \end{cases},$$

$$L_F = \begin{cases} \frac{1}{e(w^*)} R'^{-1}\left(\frac{w^*}{e(w^*)}\right) & \text{for } \bar{w} < w^* \\ \frac{1}{e(\bar{w})} R'^{-1}\left(\frac{\bar{w}}{e(\bar{w})}\right) & \text{for } \bar{w} \geq w^* \end{cases}.$$

It is obvious that in this situation the contract wage does not have any influence on the effective wage nor on employment, unless it exceeds the efficiency wage. Consequently, if $\bar{w} \geq w^*$ no wage gap will occur. In this case, the contract wage influences the employment level negatively:

$$\frac{dL_F}{d\bar{w}} = \frac{1 + e'(\bar{w}) R'(e(\bar{w})L_F)}{[e(\bar{w})]^2 R''(e(\bar{w})L_F)} - \frac{e'(\bar{w})}{e(\bar{w})} L_F < 0$$

Note that these results are independent of the disagreement profit π^{dis} . Therefore, the different assumptions about the effort level do not have any influence on this scenario.

Now consider the case where the workers have full bargaining power, $\theta = 0$. The workers' maximization problem is⁴³

$$\max_w w \quad \text{s.t.} \quad \pi \geq \pi^{\text{dis}}.$$

⁴³It would be more correct to write the maximization problem as $\max_w w - \bar{w}$ s.t. $\pi \geq \pi^{\text{dis}}$. The problems are equivalent, though, since if w is maximized, then also $w - \bar{w}$ is maximized.

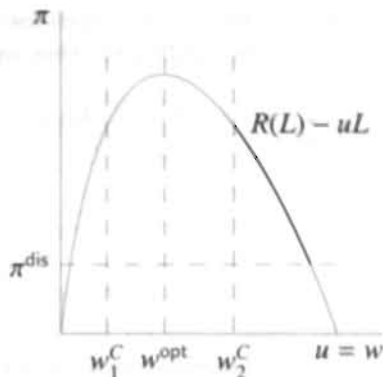


Figure 3.5: Profit-utility pairs corresponding to all feasible outcomes in the models of Muysken and van Veen, and de Gijssel for a given employment level. w^{opt} is the profit maximizing wage for this employment level. The possible outcomes are indicated by the bold part of the line.

The Kuhn-Tucker conditions yield

$$\begin{aligned} \pi &= \pi^{dis} \\ \Leftrightarrow w &= \bar{w} + \frac{1}{L} [R(e(w)L) - R(e^{dis}L)], \end{aligned} \quad (3.23)$$

which has two zeros because of the shape of the revenue function. The higher wage that satisfies this equation will optimally be chosen by the workers as the effective wage w_W :

$$w_W = \max \left\{ w \mid w = \bar{w} + \frac{1}{L} [R(e(w)L) - R(e^{dis}L)] \right\} \geq w^* \quad (3.24)$$

Taking this wage setting behaviour of the workers for given, the firm solves

$$\max_L R(e^{dis}L) - \bar{w}L$$

to find the employment level.⁴⁴ The first order condition is

$$\bar{w} = e^{dis} R'(e^{dis}L_W). \quad (3.25)$$

In this scenario, the effective wage and employment are influenced by the contract wage. Consider the case of Muysken and van Veen in which the disagreement effort

⁴⁴ Actually the firm solves $\max_L R(e(w_W)L) - w_W L$. But since it anticipates that the workers will set the effective wage such that $\pi = \pi^{dis}$, maximizing π is equivalent to maximizing π^{dis} . The original problem is solved by $w_W = e(w_W) R'(e(w_W)L_W)$.

depends on the wage paid during a delay in negotiations, i. e., $e^{\text{dis}} = e(\bar{w})$. Then $w_W = \bar{w}$ is a solution of equation (3.23). Differentiating equation (3.25) we get

$$\left. \frac{dL_W}{d\bar{w}} \right|_{e^{\text{dis}}=e(\bar{w})} = \frac{1 - e'(\bar{w}) R'(e(\bar{w}) L_W)}{[e(\bar{w})]^2 R''(e(\bar{w}) L_W)} - \frac{e'(\bar{w}) L_W}{e(\bar{w})}, \quad (3.26)$$

which has positive sign if

$$\frac{e'(\bar{w})}{e(\bar{w})} \bar{w} > 1 - e(\bar{w}) e'(\bar{w}) L_W R''(e(\bar{w}) L_W).$$

Because of the shape of the revenue function the right hand side of the inequality above is greater than one. Therefore, a necessary but still not sufficient condition for $\left. \frac{dL_W}{d\bar{w}} \right|_{e^{\text{dis}}=e(\bar{w})}$ to be positive is that $\bar{w} < w^*$. If the contract wage is low enough, employment might be increasing with the effective wage.⁴⁵ If it is equal to the profit maximizing wage w^* or above this level, it is decreasing, though.⁴⁶ For the sketch of a graph of the employment function which has the described characteristics see figure 3.6.

The relationship between effective wage and contractual wage can be derived from the equations (3.24) and (3.25):

$$\left. \frac{dw_W}{d\bar{w}} \right|_{e^{\text{dis}}=e(\bar{w})} = \frac{1 - \frac{e'(\bar{w})}{e(\bar{w})} \bar{w}}{1 - \frac{e'(w_W)}{e(w_W)} w_W} \quad (3.27)$$

The analysis of different constellations of \bar{w} and w_W shows that the effective wage is a decreasing function of the contract wage if the latter is below the efficiency wage. If the contract wage is equal to or greater than the efficiency wage, then the effective wage is equal to the contract wage. The graph of w_W is depicted in figure 3.6. A graphical interpolation using the employment function depicted in the same figure yields equivalent results.

⁴⁵This possible behaviour is due to the specific assumption about the disagreement effort. If the workers have full bargaining power the total derivative of the disagreement effort with respect to the contract wage is in this case given by

$$\left. \frac{d\pi^{\text{dis}}}{d\bar{w}} \right|_{e^{\text{dis}}=e(\bar{w})} = L_W \left(\frac{e'(\bar{w})}{e(\bar{w})} \bar{w} - 1 \right),$$

which has positive sign if the contract wage is less than the profit maximizing wage w^* . An increase in the contract wage might then lead to a decrease in the effective wage, such that labour demand goes up. The implications of this behaviour will be discussed later on.

⁴⁶Note that the derivative of employment with respect to the effective wage has the same shape as $\left. \frac{dL_W}{d\bar{w}} \right|_{e^{\text{dis}}=e(\bar{w})}$ in the model of Muysken and van Veen. Hence, we have

$$\left. \frac{dL_W}{dw_W} \right|_{e^{\text{dis}}=e(\bar{w})} = \frac{1 - e'(\bar{w}) R'(e(\bar{w}) L_W) - e(\bar{w}) e'(\bar{w}) L_W R''(e(\bar{w}) L_W)}{[e(\bar{w})]^2 R''(e(\bar{w}) L_W)}.$$

Also a wage drift equation for the model of Muysken and van Veen can be calculated from the equations (3.24) and (3.25):

$$\begin{aligned}\hat{w}_W - \hat{w} &= w_W \frac{e'(w_W)}{e(w_W)} \hat{w}_W - \frac{\bar{w}}{w_W} \left(\bar{w} \frac{e'(\bar{w})}{e^{\text{dis}}} + \frac{w_W - \bar{w}}{\bar{w}} \right) \hat{w} \\ \Leftrightarrow \hat{w}_W - \hat{w} &= \left(\frac{\bar{w}}{w_W} \cdot \frac{1 - \bar{w} \frac{e'(\bar{w})}{e^{\text{dis}}}}{1 - w_W \frac{e'(w_W)}{e(w_W)}} - 1 \right) \hat{w}\end{aligned}\quad (3.28)$$

Under the assumption of de Gijssel (1996), that the disagreement effort does not depend on the contract wage but on the relative bargaining power of the workers, the results differ. Because we have confined to the case where the workers have full bargaining power, we can set $e^{\text{dis}} := \bar{e} = \text{const}$. In this case, the equations (3.26) and (3.27) change to

$$\begin{aligned}\frac{dL_W}{d\bar{w}} &= \frac{1}{\bar{e}^2 L_W R''(\bar{e} L_W)} < 0, \\ \frac{dw_W}{d\bar{w}} &= \frac{1}{1 - \frac{e'(w_W)}{e(w_W)} w_W} > 1\end{aligned}$$

because the effective wage must not be below the profit maximizing wage. In contrast to the model of Muysken and van Veen but similar to the results of Holden, there is a strictly inverse relationship between employment and the contract wage, and the effective wage will increase when the contract wage goes up. Typical graphs of employment and effective wage in relation to the contract wage are depicted in figure 3.6. The wage drift equation is

$$\hat{w}_W - \hat{w} = \left(\frac{\bar{w}}{w_W} \cdot \frac{1}{1 - \frac{e'(w_W)}{e(w_W)} w_W} - 1 \right)$$

Wage drift as the measure of relative changes in the effective wage compared to relative changes in the contract wage is independent of the disagreement effort. Note that this does not hold for the absolute level of the effective wage, though.

The disagreement effort e^{dis} was considered to be an exogenous variable in the model of de Gijssel. In view of its influence on the threat points and the bargaining outcome it can—same as the exponent θ —be interpreted as an expression of bargaining power. In contrast to the model of Muysken and van Veen, the labour demand and the effective-wage function are in fact a family of functions with the disagreement effort $e^{\text{dis}} = \bar{e}$ as a parameter. Differentiating (3.25) we get

$$\left. \frac{dL_W}{d\bar{e}} \right|_{\bar{w}=\text{const.}} = -\frac{1}{\bar{e}} \left(\frac{R'(\bar{e} L_W)}{\bar{e} R''(\bar{e} L_W)} + L_W \right), \quad (3.29)$$

which has positive sign if the marginal revenue of labour is not decreasing in effort, $\frac{\partial \bar{e} R'(\bar{e} L_W)}{\partial \bar{e}} \geq 0$. In this case, employment is the higher the less workers can reduce

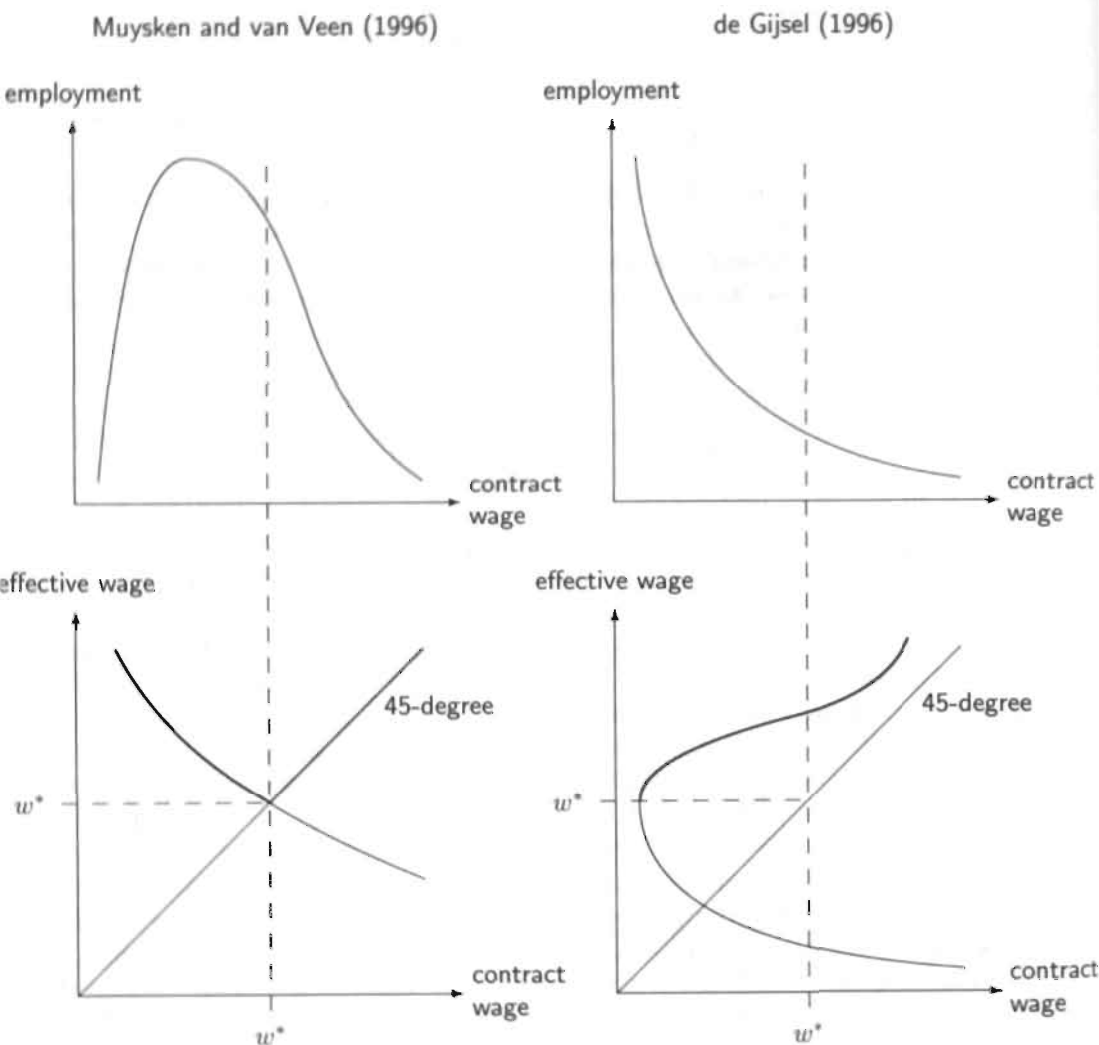


Figure 3.6: Employment and effective wage in the models of Muysken and van Veen (1996) and de Gijzel (1996). The effective wage curve for the latter model is derived by a graphical interpolation.

their effort level if there is a delay in local wage bargaining. The relationship between disagreement effort and the effective wage can be deduced from the equations (3.24) and (3.29):

$$\left. \frac{dw_W}{d\bar{e}} \right|_{\bar{w}=\text{const.}} = - \frac{R'(\bar{e}L_W)}{1 - e'(w_W)R'(e(w_W)L_W)} \quad (3.30)$$

Consequently, the effective wage will be the higher the lower the effort level is in a delay of bargaining.⁴⁷

3.5.3 Threat points and bargaining power in local wage bargaining

In view of the previous subsections the different results of the bargaining models of wage drift seem to be due to the underlying assumptions about the threat points, rather than the question whether efficiency wages matter after an agreement is achieved. The problem of motivating reasonable assumptions about the threat points therefore deserves some attention. In general, two different interpretations are given:⁴⁸

1. The threat points reflect the outside option or a status quo, i. e., a reservation utility or profit the agents can secure when abandoning the attempt to reach an agreement.
2. They represent the utility or profit to the agents in the course of a dispute, for example during a delay of bargaining.

For the described models of local wage bargaining the latter interpretation appears relevant. In disagreement the firm pays the contract wage \bar{w} as the lowest possible wage rate. The assumptions about profits in dispute, however, are different. Holden and de Gijssel allow for workers deviating from their normal effort (function), thus reducing revenues. The same effect occurs in the model of Muysken and van Veen, but they assume that one and the same effort function holds in agreement as well as in disagreement. Reduced effort is therefore induced by the firm itself. Consequently, the workers' threat is in fact no threat at all. A simple thought experiment reduces the underlying assumption to absurdum: if in an environment as described by Muysken and van Veen the firm decided to pay the profit maximizing wage w^{opt} during the dispute, the whole bargaining problem would degenerate and the pure efficiency wage w^* would be the only possible outcome.⁴⁹ Therefore, paying

⁴⁷In principle two cases have to be distinguished. The denominator of the fraction in equation (3.30) is the (simplified) first derivative of normal profits, $\pi = R(e(w_W)L_W) - w_W L_W$, with respect to the effective wage for a given employment level L_W . If effective earnings are higher than the profit maximizing wage, this derivative is negative and effective wages depend positively on the disagreement effort. If however the effective wage is below the profit maximizing wage it is increasing with \bar{e} . This case cannot occur, though, because it implies an inefficient bargaining outcome.

⁴⁸See for example Binmore, Rubinstein, and Wolinsky (1986) and Binmore (1987).

⁴⁹The Nash maximand

$$[R(e(w)L) - wL - (R(e(w^{\text{opt}})L) - w^{\text{opt}}L)]^\theta [w - w^{\text{opt}}]^{1-\theta}$$

a lower wage rate in disagreement is not a credible threat, since it is not optimal for the firm.

Though never stated explicitly, also the first interpretation of the threat points as outside option plays a role in these models, namely if one of the agents has full bargaining power. In this case, she cannot get the other agent to agree to any arbitrary demand. Instead, any bargaining outcome has to secure both agents a minimum income which prevents them from taking up the best alternative elsewhere. Hence, if firms generally pay efficiency wages workers can earn at least w^* everywhere, since other bargaining outcomes are not efficient.⁵⁰ On the other hand non-cooperative bargaining theory shows however, that under reasonable assumptions these "exit points" do not have an influence on the bargaining outcome but setting a minimum constraint.⁵¹

Still another problem is the lack of an explanation of bargaining power in local negotiations. Non-cooperative bargaining theory proposes that the asymmetric Nash bargaining approach can either be justified by assuming different discount rates or by assuming different beliefs about the probability that the negotiations will break down without having led to an agreement. Either assumptions will in turn have an effect on the choice of the threat points, too.⁵² However, it seems reasonable to assume that also other factors play an important role in the determination of bargaining power. Next to legal factors and the structure of negotiations, confidence between the bargaining partners and other long term aspects might have an important influence.

3.6 CONCLUSIONS

In view of the variety of explanations for the occurrence of wage gaps it seems that economic theory is still far from being able to explain the phenomenon. While the different assumptions about the engine of drift could be considered as complementary rather than alternative approaches, the results concerning the predicted drift patterns show fundamental differences.

With respect to their results, models of wage drift can be ordered into two main groups—models which support the adaptation hypothesis and models in favour of the mark-up hypothesis.

The former models are characterized by a reduced form of the following type: the effective wage is a function of a vector of variables $\mathbf{m} := (m_1, m_2, \dots, m_n)$, but not of the contract wage \bar{w} : $w = w(\mathbf{m})$. Wage drift is then given by

$$\hat{w} - \hat{w} = \sum_{i=1}^n \frac{\partial w}{\partial m_i} \frac{m_i}{w} \hat{m}_i - \hat{w}.$$

is negative for all $w \neq w^{\text{opt}}$ unless the workers have full bargaining power, $\theta = 0$.

⁵⁰ Allowing for unemployment, the outside option would be less than, but still be proportional to the (average) efficiency wage.

⁵¹ See Shaked and Sutton (1984).

⁵² See Binmore (1987).

The development of the effective wage does not depend on the contract wage. Depending on the underlying wage formation argument \mathbf{m} can, for example, be a vector of demand and supply parameters as in Phillips-curve models in the long run, or of behavioural parameters as in efficiency-wage models where the effort function does not depend on the contract wage. Consequently, the wage gap is just the calculatory residual between the two wage levels. These theories therefore fail to explain why unions and employers' federations attach great importance to collective bargaining. At best, the contract wage hinders the effective wage from attaining its equilibrium level immediately.⁵³ However, from a long-run perspective this argument should be void. If the contract wage has indeed no significant influence on effective earnings, it might be doubted whether the wage-drift phenomenon deserves any scientific interest anyway.

The mark-up hypothesis is therefore more promising. In the reduced form of the related models the effective wage also depends on the contract wage, $w = w(\mathbf{m}; \bar{w})$. The resulting wage drift equation,

$$\dot{w} - \hat{w} = \sum_{i=1}^n \frac{\partial w}{\partial m_i} \frac{m_i}{w} \hat{m}_i + \left(\frac{\partial w}{\partial \bar{w}} \frac{\bar{w}}{w} - 1 \right) \hat{w},$$

does not necessarily imply that wage drift is positive in these models, as the elasticity of the effective wage with respect to the contract wage might be less than one. However, as long as $\frac{\partial w}{\partial \bar{w}}$ is positive, the effective wage is driven by the contract wage and firms are forced to pay a mark-up. Models with this characteristic are the search model of Kleinhüskelskoten and Spaetling (1980), the efficiency-wage model of Schlicht (1992) and the models of Holden (1988, 1989, 1998), Muysken and van Veen (1996) and de Gijssel (1996). Though sharing the same reduced form, they are distinguished by the assumed engine of wage drift. In terms of their reduced form, the models' wage function $w(\mathbf{m}; \bar{w})$ and their vector \mathbf{m} are different.

In the model of Kleinhüskelskoten and Spaetling, drift is due to incomplete information of workers and firms about the other side of the labour market. Being the legal minimum wage, the contract wage truncates the distribution of wage offers. In this way, it provides information that influences the reservation wage of workers, which in turn will have to be taken into account by the firms. As a result, an increase in the contract wage will induce an increase in effective wages. As it explicitly models the wage formation process on the labour market, this approach might be seen as a counterfactual against the Phillips-curve theory. Although wages are

⁵³If effective wages do not adapt immediately to their equilibrium level, the wage equation becomes

$$w = \bar{w} + \lambda(w^* - \bar{w}),$$

where $w^* = w^*(\mathbf{m})$ is the equilibrium wage and $\lambda \in [0, 1]$ the adaptation velocity. In such an environment wage drift is—at least in the short run—equal to

$$\dot{w} - \hat{w} = \lambda \sum_{i=1}^n \frac{\partial w^*}{\partial m_i} \frac{m_i}{w} \hat{m}_i - \lambda \frac{w^*}{w} \hat{w},$$

which is for example equivalent to the drift equation (3.5) on page 37.

market determined, the contract wage matters. It must not be overlooked, however, that the scope of search models is rather limited. Originally aiming at the analysis of frictional unemployment, they certainly cannot provide *the* theory of wage formation.

In the efficiency-wage model of Schlicht, the contract wage enters the effective-wage function via the fairness perception (or equivalently, envy) of workers. The higher a worker's earnings are in comparison with the wage that other comparable workers receive, and in comparison with the contract wage, the more the worker is willing to exert. To put it the other way around: an increase in the contract wage creates a claim to increase the effective wage. If the firm does not give in, workers are dissatisfied and decrease their effort level. With different groups of workers, once the effective wage for one group increases, the fairness argument can cause a wage gap in the other groups. On the one hand this theory of primary and secondary drift is able to explain the observed persistence of relative wage structures. On the other hand it provides one possible explanation for the persistence of wage drift itself: secondary drift can spread to other firms and other sectors of the economy and finally re-effect the initiating groups.

Next to the fairness perceptions, other efficiency-wage arguments could be employed. In the model of Gahlen and Ramser (1987) effort is assumed to depend on the effective wage, the contract wage and the unemployment rate. This is a rather general approach that is in line with several efficiency-wage arguments.⁵⁴ On the one hand their study shows that the Efficiency-wage theory can be a powerful and flexible instrument for the analysis of the drift phenomenon. On the other hand this flexibility might easily turn into arbitrariness in economic theorizing.

Finally, in models of central and local bargaining, workers' representatives and the management of a firm explicitly negotiate a wage gap. In this way, unions can exert influence on *all* levels of the wage setting procedure.

An overview of the different strands of theory is given in table 3.1. Concerning the described outcomes it should be noted that most of the studies are models of the wage gap rather than wage drift. In a dynamic environment, the relationship

⁵⁴Gahlen and Ramser themselves, p. 130, mention the adverse selection, the shirking and the labour-turnover problem. The main difference to the model of Schlicht (1992) is their assumption that the contract wage influences effort in two ways: first, it has a direct negative effect on effort by its impact on the workers' outside option or by creating entitlements to effective-wage increases. Second, it might have a negative effect on unemployment, which in turn influences effort positively by reducing the outside option. The effective-wage function is derived as

$$w = \bar{w}\phi(1 - u), \quad \phi' > 0$$

(Gahlen and Ramser 1987, p. 135), where u is the unemployment rate which depends on a vector of variables \mathbf{x} and the contract wage, such that $\frac{\partial u}{\partial \bar{w}} \geq 0$. The impact of the contract wage on effective earnings, $\frac{dw}{d\bar{w}} = \phi - \phi' \frac{\partial u}{\partial \bar{w}} \bar{w}$, is positive if its direct effect on effort outweighs the indirect effect. The related wage drift equation is

$$\dot{w} - \dot{\bar{w}} = \frac{d\phi}{du} \frac{u}{\phi} \left(\sum_{i=1}^n \frac{\partial u}{\partial x_i} \frac{x_i}{u} \dot{x}_i + \frac{\partial u}{\partial \bar{w}} \frac{\bar{w}}{u} \dot{\bar{w}} \right),$$

where $\frac{d\phi}{du} < 0$.

between contract and effective wage might change significantly. The effects of long term labour relations, repeated negotiations, learning, etc. are beyond the scope of static models. Moreover, most studies focus on the impact of the contract wage on effective earnings. This impact, however, is likely to be taken into account in collective bargaining.

Theory	Related Studies	Engine of Drift/Gap	Locus of Determination of Drift/Gap	Relationship between Contract and Effective Wage
Phillips curve	i.a. Hansen and Rehn (1956), Gerfin (1969)	excess demand for labour	(neoclassical) labour market	In the long run wage drift nullifies the outcome of collective bargaining. The contract wage follows development of the effective wage (adaptation hypothesis). In the short run it might influence the effective wage.
Search Theory / Labour Turnover	Kleinhückelskoten and Spaetling (1980)	incomplete information / turnover cost	labour market	A mark-up, which might change over the business cycle, is paid over the contract wage. The contract wage drives effective earnings persistently.
Efficiency wages	i.a. Gahlen and Ramser (1987), Schlicht (1992)	incomplete contracts	firm	If the effort or its arguments are a function of the contract wage a mark-up is paid which in turn might depend on the contract wage.
Local Wage Bargaining	i.a. Holden (1988, 1989, 1998), Muysken and van Veen (1996), de Gijssel (1996)	explicit bargaining	firms	Firms pay a mark-up depending on the distribution of bargaining power and the threat points.

Table 3.1: Theories of the wage gap and wage drift and their characteristics.

As we have argued before, a meaningful theory of the wage gap and wage drift should focus on the mark-up hypothesis, if only for the reason that it is common wisdom that the contract wage matters. Of the theories discussed in this chapter, the efficiency-wage models and models of local wage bargaining seem to have the greatest potential to provide an explanation for the occurrence of wage gaps. Both stories are appealing because they are backed by everyday observations of the behaviour of firms and unions. An integrative approach therefore seems to be promising.

However, the discussion has also shown that the models have their weak points. Chapter 4 is related to one problem of local wage bargaining, namely that the outcomes of the related models are very sensitive to the assumptions about the union's threat points. Local wage bargaining models also fail to explain why firms

do not strongly resist a second round of bargaining that obviously can only reduce their profits. The answer that unions enforce these negotiations by reducing workers' effort and blackmailing the firms only shifts the problem to another level. The second part of this thesis (chapters 6 and 7) deals with this problem by showing that a two-tier wage setting system can enhance the efficiency of employment relations.

All the models developed in this book use some kind of efficiency-wage argument. In order to avoid the problem of formulating arbitrary assumptions about how the contract wage enters the effort function, all but chapter 6 explicitly derive the effort function. Especially chapters 4 and 5 are based explicitly on the shirking model. This choice is not motivated by the conviction that fairness considerations, as in the model of Schlicht (1992), are necessarily wrong, but by the fact that shirking models do not only provide a convincing story for an economist trained to believe in rational agents, but also open a convenient way to explain the influence of unions on effort, as is shown in chapter 5.

part of the... (faint text)

(faint text)

Part I

Bargaining in efficiency wage models

The following two chapters 4 and 5 analyse bargaining situations between (local) unions and firms when efficiency wages are paid. In both chapters it is assumed that firms take the incentive problem into account by paying efficiency wages. Chapter 4 focuses on the problem that has already been touched on in chapter 3: what determines the distribution of bargaining power in local wage bargaining, and what are the consequences of different assumptions about the bargaining game for the outcome?

Chapter 5 abandons the assumption that unions only bargain over wages, but introduces effort as a second object of negotiations. However, unions do not bargain directly over effective wages and effort levels, but over a contract wage and a reference effort level that determine the shape of the no-shirking condition. Though chapter 5 provides an explanation for the occurrence of a wage gap that stands on its own, its reference effort level can also help to solve the problem of the determination of the workers' threat point in local wage bargaining.

CHAPTER 4

Threat points in local wage bargaining when firms pay efficiency wages¹

4.1 INTRODUCTION

In a lot of European countries wages are determined by a two-tiered bargaining process. First, nation- or industrywide negotiations about the contract wages take place. Afterwards, on the individual firm level, workers or their representatives and the management bargain, explicitly or implicitly, about a mark-up on this contract wage—the wage gap.

This chapter focuses on local wage bargaining, the second stage of the two-tier negotiations, and discusses the problem how rents are shared between a firm and its employees in the process of bargaining.

Necessarily, wage bargaining on the local level must obey different rules than collective bargaining. The main reason is that collective agreements usually provide for a peace clause, which prohibits strikes and lockouts during the contract period. The bargaining parties have to find other measures to put pressure on their respective opponent. For theoretical models of local wage bargaining this leads to the question which assumptions about the parties' threat points are appropriate. There is a consensus in the literature that the firm's threat point is given by the contract wage from the collective agreement, simply because the contract wage serves as a legal minimum wage. However, the assumptions about the workers' threat point diverge greatly.² Moene (1988) and Holden (1988) assume that workers can cut back their effort to a certain extent and work to rule during a clash, while Muysken and van Veen (1996) criticize just this "exogeneity of disagreement". Instead, they use an efficiency-wage argument, assuming that one and the same no-shirking condition applies during negotiations and after an agreement has been reached. In this case, a reduced effort is the result of the firm's cutting back wages to the contractual level. Finally, de Gijssel (1996), though taking on the efficiency wage argument, argues that during a clash over the effective wage the effort function may shift to a lower level, reflecting—as in Holden's model—workers' dissatisfaction with a disagreement. The

¹ An earlier version of this chapter, de Gijssel and Olthoff (1998) has been published in Haslinger and Stöckner-Venkatarama (1998), *Aspects of the Distribution of Income*, pp. 441–465.

² See also the discussion in chapter 3.

extent of deviation from the normal effort level is considered to be a sign of the workers' bargaining power.

The discussion of the different models in chapter 3 has shown that the outcome of bargaining turns out to be highly sensitive to the models' different assumptions about the threat points, which are, next to the degree of impatience, the only variables in which bargaining power finds an expression in these models. Furthermore, only little attention is paid to the bargaining structure.

In order to fill this apparent gap, this chapter focuses explicitly on the bargaining process, while allowing for different constellations of threat points. The basic idea is that bargaining power influences the outcome of negotiations in different ways. On the one hand, it finds its expression in the disagreement payoffs of the bargaining parties. Like in the models of Holden and de Gijzel, it is assumed that the firm has to pay the contract wage, whereas the workers can vary their effort while there is a disagreement. On the other hand, the distribution of bargaining power also influences the process of bargaining itself. The model of this chapter uses a non-cooperative Rubinstein-like bargaining model³ in which Nature chooses the opponent who has the right to make a wage proposal up to a certain probability which reflects that opponent's activity in the negotiations.

This set-up has two advantages. First, it will be shown, that this probability has a strong impact on the outcome of bargaining. The more often a party can propose a wage, the higher its payoff will be when an agreement is reached.⁴ Second, by assuming that the right to make a wage proposal depends on a certain probability, the problem of the last mover advantage is avoided, which otherwise occurs in finite-horizon Rubinstein bargaining models.⁵

The chapter is organized as follows. In the following section 4.2 the payoff functions of the workers and the firm are derived using an efficiency-wage argument. Section 4.3 introduces the structure of the bargaining game and discusses the outcome of bargaining for different constellations of the contract wage, the profit maximizing wage and the disagreement payoffs. Finally, the conclusion compares these results with those of Holden (1988), Muysken and van Veen (1996) and de Gijzel (1996).

4.2 THE PAYOFF FUNCTION OF THE FIRM AND THE WORKERS

4.2.1 *The no-shirking condition*

In this section the payoffs of the firm and its workers are derived using an efficiency-wage argument. It is assumed that the relationship between wage and effort which is calculated in the following holds in a situation of agreement about the wage. Notice,

³See Rubinstein (1982).

⁴To illustrate this, assume that one opponent can make all proposals during the bargain. Hence, the passive party can only accept or reject them. Then the optimal strategy for the active party is to propose a wage which on the one hand is acceptable for its opponent, but on the other hand maximizes its own payoff.

⁵See Hoel (1987).

that for a situation of disagreement different payoff functions are assumed. These will be derived in section 4.3.1.

Assume that all workers are identical, so that we can restrict our analysis to a representative worker. When she is employed her utility in a given period is given by the difference between the effective wage w she earns and the effort e demanded by her employer. In the case she is not employed she receives some payment a , which we assume to be constant. Let furthermore $w, e, a \in \mathbb{R}_0^+$.

Each worker begins her life being employed and maximizes her utility over an indefinitely long lifetime by deciding whether to be a shirker, which means to work with an effort lower than e , or not to shirk and, hence, to supply e . In contrast to the literature about efficiency wages⁶ we thus assume that an employer considers shirking not to imply an effort of zero but any effort which is lower than the one the employer demands. However, it is clear that also in this case it is optimal for a worker who once has decided to shirk only to offer an effort of zero. Using this assumption we can now derive the maximal effort related to a given wage rate the employer can demand from her workers without making them shirk.

When a worker is caught as a shirker, she will be fired and receive the (net) payment a henceforth, which can be an unemployment benefit or an expected wage in other firms. Hence, in a period of employment a non-shirking worker has a utility of $w - e$, whereas in a period when she is employed but shirks, her utility equals w , and finally in a period of unemployment it equals a .

Let δ be the worker's time preference factor and p the probability that a shirker is caught in one subperiod. Then, the present value of the expected life-time utility of a non-shirking worker is

$$E[U_N] = \sum_{z=0}^{\infty} \delta^z (w - e) = \frac{w - e}{1 - \delta}, \quad (4.1)$$

and

$$E[U_S] = \sum_{z=0}^{\infty} \delta^z [(1 - (1 - p)^z) a + (1 - p)^z w] = \frac{a}{1 - \delta} + \frac{w - a}{1 - \delta + \delta p}, \quad (4.2)$$

is the corresponding value for a shirking worker.⁷

A worker will decide not to shirk if this grants her at least not less utility, $E[U_N] \geq E[U_S]$. This implies for the given wage rate

$$e \leq \frac{\delta p}{1 - \delta + \delta p} (w - a) \equiv \gamma(w - a). \quad (4.3)$$

Hence, the highest feasible effort an employer can demand from his workers without making them shirk is $e = \gamma(w - a)$, where $\frac{\delta p}{1 - \delta + \delta p}$ has been substituted by γ .

⁶Compare for example Shapiro and Stiglitz (1984).

⁷In the last equation it is implicitly assumed that a worker, in the same (sub)period she is caught shirking, does not receive her prior wage anymore, but the outside option a .

Inserting this relationship into the utility function for one subperiod of a non-shirking worker gives

$$u = (1 - \gamma)w + \gamma a. \quad (4.4)$$

This means that the workers' utility in agreement is an increasing function in the wage rate. Therefore, workers will, in order to maximize their utility, strive for an effective wage as high as possible.

4.2.2 The problem of the firm

Once efficiency wages are introduced, there is no simple monotonically decreasing relationship between wages and profit anymore. We normalize the price of the firm's output to one, and assume that the firm produces with labour as the only variable production factor using a Cobb-Douglas technology. Then, profits in one (sub)period are given by the difference between total revenues and labour costs

$$\pi = A(\epsilon L)^\alpha - wL, \quad (4.5)$$

where A is a factor representing total factor productivity, L is the (fixed) number of workers,⁸ and $\alpha \in (0, 1]$ is the production elasticity of labour. Substituting (4.3) with sign of equality into the profit function (4.5), the latter also reduces to a function only in w , as long as L is assumed to be fixed:

$$\pi = A[\gamma L(w - a)]^\alpha - wL \quad (4.6)$$

In contrast to the objective function of the workers, there is no simple monotonic relationship between the effective wage and the profit here. The profit function (4.6) is first increasing in w , has a maximum at some wage w^* , and is decreasing for wages higher than this. The profit maximizing wage w^* is

$$w^* = \frac{(\alpha\gamma A)^{\frac{1}{1-\alpha}}}{\gamma L} + a. \quad (4.7)$$

The related profit is

$$\pi^* \equiv \pi(w^*) = A(\alpha\gamma A)^{\frac{\alpha}{1-\alpha}} - \frac{(\alpha\gamma A)^{\frac{1}{1-\alpha}}}{\gamma} - aL.$$

Unfortunately, the functional form of the profit function (4.6) does not allow to solve explicitly for the bargaining outcome in the following section. Because just the bargaining game is at the very centre of the model, two linear approximations $\tilde{\pi}$ will be used instead, having the zeros $(w_0^1, 0)$ and $(w_0^2, 0)$, $w_0^2 > w_0^1$ and the maximum (w^*, π^*) with π in common:

$$\tilde{\pi} = \begin{cases} \pi^* \frac{w - w_0^1}{w^* - w_0^1} & \text{for } w < w^* \\ \pi^* \frac{w_0^2 - w}{w_0^2 - w^*} & \text{for } w \geq w^* \end{cases} \quad (4.8)$$

⁸We can call the product ϵL the effective amount of labour. As L is fixed, labour input is variable only because effort ϵ can be changed.

Note that because of the concavity of π , its approximation $\bar{\pi}$ underestimates the true values except for the three points mentioned above. Nevertheless, the basic characteristics of the original profit function are preserved. For an illustration see figure 4.1.

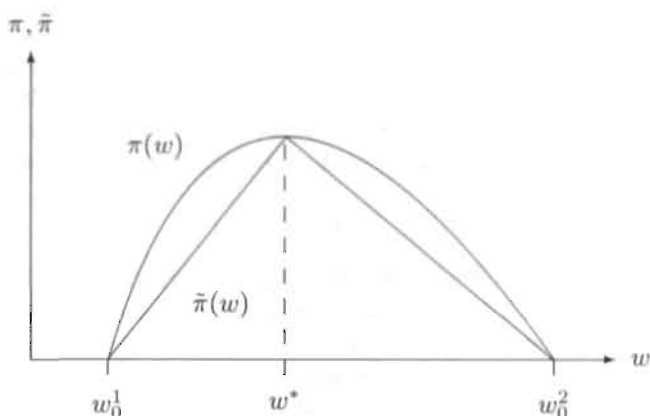


Figure 4.1: The profit function π and its approximation $\bar{\pi}$.

4.3 LOCAL WAGE BARGAINING

4.3.1 The bargaining structure

Consider the time span $[0, 1]$ between two subsequent central wage negotiations. At time $h = 0$ a contract wage \bar{w} is negotiated on a central level by firms' and workers' representatives and it will be accepted by all firms and workers as a minimum wage during $[0, 1]$. At $h = 1$ central bargaining on \bar{w} starts again.

Assume that after the contract wage has been negotiated, an individual firm and its workers can immediately enter into bargaining about the effective wage. Negotiations take place at most at $T + 1$ points in time during the interval $[0, 1]$, depending on whether an agreement is reached or not. The points in time at which negotiations take place are assumed to be equally spread over $[0, 1]$. When the firm and its workers meet at $t \in [0, T]$ there is a probability θ respectively $1 - \theta$ that the firm or the workers will make a wage proposal. It is plausible that a bargaining party will take the initiative if she has bargaining power: in this case she knows that there is a chance that the other party will have to accept her wage proposal. Therefore, the probabilities θ and $1 - \theta$ can be interpreted as a measure of the opponents'

bargaining power.

When negotiations on the effective wage start immediately after the contract wage has been bargained (i.e., at $h = t = 0$), it may happen that both parties achieve an agreement right away, without even entering a round of bargaining. In this case the firm will pay either the contract wage \bar{w} or the profit maximizing wage w^* , depending on which of these wage rates is higher:⁹

$$\hat{w} = \max \{ \bar{w}, w^* \}$$

The wage \hat{w} will be paid throughout the period until central bargaining starts again.

If, however, the workers or the firm do not agree upon this wage rate, one of the parties can initiate negotiations by making another proposal. If this is rejected at $t = 0$ negotiations will be interrupted until $t = 1$. During the period of disagreement the firm has to pay a wage rate w^{dis} which equals the contract wage:

$$w^{\text{dis}} = \bar{w}$$

Workers on the other hand have the opportunity to deviate from their normal effort (compare inequality (4.3)) in a situation of disagreement within certain limits.¹⁰ The effort level during negotiations is denoted by e^{dis} . Next to that, both the firm and the workers encounter some constant bargaining costs c^F respectively c^W per point in time of disagreement. These can for instance represent the costs of discussing the other party's wage proposal, organizing and calling meetings, etc. Then in one point in time the disagreement payoffs of the firm and the workers can be written as

$$\pi^{\text{dis}} = A(e^{\text{dis}}L)^\alpha - w^{\text{dis}}L - c^F, \quad (4.9)$$

$$u^{\text{dis}} = w^{\text{dis}} - e^{\text{dis}} - c^W. \quad (4.10)$$

Now consider an arbitrarily chosen time $t \in [0, T]$ and assume that at all preceding times the proposals have been rejected so that bargaining continues and both the firm and the workers have earned up to now their disagreement payoffs π^{dis} and u^{dis} given in equations (4.9) respectively (4.10). In t either the firm (with probability θ) or the workers (with probability $1 - \theta$) may propose a new effective wage w_t^F respectively w_t^W . If the other bargaining party accepts, the negotiations will end, and the proposal will be the valid wage for the remaining time $[t, T]$ in which the firm and the workers will earn the related payoffs $\pi(w_t^F)$, $u(w_t^F)$ respectively $\pi(w_t^W)$, $u(w_t^W)$.

⁹The reason why the firm does not choose for the lowest possible wage is the efficiency-wage argument. As long as $w^* > \bar{w}$ the firm will pay voluntarily the profit maximizing wage. If $\bar{w} > w^*$ it has to pay the tariff wage because it is the minimum wage agreed upon in central bargaining.

¹⁰Note, that after the central bargaining has ended, in most countries no industrial actions are allowed. However, normally the effort level is only fixed implicitly. During a disagreement workers may work less intensively—for instance work to rule—in order to express their impatience, dissatisfaction or just to put pressure on the firm. This can hardly be interpreted as a contradiction to the effort function derived in section 4.2, because in the case considered here, a deviation from the "normal" effort level is not the decision of an individual worker (as was the decision about shirking) but of the collective of workers. An alternative interpretation—in line with the shirking argument—is that dissatisfaction about the clash increases the disutility of effort during a disagreement.

If however the proposal is rejected, the negotiations are interrupted; firm and worker receive their disagreement payoff in time t , and a new round of bargaining is entered in time $t + 1$. The structure of this local wage bargaining game is also depicted in figure 4.2.

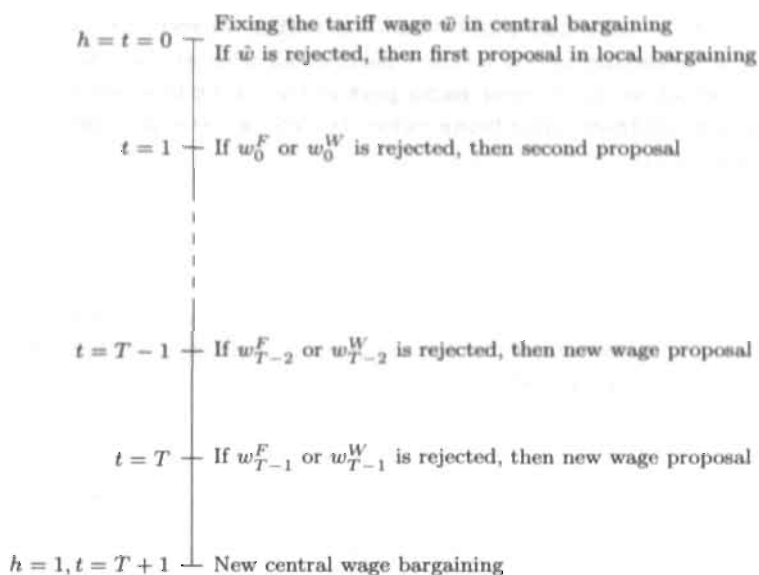


Figure 4.2: The time structure of the bargaining game

In order to preclude any misconceptions it should be noted that the assumption of perfect information implies that firm and workers will typically have no interest in an unnecessary prolongation of the negotiations. In any time t in which negotiations take place, both parties will compare their possible agreement payoffs with the expected payoff of enduring another time of disagreement and possibly achieving a more favourable result in the future. Because a disagreement is costly they will attempt to reach a mutually acceptable agreement as soon as possible. As will be shown in the following subsection the equilibrium solution of the bargaining game described above is the first wage proposal, be it by the firm or by the workers. This proposal is accepted as the effective wage henceforth. So actually negotiations do not take place in times $t > 0$ at all. However, in order to be able to calculate the optimal initial wage proposal the bargaining parties have to know how to act in every contingency that *might* arise. That is, they have to go through *all* points in time $t \in [0, T]$ in their mind, playing the bargaining game virtually.¹¹

¹¹See also Binmore (1987) who states that in this kind of bargaining models only "harmless"

The next subsections discuss the outcome of bargaining for different constellations of the disagreement payoff of the workers u^{dis} , the contract wage \bar{w} and the profit maximizing wage w^* . As it will turn out, only for the first scenario to be considered a general solution for all values of θ can be given. For the other scenarios the analysis has to be confined to the extreme cases that one of the opponents has the "full bargaining power", so that θ equals either zero or one.

Since the profit maximizing wage w^* is dominating all wage levels smaller than w^* , we can confine the discussion to the case that all wage proposals will be in the area $[w^*, w_0^2]$, which is on the decreasing part of the profit function.¹² Thus, even in the case of the contract wage being below the efficiency wage, only this area is relevant for bargaining.

4.3.2 Scenario 1: $u^{\text{dis}} \geq \max\{u(\bar{w}), u(w^*)\}$

The disagreement payoffs of the bargaining parties can be interpreted as their threat points. As can be seen in figure 4.3, with the parameter constellation $u^{\text{dis}} \geq \max\{u(\bar{w}), u(w^*)\}$ to be considered here, these threat points determine the boundaries of the outcome of bargaining.

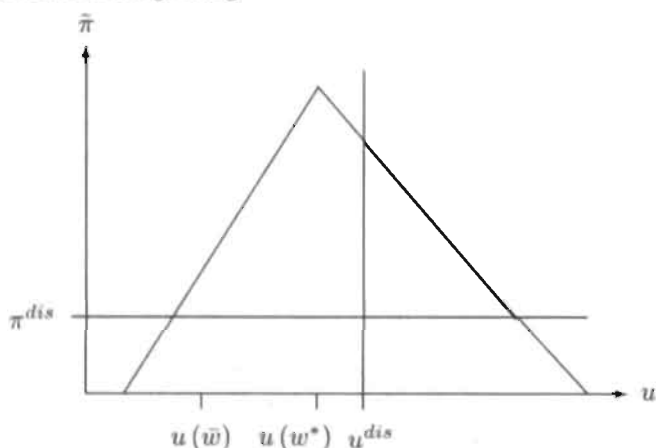


Figure 4.3: Combinations of profit and utility in scenario 1. The bold area on the graph of the profit function depicts the possible outcomes of local wage bargaining when the proposals are only restricted by the disagreement payoffs and $u^{\text{dis}} \geq \max\{u(\bar{w}), u(w^*)\}$. Then a wage gap occurs irrespectively of the outcome of local bargaining. In order to illustrate the parameter constellation profits are expressed in terms of workers' utility

At each point in time $t \in [0, T]$ either the firm or the workers can make a wage proposal with the probabilities mentioned above. If this proposal is accepted by the

assumptions have to hold for the existence of unique subgame-perfect equilibria.

¹²I. e., this part of the profit function is the Pareto frontier.

other party, it will be the valid effective wage \hat{w} for the following periods until the central bargaining starts again. The related payoffs are given by the equations (4.4) and (4.6) in the previous section.

In principle, the initial (and acceptable) wage proposal—i.e., the equilibrium solution of this game—can be found by backward induction, considering all wage offers and demands in each point in time. However, in order to keep things simple, a different but nevertheless equivalent method is used here.

Consider an arbitrarily chosen time $t \in [0, T - 1]$ and assume that all prior proposals have been rejected so that bargaining takes place in this point in time and both the firm and the workers have up to now earned their disagreement payoffs π^{dis} and u^{dis} given in equations (4.9) and (4.10) respectively.¹³

Now the firm, if rejecting a proposal w_t^W made by the workers and entering another period of disagreement, expects future payoffs to have a present value of

$$\begin{aligned} \pi^{\text{dis}} + \text{E}[\pi(w_{t+1})] & \sum_{\tau=t+1}^T \delta^{\tau-t} \\ & \approx \pi^{\text{dis}} + \frac{\delta - \delta^{T-t+1}}{1 - \delta} (\theta \bar{\pi}(w_{t+1}^F) + (1 - \theta) \bar{\pi}(w_{t+1}^W)) \\ & = \pi^{\text{dis}} + \frac{\delta - \delta^{T-t+1}}{1 - \delta} \cdot \pi^* \frac{w_0^2 - (\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W)}{w_0^2 - w^*}, \end{aligned}$$

with δ being the discount rate for both, workers and the firm. In contrast, if the firm accepts the workers' proposal w_t^W the present value of the its profit equals

$$\pi(w_t^W) \sum_{\tau=t}^T \delta^{\tau-t} \approx \frac{1 - \delta^{T-t+1}}{1 - \delta} \cdot \frac{w_0^2 - w_t^W}{w_0^2 - w^*}.$$

In order to make the firm accept, the latter payoff has to be (at least) equal to the former:

$$\begin{aligned} \pi(w_t^W) \sum_{\tau=t}^T \delta^{\tau-t} & \geq \pi^{\text{dis}} + \text{E}[\pi(w_{t+1})] \sum_{\tau=t+1}^T \delta^{\tau-t} \\ \Rightarrow w_t^W & = \frac{1 - \delta^{T-t}}{1 - \delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W) \\ & \quad + \frac{1 - \delta}{1 - \delta^{T-t+1}} \left[\left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^* \right] \end{aligned} \quad (4.11)$$

The same reasoning also holds for the workers. By rejecting the firm's proposal

¹³ Recall that in equilibrium no negotiations will occur at points in time $t > 0$ because the firm or the workers will make acceptable wage proposals already at $t = 0$. However, we need to know the sequence of wage proposals firm and worker would make at any time $t \in [1, T]$ in order to be able to solve for the proposals at $t = 0$.

w_t^F the present value of their expected future utilities equals

$$\begin{aligned} & u^{\text{dis}} + E[u(w_{t+1})] \sum_{\tau=t+1}^T \delta^{\tau-t} \\ &= u^{\text{dis}} + \frac{1-\delta^{T-t+1}}{1-\delta} [\theta((1-\gamma)w_{t+1}^F + \gamma a) + (1-\theta)((1-\gamma)w_{t+1}^W + \gamma a)] \\ &= u^{\text{dis}} + \frac{1-\delta^{T-t+1}}{1-\delta} [(1-\gamma)(\theta w_{t+1}^F + (1-\theta)w_{t+1}^W) + \gamma a]. \end{aligned}$$

The present value of the utility resulting from the firm's offer is

$$u(w_t^F) \sum_{\tau=t}^T \delta^{\tau-t} = \frac{1-\delta^{T-t+1}}{1-\delta} [(1-\gamma)w_t^F + \gamma a].$$

Again, this has to be greater than or equal to the expected utility when rejecting the proposal:

$$\begin{aligned} u(w_t^F) \sum_{\tau=t}^T \delta^{\tau-t} &\geq u^{\text{dis}} + E[u(w_{t+1})] \sum_{\tau=t+1}^T \delta^{\tau-t} \\ &\Rightarrow w_t^F = \frac{1-\delta^{T-t}}{1-\delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1-\theta)w_{t+1}^W) \\ &\quad + \frac{1-\delta}{1-\delta^{T-t+1}} \cdot \frac{u^{\text{dis}} - \gamma a}{1-\gamma} \end{aligned} \quad (4.12)$$

The two equations (4.11) and (4.12) form a system of indifference equations with variable coefficients,

$$\begin{aligned} w_t^F &= \frac{1-\delta^{T-t}}{1-\delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1-\theta)w_{t+1}^W) + \frac{1-\delta}{1-\delta^{T-t+1}} \Phi, \\ w_t^W &= \frac{1-\delta^{T-t}}{1-\delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1-\theta)w_{t+1}^W) + \frac{1-\delta}{1-\delta^{T-t+1}} \Psi, \end{aligned}$$

where for convenience $\Phi = \frac{u^{\text{dis}} - \gamma a}{1-\gamma}$ and $\Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^*$ have been set.¹⁴

The general solutions are¹⁵

$$w_t^F = \frac{1-\delta}{1-\delta^{T-t+1}} \delta^{T-1} w_T^F + \frac{1-\delta^{T-1}}{1-\delta^{T-t+1}} [\Phi + (1-\theta)\delta(\Phi - \Psi)], \quad (4.13)$$

$$w_t^W = \frac{1-\delta}{1-\delta^{T-t+1}} \delta^{T-1} w_T^W + \frac{1-\delta^{T-1}}{1-\delta^{T-t+1}} [\Psi - \theta\delta(\Phi - \Psi)], \quad (4.14)$$

¹⁴ Φ is equal to the wage that makes the workers' utility after an agreement has been reached equal to their disagreement utility, and Ψ is equal to the wage that makes the firm's agreement profit equal to its disagreement profit. Consequently, Ψ is the greatest possible wage and Φ the lowest possible wage firm and workers can ever agree upon. Notice that Ψ is always greater than the contract wage \bar{w} : due to the bargaining cost and the workers' ability to lower their effort during disagreement, the disagreement profit cannot be greater than the agreement profit at the contract wage, $\pi^{\text{dis}} \leq \pi(\bar{w}) \approx \bar{\pi}(\bar{w}) = \pi^* \frac{w_0^2 - \bar{w}}{w_0^2 - w^*}$. Because Ψ is decreasing in π^{dis} it follows that

$\Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^* \geq \left(1 - \frac{\bar{\pi}(\bar{w})}{\pi^*}\right) w_0^2 + \frac{\bar{\pi}(\bar{w})}{\pi^*} w^* = \bar{w}$.

¹⁵ For the derivation of the solution see the appendix of this chapter on page 87 ff.

with w_T^F and w_T^W being the proposals the firm respectively the workers would make in the last point in time T if they can take the initiative. As was already stated, with the parameter constellation considered here, these last wage proposals are just bound by the threat points. Assume that the bargain reached T , then the firm will have to propose some wage w_T^F which causes the workers to be at least indifferent between accepting it or continuing the disagreement until the next central wage negotiations begin. Hence,

$$\begin{aligned} u(w_T^F) &\geq u^{\text{dis}} \\ \Rightarrow w_T^F &= \frac{u^{\text{dis}} - \gamma a}{1 - \gamma} \equiv \Phi. \end{aligned} \quad (4.15)$$

A similar argument¹⁶ holds for the workers' proposal, resulting in¹⁷

$$\begin{aligned} \pi(w_T^W) &\geq \pi^{\text{dis}} \\ \Rightarrow w_T^W &= \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^* \equiv \Psi. \end{aligned} \quad (4.16)$$

Inserting (4.15) and (4.16) into the general solutions (4.13) and (4.14) for w_0^F and w_0^W , i.e., for the first wage proposals made by the firm respectively the workers in time $t = 0$ after having declared the disagreement, we get

$$w_0^F = \Phi + (1 - \theta) \frac{1 - \delta^T}{1 - \delta^{T+1}} \delta(\Psi - \Phi), \quad (4.17)$$

$$w_0^W = \Psi - \theta \frac{1 - \delta^T}{1 - \delta^{T+1}} \delta(\Psi - \Phi). \quad (4.18)$$

Hence, the initial wage proposals are a weighted average of the wages Φ and Ψ , that are related to the threat points.¹⁸ The higher the bargaining power of one of

¹⁶In order to show that this argument really implies the firm's optimal wage proposal, consider the following game for time T :

1. The firm proposes some wage W_T^F , which can be higher, less or equal to w_T^F from equation (4.15).
2. The workers decide about accepting this proposal or rejecting it.
3. If the workers decide to accept, the payoffs are $u(W_T^F)$ and $\pi(W_T^F)$, otherwise u^{dis} and π^{dis} .

Solving for the outcome of this game by backward induction will lead to w_T^F from equation (4.15). A corresponding argument holds for the workers' last proposal.

¹⁷Note that because of the concavity of the profit function π^{dis} is actually related to two different wages, unless $\pi^{\text{dis}} = \pi^*$. The workers will, however, decide for the higher wage rate in order to maximize their payoffs.

¹⁸Note that the proposals of the firm and the workers do *not* converge, even if the number T of potential bargaining times approaches infinity. As a consequence, there is a kind of first mover advantage. Yet, the expected value of the first wage proposals is an average of Φ and Ψ weighted by the probability θ :

$$E[w_0] = \theta\Phi + (1 - \theta)\Psi$$

Note, that this is exactly the outcome of the corresponding Nash bargaining game. However, for the scenarios 2 and 3 (subsections 4.4 and 4.5) the results will differ.

the parties, the closer its initial wage proposal is to the disagreement payoff of its opponent. A party with full bargaining power is able to push its opponent on its reservation, i. e., its disagreement payoff. If the workers have full bargaining power, $\theta = 0$, their initial proposal will be

$$w_0^W(\theta = 0) = \Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^*.$$

In the case the firm has full bargaining power, $\theta = 1$, it will propose

$$w_0^F(\theta = 1) = \Phi = \frac{u^{\text{dis}} - \gamma a}{1 - \gamma}.$$

Because this is the lowest possible effective wage in this scenario, it can be concluded that there will be a wage gap irrespectively of the outcome of bargaining, unless the firm has full bargaining power, $\theta = 1$, and the disagreement utility is exactly equal to the (agreement) utility at the contract wage, $u^{\text{dis}} = (1 - \gamma)\bar{w} - \gamma a \equiv \bar{w} - e(\bar{w})$. In this case, the workers' bargaining costs would exactly offset the utility gain from supplying less effort during a disagreement.

In general, the wage gap is influenced by the disagreement payoffs. The higher the disagreement payoff of the workers and the lower that of the firm, the greater will be the bargained wage and consequently the wage gap. Because $\frac{\partial w_0^F}{\partial u^{\text{dis}}}, \frac{\partial w_0^W}{\partial u^{\text{dis}}} > 0$ and $\frac{\partial w_0^F}{\partial e^{\text{dis}}}, \frac{\partial w_0^W}{\partial e^{\text{dis}}} < 0$,¹⁹ the firm will strive to pay a wage as low as possible and the workers will strive to work as less as possible during a disagreement.

4.3.3 Scenario 2: $u(\bar{w}) \geq \max\{u^{\text{dis}}, u(w^*)\}$

Compared to the first scenario the case to be considered here is complicated, because now the level of the contract wage is decisive for the outcome of bargaining, too.²⁰ Since the contract wage is accepted by both opponents to be the minimum wage, it is the lower boundary of all wage proposals (compare figure 4.4).

As a consequence, the workers' payoff related to the lowest possible wage offer of the firm deviates from her disagreement payoff, and it is impossible to solve for the outcome of bargaining in general terms. To show this, consider first the last point in time T . The firm, if it is its turn to make a proposal, has to offer at least

$$w_T^F = \bar{w},$$

whereas the workers would propose

$$w_T^W = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^* \equiv \Psi$$

¹⁹ Recall that $\Phi = \frac{u^{\text{dis}} - \gamma a}{1 - \gamma}$, $\Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^*$ and $u^{\text{dis}} = w^{\text{dis}} - e^{\text{dis}} - c^W$, $A(e^{\text{dis}}L)^\alpha - w^{\text{dis}}L - c^F$.

²⁰ Note, that because of the monotonicity of the workers' utility function this scenario holds for $\bar{w} \geq w^*$.

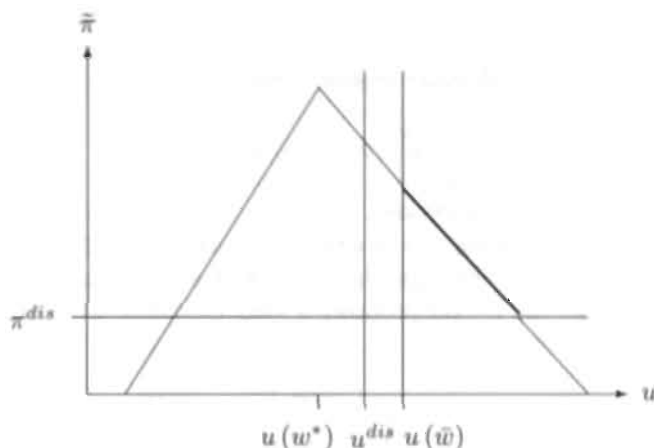


Figure 4.4: Combinations of profit and utility in scenario 2. The bold area on the graph of the profit function depicts the possible outcomes of local wage bargaining when the proposals are restricted by the contract wage \bar{w} and the disagreement payoff of the firm π^{dis} . In this scenario, whether a wage gap occurs or not depends only on the bargaining power of the opponents.

(compare (4.16)). In $T-1$ the firm has to propose a wage making the workers indifferent between accepting or rejecting it. It follows

$$\begin{aligned} (1+\delta)u(w_{T-1}^F) &\geq u^{\text{dis}} + \delta E[u(w_T)] \\ \Rightarrow w_{T-1}^F &\geq \frac{1}{1+\delta}\Phi + \frac{\delta}{1+\delta}(\theta\bar{w} + (1-\theta)\Psi) \equiv Q. \end{aligned} \quad (4.19)$$

Note, that Q is “acceptable” for the workers since they cannot expect to improve by rejecting it. However, if $\bar{w} > Q$ it is not “legal”, because the contract wage is the minimum wage rate. Depending on whether the contract wage is greater or less than the right-hand side of inequality (4.19) the wage proposal of the firm will be

$$w_{T-1}^F = \begin{cases} Q & \text{if } Q > \bar{w} \\ \bar{w} & \text{else} \end{cases}.$$

This ambiguity of the firm’s wage proposal can also occur in points in time $t < T-1$. It is more likely in periods close to T and the greater the difference between the contract wage \bar{w} and $\Phi \equiv \frac{u^{\text{dis}} - \gamma a}{1-\gamma}$ is.

For this reason, we can only derive the outcomes of bargaining for the cases in which the opponents have full bargaining power. A solution for values of θ other than 0 or 1 can only be found numerically.

Consider first the firm. If $\theta = 1$ it will make a proposal in each point in time until an agreement is reached. Since their disagreement payoff is lower than the

utility related to the contract wage, the workers cannot improve by rejecting a firm's proposal. This can be seen by inequality (4.19), which is satisfied by $w_{T-1}^F = \bar{w}$ when $\theta = 1$. The same holds for all other points in time. Therefore, in $t = 0$ the firm will make a proposal

$$w_0^F(\theta = 1) = \bar{w}, \quad (4.20)$$

which will be accepted by the workers.

In contrast, the outcome for $\theta = 0$, when the workers have full bargaining power, is the same as in scenario 1, because they do not face any legal restrictions on their wage proposals. Thus, the worker proposes a wage that makes the firm's profit equal to its disagreement profit,²¹

$$w_0^W(\theta = 0) = \Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi^*}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi^*} w^* > \bar{w}. \quad (4.21)$$

We can conclude that a wage gap will occur unless the firm has full bargaining power, $\theta = 1$. However, the difference between the contract and the bargained wage will be smaller than in scenario 1 if the firm can propose $w_t^F = \bar{w}$ in points in time $t \in [\tau, T]$ for $\tau \geq 0$.

4.3.4 Scenario 3: $u(w^*) > \max\{u^{\text{dis}}, u(\bar{w})\}$

Formally this scenario is identical with scenario 2. Consider the proposal the firm would make at the last point in time T . Although it does not need to suggest more than the contract wage or $\Phi \equiv \frac{u^{\text{dis}} - \gamma\alpha}{1 - \gamma}$ in order to be accepted, it will propose the higher wage w^* in order to maximize its profits. Now a similar reasoning as in the previous subsection holds for the firm's proposals. There may be points in time $t \in [\tau, T]$, $\tau \geq 0$ in which $w_t^F = w^*$ is acceptable for the workers. Furthermore, if the firm has full bargaining power, $\theta = 1$, the wage to be agreed upon is the profit maximizing wage $w_0^F = w^*$.²²

However, in this scenario a wage gap will occur in any case due to the underlying assumption about the parameter constellation, $u(w^*) > \max\{u^{\text{dis}}, u(\bar{w})\}$, as can be seen in figure 4.5.

²¹Since in scenario 1 the outcome of our model and the corresponding Nash bargaining game were equivalent (compare footnote 18), one could think that this also may be the case here. In fact, maximizing the Nash maximand $\Omega = (\bar{\pi}(w) - \pi^{\text{dis}})^\theta (u(w) - u(\bar{w}))^{1-\theta}$ over w leads to $w = \theta\bar{w} + (1-\theta)\Psi$, which is a weighted average of the first wage proposals (4.20) and (4.21). However, with these it was assumed that the firm respectively the workers have full bargaining power. Second, the outcome of the Nash bargaining game is differentiable in θ , whereas we concluded that the outcome of our bargaining game in this scenario is not.

²²Of course the outcome for the workers having full bargaining power is again equal to that in scenario 1:

$$w_0^W(\theta = 0) = \Psi = \left(1 - \frac{\pi^{\text{dis}}}{\pi(w^*)}\right) w_0^2 + \frac{\pi^{\text{dis}}}{\pi(w^*)} w^* \quad (4.22)$$

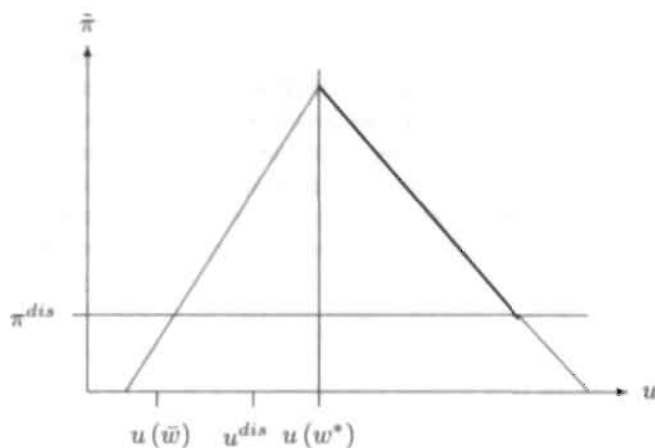


Figure 4.5: Combinations of profit and utility in scenario 3. The bold area on the graph of the profit function depicts the possible outcomes of local wage bargaining when the proposals are restricted by the profit maximizing wage w^* and the disagreement payoff of the firm π^{dis} . In this case a wage gap always occurs.

4.3.5 The decision about bargaining

Having solved for the outcome of a possible bargain, it will now be shown that local wage bargaining will occur under virtually any circumstances.

Recall that without local bargaining the firm would voluntarily pay $\hat{w} = \max\{\bar{w}, w^*\}$ (see section 4.3.1). Both the firm and the workers will agree upon this wage, if they do not expect to be better off by rejecting it and entering a disagreement. In fact, the results of the preceding subsections suggest that only the workers can have an interest in local wage bargaining, while the firm's profits will be reduced if it is forced to pay a mark-up on the contract wage. Nevertheless, also the firm might call for local wage bargaining in order to prevent the workers from gaining a first-mover advantage in the negotiations.

In spite of the assumption that the right to make a wage proposal in some point in time is assigned by certain probabilities, it is reasonable to assume that this rule does not apply for the initial round of bargaining. Instead, assume that when only one party asks for bargaining, it can force negotiations by imposing disagreement costs on its opponent (and on itself), and that it has the right to make the initial proposal w_0^F respectively w_0^W . If both parties ask for bargaining the right to make a proposal is stochastic as before. Then the parties compare the present value of their payoffs from \hat{w} with the expected payoffs from the bargained wage, $\theta\pi(w_0^F) + (1-\theta)\pi(w_0^W)$ and $\theta u(w_0^F) + (1-\theta)u(w_0^W)$.

Since in any case \hat{w} is the lower boundary for the outcome of local bargaining, the workers will always have an incentive to ask for local wage negotiations unless the firm has full bargaining power.

The firm on the other hand cannot be better off if it comes to negotiations. Nevertheless, it knows that if only the workers claim for negotiations, they will have a first-mover advantage by making the first proposal and thereby determining the effective wage.²³ If however also the firm claims for bargaining, it will have the right to make a proposal with probability θ and can expect a lower wage than when leaving the initiative to the workers alone. Hence, when the firm knows that the workers will claim for negotiations, its best strategy is to do so, too.

4.4 CONCLUSIONS

This chapter has shown that a wage gap occurs as a result of local wage negotiations for almost all possible parameter constellations under consideration. Only in case the firm has full bargaining power and, at the same moment, does not voluntarily pay a wage higher than the contract wage from collective bargaining in order to maximize its profits, the effective wage will equal the contract wage. However, under these presumably exceptional circumstances no local wage bargaining will take place at all, since neither the workers nor the firm can expect to be better off by negotiating.

The results of this chapter are related to those of Holden (1988), Muysken and van Veen (1996) and de Gijssel (1996). As it was already stated in the introduction, a basic characteristic of these models is that their to some respect highly different results are mainly due to their assumptions especially about the disagreement payoffs, whereas only little attention is being paid to bargaining power and its effect on the structure of bargaining.

As it is shown in this chapter, their results can be derived as special solutions of a local wage bargaining game in which the firm faces a shirking problem and knows that its workforce will work to rule in the case of an unresolved wage conflict. Given the specific bargaining structure the different results depend essentially on the relation between the workers' net utility of the exogenously given contract wage, the technologically determined efficiency wage and the workers' disagreement utility.

Holden (1988) does not include efficiency wages in his model.²⁴ Hence, neglecting the existence of a profit maximizing wage his profit function is strictly decreasing in the wage rate. As the lower boundary of the outcome of bargaining is given by the tariff wage \bar{w} , his model is best reflected by scenario 2 of the model of this chapter, though here the assumption that underlies the profit function is different. Indeed, Holden's result about the occurrence of a wage gap is the same as in scenario 2: there will be a wage gap unless the firm has full bargaining power.

Muysken and van Veen (1996), considering the exogenously given effort level in disagreement as "a drawback of the Moene/Holden type of models", assume a functional relationship between the wage rate and the effort that holds in agreement

²³Recall that the first wage proposal w_0 will be the equilibrium outcome of the bargaining game if both parties are acting rationally (see also section 4.3.1), and that the first wage proposal of the firm will always be lower than that of the workers.

²⁴However, effort does play a role, because he assumes that in disagreement workers can choose a lower effort than normal.

as well as in a disagreement situation.²⁵ The essence of this assumption is that bargaining is not costly: the payoffs of firm and workers during a disagreement are not reduced by a lower effort or other bargaining cost. A wage gap can therefore only occur if it leads to a Pareto-improvement. Consequently, two cases can be distinguished in their model. The first, with the profit maximizing wage being higher than the tariff wage ($w^* > \bar{w}$), can be compared to scenario 3 and leads to exactly the same result: a wage gap will occur, because the firm voluntarily pays a higher wage than the tariff wage in order to maximize its profits. For the contrary case, $\bar{w} \geq w^*$, Muysken and van Veen conclude that no wage gap will occur, whatever the bargaining power of the workers is. Like in de Gijssel's (1996) model, the related result in this chapter is different: this case is again best described by scenario 2. With the workers' bargaining power being above zero, the bargained wage will always be higher than the contract wage because the workers can impose costs of disagreement on the firm. In contrast to the claim of Muysken and van Veen, there is apparently no reason why the contract wage should always be below the profit maximizing wage.

Next to the parameter constellations which can already be found in the earlier models, this chapter has identified another one, that is described in scenario 1. In contrast to the other scenarios, here the outcome of bargaining does not depend on the constellation of the tariff wage and the profit maximizing wage, but only on the disagreement payoffs of the bargaining parties.²⁶ Then a wage gap will occur, apart from the extremely unlikely case in which, simultaneously, the firm has full bargaining power, and the union cannot impose any bargaining cost on the firm, i.e., $u(\bar{w}) = u^{\text{dis}} \Leftrightarrow \bar{w} = \Phi$.

After having assessed the different parameter constellations in local wage bargaining, the question arises whether they are all relevant. Taking into account that in central wage bargaining unions negotiate with employer federations, which represent different firms with different profit functions, the answer is that most likely none of the scenarios can be excluded *a priori*.

Nevertheless, the determination of the disagreement payoffs remains a problem in models of local wage bargaining. Though the probabilities θ and $1 - \theta$ are important expressions of the firm's respectively the workers' bargaining power, as they influence not only the outcome but also the structure of negotiations, it is clear that the choice of the disagreement wage w^{dis} and effort e^{dis} also has a strong impact because of the influence of these variables on the disagreement payoffs. While it seems to be quite reasonable to set $w^{\text{dis}} = \bar{w}$,²⁷ the choice of e^{dis} is still prone to arbitrariness.

²⁵In terms of the model of this chapter this would mean $u^{\text{dis}} = (1 - \gamma)w^{\text{dis}} + \gamma a, \pi^{\text{dis}} = A[\gamma L(w^{\text{dis}} - a)]^\alpha - w^{\text{dis}}L$. This assumption has an important shortcoming: as long as the firm can set $w^{\text{dis}} = w^*$, there is only one possible outcome of bargaining—the profit maximizing wage, since it is the only acceptable wage for both, workers and the firm, in this case. Once bargaining costs are introduced such that especially $\pi^{\text{dis}} < \pi(w^*)$, bargained effective wages higher than w^* are possible.

²⁶However, due to the assumption that the firm has to pay the contract wage in disagreement, \bar{w} has some influence on the bargained effective wage.

²⁷In a model of repeated central and local wage negotiations it may be more likely that the firm

It is likely that there is a lower boundary of the disagreement effort, e_{\min}^{dis} . Since industrial actions are not allowed in most countries after a tariff wage has been set, e_{\min}^{dis} may be technologically given and represent, for example, work-to-rule. The real disagreement effort must then be somewhere between e_{\min}^{dis} and $e(\bar{w})$. But though, at first sight, workers have an incentive to choose for a disagreement effort as low as possible, they may face some obstacles. First, it may worsen the atmosphere of negotiations, lead to a breach of trust and cause the firm to make less concessions in later negotiations. Second, if having high bargaining power the workers must take into account that the firm may be forced to shut down if the bargained effective wage implies negative profits. Hence, self-restraint with respect to e^{dis} can be necessary in order to prevent the firm's disagreement payoff from falling below a certain threshold. It follows that also the choice of e^{dis} is an expression of bargaining power and it seems as if this variable is far away from being unidimensional as is assumed implicitly in most of the literature about bargaining.

Yet, the solution might be found on another level. This chapter has focussed on local wage bargaining and ignored the presumable interdependence between local and central negotiations. However, if the workers' threat in local bargaining really consists of work-to-rule, i.e., "work according to working rules" (Moene 1988), it is unlikely that the disagreement payoff is just technologically given. Rather, it is very probable that the parties to collective bargaining on the centralized level have a stake in their determination. The following chapter, though it does not deal with local bargaining, might shed some light on this problem.

has to pay the last period's effective wage w_{-1} during a disagreement.

4.A APPENDIX—SOLUTION OF THE DIFFERENCE EQUATIONS (4.11) AND (4.12)

The two equations (4.11) and (4.12) form a system of difference equations with variable but known coefficients,

$$\begin{aligned}w_t^F &= \frac{1 - \delta^{T-t}}{1 - \delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W) + \frac{1 - \delta}{1 - \delta^{T-t+1}} \Phi, \\w_t^W &= \frac{1 - \delta^{T-t}}{1 - \delta^{T-t+1}} \delta (\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W) + \frac{1 - \delta}{1 - \delta^{T-t+1}} \Psi,\end{aligned}$$

where we have set $\Phi \equiv \frac{\pi^{dis} - \gamma \alpha}{1 - \gamma}$ and $\Psi \equiv \left(1 - \frac{\pi^{dis}}{\pi(w^*)}\right) w_0^2 + \frac{\pi^{dis}}{\pi(w^*)} w^*$. Subtracting the first from the second equation and rearranging we get

$$\begin{aligned}w_t^F &= w_t^W + \frac{1 - \delta}{1 - \delta^{T-t+1}} \Phi - \frac{1 - \delta}{1 - \delta^{T-t+1}} \Psi \\ \Leftrightarrow w_t^W &= w_t^F - \frac{1 - \delta}{1 - \delta^{T-t+1}} \Phi + \frac{1 - \delta}{1 - \delta^{T-t+1}} \Psi.\end{aligned}$$

Notice the implication of this equations: the differences between the wage proposals of firm and worker increase the longer the negotiations go on, but the difference does not vanish for t close to zero.

Except of the changing time indices, the same relation must hold for the next point of time $t + 1$:

$$\begin{aligned}w_{t+1}^F &= w_{t+1}^W + \frac{1 - \delta}{1 - \delta^{T-t}} \Phi - \frac{1 - \delta}{1 - \delta^{T-t}} \Psi \\ \Leftrightarrow w_{t+1}^W &= w_{t+1}^F - \frac{1 - \delta}{1 - \delta^{T-t}} \Phi + \frac{1 - \delta}{1 - \delta^{T-t}} \Psi\end{aligned}$$

Resubstituting this into the original difference equations leads to a system of two unjoined difference equations:

$$w_t^F = \frac{1 - \delta^{T-t}}{1 - \delta^{T-t+1}} \delta w_{t+1}^F + \frac{1 - \delta}{1 - \delta^{T-t+1}} [\Phi + (1 - \theta) \delta (\Psi - \Phi)] \quad (4.23)$$

$$w_t^W = \frac{1 - \delta^{T-t}}{1 - \delta^{T-t+1}} \delta w_{t+1}^W + \frac{1 - \delta}{1 - \delta^{T-t+1}} [\Psi + \theta \delta (\Phi - \Psi)] \quad (4.24)$$

Because the coefficients are not constant over time, each single equation will be solved by developing the series. Consider first equation (4.23) for $t = T - 1$,

$$w_{T-1}^F = \frac{1 - \delta}{1 - \delta^2} \delta w_T^F + \frac{1 - \delta}{1 - \delta^2} [\Phi + (1 - \theta) \delta (\Psi - \Phi)],$$

and insert this into the related equation for $t = T - 2$,

$$\begin{aligned}w_{T-2}^F &= \frac{1 - \delta^2}{1 - \delta^3} \delta w_{T-1}^F + \frac{1 - \delta}{1 - \delta^3} [\Phi + (1 - \theta) \delta (\Psi - \Phi)] \\ &= \frac{1 - \delta}{1 - \delta^3} \delta^2 w_T^F + \frac{1 - \delta}{1 - \delta^3} (1 + \delta) [\Phi + (1 - \theta) \delta (\Psi - \Phi)].\end{aligned}$$

Continuing, we get for each time t

$$\begin{aligned} w_t^F &= \frac{1-\delta}{1-\delta^{T-t+1}} \delta^{T-t} w_T^F + \frac{1-\delta}{1-\delta^{T-t+1}} [\Phi + (1-\theta)\delta(\Psi - \Phi)] \sum_{\tau=0}^{T-t-1} \delta^\tau \\ &= \frac{1-\delta}{1-\delta^{T-t+1}} \delta^{T-t} w_T^F + \frac{1-\delta^{T-t}}{1-\delta^{T-t+1}} [\Phi + (1-\theta)\delta(\Psi - \Phi)]. \end{aligned}$$

Equation (4.24) is symmetric, so that we immediately can conclude

$$w_t^W = \frac{1-\delta}{1-\delta^{T-t+1}} \delta^{T-t} w_T^W + \frac{1-\delta^{T-t}}{1-\delta^{T-t+1}} [\Psi + \theta\delta(\Phi - \Psi)].$$

CHAPTER 5

Bargaining over the effort function

5.1 INTRODUCTION

This chapter studies a possible role of trade unions in a shirking model. The shirking approach of efficiency wage theory usually studies the determination of wages and effort levels in a framework in which firms are assumed to have monopsony power in the labour market, while labour supply is competitive. If firms cannot monitor the effort of their employees perfectly, or simply find it too costly, they can raise the wage above its market clearing level and threaten their employees with layoff once they are caught shirking. The apparent attractiveness of efficiency wage models is that they are backed by the trinity of personnel economics "to recruit, retain, and motivate". As Lazear (1995, p. 70) puts it, "the story is appealing".

However, most efficiency wage models neglect the existence of unions. Partly, this might be for historical reasons, as the original intention of these models was to explain wage rigidities and involuntary unemployment in a perfectly rational world *without* unions (see e.g. Yellen 1984, Shapiro and Stiglitz 1984). Partly, the idea that a firm can exert monopsony power in the labour market only when facing a competitive supply side might have contributed to the tendency to ignore the role of unions.

Still, trade unions do affect the wage-setting behaviour in different ways. The most apparent is collective bargaining: the resulting contract wage has to be accepted as a minimum wage by the firm. But the scope of unions is usually even broader. They try to interfere in most points governing the relation between firms and their employees. Consequently, unions do not only negotiate wages but they also bargain over working conditions, work standards, layoff procedures, and sometimes even employment. Although this "bargaining over everything" might suggest otherwise, firms still have some freedom in determining the employment conditions. This freedom stems from at least two factors. First, in contrast to wages, the collective contract must necessarily be incomplete concerning the other items, because they are not easily controlled for and because the contract cannot possibly account for all circumstances in the different firms and workplaces. Second, in a world of incomplete information the union cannot behave like a monopolist governing labour supply at its own discretion, because it faces the same moral hazard problem as the firm.

The problem of incomplete contracts can be partially mitigated by bargaining on different levels, e.g. by collective bargaining over wages on a nation- or industry-wide level, and local bargaining over working conditions etc. on the level of individual firms or even plants. However, though the information problem is less severe on the local level, this two-tier bargaining raises an agency problem because the negotiators on the local level might pursue a different agenda than their counterparts on the centralized level.

The moral hazard problem is perhaps more severe. Although the union might attempt to act in the best interest of all its members, the latter cannot commit themselves to keep the contracts the union has signed on their behalf. If for example the union attempts to trade wages for employment (at a given effort level) the resulting wage-effort combination might just lead to shirking, and the firms will reject the wage cut. Consequently, if unions are to play a role in shirking models, they must be restricted by the no-shirking condition.¹

Yet even under this restriction there is room for unions. In the stylized framework of a shirking model in the tradition of Shapiro and Stiglitz (1984), unions could accept a given no-shirking condition and bargain over wages, or wages and effort, or they could attempt to influence the no-shirking condition in their favour. The (few) studies over unions in shirking models favour the first possibility. Bulkley and Miles (1996) for example study the behaviour of a monopoly union in a shirking model. Given the worker's effort function the union can either set wages, or wages *and* effort. In the first case, employers are still free to choose the effort level, and the resulting wage-effort combinations are therefore confined to points on the no-shirking condition. However, whether the union will pick a wage that is greater than the profit-maximizing wage the firm would choose, hinges critically on the assumptions about the union's utility function (see Goerke 1998, Bulkley and Miles 1998). In the second case, the union can achieve wages that are higher than necessary to achieve a certain effort level, or effort levels lower than what workers would maximally provide at a given wage. Rocheteau (2001) analyses the conditions under which either of these bargaining regimes prevail. In his model workers enjoy rents not only due to a moral hazard problem but also because recruitment is costly. A low unemployment level makes recruitment more difficult and improves the workers' bargaining power vis-à-vis the firm, which enables them to raise the wage above the level implied by the no-shirking condition.

Though these models attempt to integrate unions into an efficiency wage model, their results do not qualify for an explanation of the occurrence of wage gaps. They predict a wage gap pattern that is not only in line with the adaptation hypothesis,² but they also lead to the counterintuitive result that the wage gap will be high if the union's bargaining power is low. On the other hand, the more grip the union has on effort and wages, the less likely the effective wage will be greater than the contract

¹The no-shirking condition states that, given the wage w and a vector of parameters \mathbf{x} , there is a maximum effort level e^d that the firm can push through without provoking shirking, i.e., $e^d \leq e(w, \mathbf{x})$, with $\partial e / \partial w > 0$. The effort function reflects the same relation, but with sign of equality, $e^d = e(w, \mathbf{x})$.

²See also chapter 3.2 of this thesis.

wage.

Another drawback is that, if the union strives at an effort level greater than the level the firm would choose (naturally at a greater wage than the firm would like to pay, as the no-shirking condition has to be satisfied), the union will become responsible for the workers keeping their part of the contract. In other words, the firm could pass on the responsibility for monitoring effort to the union. It is questionable whether a union can accept this role without losing the support of the workers.³

Therefore, this chapter takes a different direction. Its basic idea is that unions do indeed try to influence the effort level, but that they bargain over the parameters *influencing the shape* of the effort function rather than over wage-effort combinations on or off a *given* effort function. This approach has two advantages. First, to a great extent it excuses the union from monitoring the workers. It can leave this responsibility to the firm, which itself is free—under certain restrictions—to set wages and effort at its will. Second, much of what workers' representatives do on the firm level can be interpreted as attempts to influence the shape of the effort function.

In order to illustrate this argument in the framework of shirking models consider the influence unions have on layoff procedures. An implicit assumption of most shirking models is that the detection of a shirker is synonymous with her being fired. Though this might be feasible in a country like the U.S., where employers are in principle free to decide about employment,⁴ labour legislation in European countries tends to curtail their property rights. An employer who wants to get rid of a shirker might find himself summoned to appear in court where he has to *proof* his case.⁵ Unions on the other hand often provide assistance to workers facing a dismissal, which drives another wedge between the probability that a shirker is detected and the probability that she is sanctioned.

The model presented in this chapter takes this stake of unions in layoff procedures into account by assuming that there are two different sanctioning probabilities and two different sanctions, both depending on the extent of shirking. The underlying idea is that a union may find it difficult to justify the protection of a worker who falls short of certain performance standards, while it will protect workers who come up to these standards but fall short of higher, "exaggerated" demands of the employer. This means that there is a performance standard, that is supported by the union against the workers. In the following, this standard will be called the "supported effort level". Another ingredient of the model is contract wage, which serves as a minimum wage, but which also determines the possible sanction of a

³Common wisdom is that the union acts as an agent of the workers (see e.g. Faith and Reid 1987). In this situation however, the union would assume the role of a principal. Indeed, a union accepting the task of monitoring would be much like the "unions" in the socialist countries of East Europe before 1989.

⁴This means that employers in the U.S. are free to exert their property rights. In principle, laying off a worker is then not much different from "laying off" capital (see e.g. Malcomson 1997).

⁵Anecdotal evidence suggests that things look even grimmer for firms. Franz (1994, p. 299) reports a German case where a shift-worker repeatedly appeared much to late at work. The judge nevertheless declared the dismissal of the worker null and void, because the employer was not able to proof that being late has led to a concrete disruption of the employee-employer relationship.

worker who meets the supported effort level but provides less effort than the employer demands. The union protects this "minimalist" by creating a link between the contract wage and the supported effort level. It resists a layoff of the minimalist, such that the employer can only sanction her by cutting her wage to the contracted level. This framework designates the union for a more passive role than models of effort-bargaining do. The union abstains from setting the effective wage and effort level, and leaves the agency problem to solve for the firm. Instead, the union sets (or bargains) a reference wage and a reference effort level, that determines in which situation the union gives free rein to the employer in dealing with shirkers. As it will be shown in this chapter, this approach provides a ready explanation for the occurrence of a wage gap. By raising the contract wage, the union also raises the utility of a sanctioned minimalist, which in turn forces the employer to raise the workers' moral-hazard rent by increasing the effective wage. This mechanism does not only make the occurrence of a wage gap likely, but it establishes a strong positive relationship between the contract wage and the size of the wage gap.

In contrast to the previous chapter, this chapter does not focus on the bargaining structure, in order to avoid any arbitrariness about the sequence of bargaining on the parameters of the effort function. Rather, we will confine ourselves to analysing the effects of these parameters on the effective wage and effort. The reason for this choice is that unions can influence the effort function on different levels. While contract wages are determined by collective bargaining on a more centralized level, other parameters like the supported effort level and the sanctioning probability can be determined on the centralized level or on the level of individual firms. In fact, in Western European countries like Germany or the Netherlands, these parameters are determined on both levels: collective agreements determine the general working conditions,⁶ while some supplementary issues, and the more practical issues are discussed on the level of individual firms. However, it will become clear that the model of this chapter can be extended easily to allow for bargaining on different stages.

The chapter is organized as follows. In section 5.2 the no-shirking condition is derived. The approach is essentially the same as that of Shapiro and Stiglitz (1984), but it accounts for the different types of sanctions. Section 5.3 shortly describes the behaviour of firms. Section 5.4 discusses the effects of the contract wage and the supported effort level in general, while section 5.5 focuses on the consequences for the behaviour of different types of unions. Finally, in section 5.6, the conclusion is drawn.

⁶In Germany for example, there are two types of collective agreements, the so called *Lohn- und Tarifverträge* (collective agreements on wages), and the *Manteltarifverträge* (collective agreements on working conditions). The latter are usually valid over a couple of years, while the former are typically negotiated once in a year.

5.2 THE EFFORT FUNCTION

The basic framework is essentially the same as that of the shirking model of Shapiro and Stiglitz (1984) but with two extensions: first, effort e is assumed to be a continuous variable, and second, next to the usual "nonshirker" and "shirker" a third type of workers is introduced, who supply a certain positive level of effort, which is however lower than the effort the firm requires.

All workers are identical and infinitively lived. Their instantaneous utility v is an additively separable function which is concave in income w and linear in effort e :

$$v(w, e) = u(w) - e \quad \text{with} \quad u(w) = 0, \quad u'(w) \quad \text{and} \quad u''(w) < 0 \quad (5.1)$$

The present value of the utility depends on the trajectory of the instantaneous utility $v(t)$,

$$V = \int_0^{\infty} \exp(-rt) v(t) dt,$$

where r is the rate of time preference.

We assume that the union negotiates a contract (\bar{w}, \bar{e}, p, q) with the firm. The first variable \bar{w} is a contract (or minimum) wage that the firm must not undercut. In exchange, the union promises to support an effort level \bar{e} against the workers: if they provide an effort less than \bar{e} , they face a probability of p to be laid off. On the other hand, if the firm demands an effort level of $e^d > \bar{e}$ but the worker provides less but at least \bar{e} , the union does not support the firm, and the probability of a sanction is only $q < p$. Moreover, as such a worker still provides at least the contracted effort level, the firm cannot fire the worker but has to continue employment paying at least the contract wage. Restricted by the contract the firm is free to set the terms of employment: it chooses a combination of required effort e^d and wage $w \geq \bar{w}$. Moreover, we assume that the firm is free to choose the employment level L .

In this situation, workers face the problem of maximizing their utility over the effort they provide. In order to simplify the analysis, we assume that for a sanction it does not matter by how much a worker undercuts either e^d or \bar{e} . For instance, whether she provides 99 % of the supported effort \bar{e} or no effort at all, the probability of a sanction is always equal to p , and the sanction will always be a layoff. Consequently, workers will effectively choose only from two or three effort levels, $e \in \{0, \min\{\bar{e}, e^d\}, e^d\}$. We will call a worker who does not provide any effort ($e = 0$) a shirker, a worker who provides only the effort level supported by the union ($e = \bar{e}$ if $\bar{e} < e^d$) a minimalist, and a worker providing exactly the effort level required by the firm ($e = e^d$) a nonshirker.

Consider first the utility of a nonshirker who obeys to the terms set by the firm and provides the effort level e^d and receives the wage w in return, as long as she is employed. Naturally, the nonshirker does not face any sanctions. Nevertheless, we assume that there is an exogenously given probability s that the worker loses her job, e.g. because the firm ceases to exist. The present value of the nonshirker's

lifetime utility V_N must satisfy the asset pricing equation

$$rV_N = u(w) - e^d + s(V_U - V_N), \quad (5.2)$$

where V_U is the expected lifetime utility of an unemployed worker, which is assumed to be exogenous for the moment. Equation (5.2) states that the nonshirker's value of being employed is equal to her instantaneous utility $u(w) - e^d$ while employed plus the expected loss due to the possibility of becoming unemployed.

A similar equation determines the expected lifetime utility V_S of a shirker. Unlike a nonshirker she provides zero effort but faces an additional probability p of being sanctioned and becoming unemployed. It is assumed that unemployed shirkers are not discriminated and have the same unemployment utility as a nonshirker who is laid off for structural reasons:

$$rV_S = u(w) + (s + p)(V_U - V_S) \quad (5.3)$$

There is no difference between a minimalist and a nonshirker as long as the firm requires an effort level of less than the supported effort \bar{e} . In this case, the minimalist will simply provide the required effort. If the firm requires a higher effort level, $e^d > \bar{e}$, a minimalist still faces the same probability of becoming unemployed, s , as the nonshirker, because she cannot not be sanctioned by being laid off. However, with probability $q < p$, the minimalist is sanctioned by a reduced utility of \bar{V} . Her asset equation is then

$$rV_M = u(w) - \bar{e} + s(V_U - V_M) + q(\bar{V} - V_M). \quad (5.4)$$

We assume that a sanctioned minimalist stays employed at the contract wage \bar{w} and that she has to continue providing the supported effort \bar{e} . Consequently, \bar{V} is determined by

$$r\bar{V} = u(\bar{w}) - \bar{e} + s(V_U - \bar{V}). \quad (5.5)$$

However, for two reasons \bar{V} might not be feasible. First, it might be less than the utility of a shirker, V_S , at $w = \bar{w}$. In this case, a sanctioned minimalist will reduce her effort to zero (and become a shirker), rather than continuing to provide \bar{e} . It is reasonable to assume that under such circumstances the firm will resist to carry on employing the sanctioned minimalist, while the union can hardly insist, knowing that the minimalist will become a shirker for sure. In order to avoid such a situation, \bar{w} and \bar{e} have to be chosen such that $\bar{V} \geq V_S(\bar{w})$, which is equivalent to

$$\bar{e} \leq \frac{p(u(\bar{w}) - rV_U)}{r + s + p}. \quad (5.6)$$

Second, a sanctioned minimalist must prefer staying with the firm to quitting. The related condition $\bar{V} \geq V_U$ is equivalent to $\bar{e} \leq u(\bar{w}) - rV_U$. It is therefore satisfied as long as $\bar{V} \geq V_S(\bar{w})$.

After defining the utilities of a nonshirker, a shirker and a minimalist the nonshirking conditions can be derived. In order to achieve the required effort level, the

firm must set the wage w and the required effort level e^d such that the workers prefer being nonshirkers to being shirkers or minimalists. Comparing equations (5.2) and (5.3) we see that the first condition, $V_N \geq V_S$ is equivalent to

$$e^d \leq \frac{p(u(w) - rV_U)}{r + s + p}, \quad (5.7)$$

which is precisely the usual no-shirking condition as in Shapiro and Stiglitz (1984). According to this condition, the worker is motivated to provide a higher effort level if the potential loss in utility by becoming unemployed, $u(w) - rV_U$, is increasing, or if a shirker faces a greater probability p of being sanctioned.

The second condition, $V_N \geq V_M$, is by equations (5.2), (5.4) and (5.5) equivalent to

$$e^d \leq \frac{q(u(w) + sV_U - (r+s)\bar{V}) + (r+s)\bar{e}}{r+s+q} = \bar{e} + \frac{q(u(w) - u(\bar{w}))}{r+s+q}. \quad (5.8)$$

This condition is novel because it does not depend on the unemployment utility but on the contract wage. Same as before, a worker will provide more effort the higher her wage or the sanctioning probability is. The same holds for a higher supported effort \bar{e} , because it decreases the utility of a minimalist, who by definition provides \bar{e} . A higher contract wage, in contrast, will decrease the provided effort as it makes the utility loss that comes with a sanction less severe.

The firm must obey the two no-shirking conditions (5.7) and (5.8). As a profit maximizer the firm will set the required effort such that they are satisfied with sign of equality, which results in the equation

$$e^d = \min \left\{ \frac{p(u(w) - rV_U)}{r + s + p}, \bar{e} + \frac{q(u(w) - u(\bar{w}))}{r + s + q} \right\} \\ = \begin{cases} \frac{p(u(w) - rV_U)}{r + s + p} & \text{if } u(w) < \frac{(r+s+p)(r+s+q)}{(r+s)(p-q)} \left(\bar{e} + \frac{prV_U}{r+s+p} - \frac{qu(\bar{w})}{r+s+q} \right) \\ \bar{e} + \frac{q(u(w) - u(\bar{w}))}{r+s+q} & \text{if } u(w) \geq \frac{(r+s+p)(r+s+q)}{(r+s)(p-q)} \left(\bar{e} + \frac{prV_U}{r+s+p} - \frac{qu(\bar{w})}{r+s+q} \right) \end{cases} \quad (5.9)$$

It is depicted in figure 5.1. It is a kinked function with a steep lower branch $e^d = \frac{p(u(w) - rV_U)}{r+s+p}$ reflecting the condition $V_N \geq V_S$ and a flatter upper branch $e^d = \bar{e} + \frac{q(u(w) - u(\bar{w}))}{r+s+q}$ derived from the condition $V_N \geq V_M$. The kink, the intersection point between the two branches, is at the point (w_0, e_0) with

$$u(w_0) = \frac{(r+s+p)(r+s+q)}{(r+s)(p-q)} \left(\bar{e} + \frac{prV_U}{r+s+p} - \frac{qu(\bar{w})}{r+s+q} \right), \\ e^d(w_0) = \frac{p}{(r+s)(p-q)} ((r+s+q)\bar{e} + qrV_U - qu(\bar{w})).$$

These two equations allow to derive a preliminary result about the effective wage w and effort e : combined with the feasibility condition (5.6) for the sanction for a minimalist, they become

$$u(w_0) \leq u(\bar{w}), \quad e(w_0) \leq \bar{e}.$$

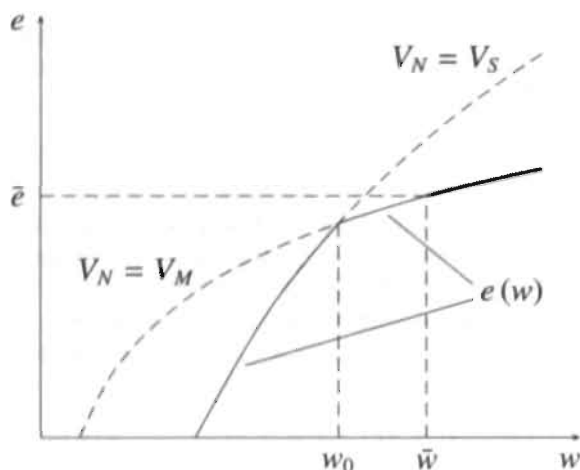


Figure 5.1: The effort function $e(w)$ consists of the lower branches of the two incentive compatibility conditions $V_N = V_S$ and $V_N = V_M$. Due to the contract wage restriction $w \geq \bar{w}$, feasible combinations of effective wage and effort are restricted to the bold part of the effort function: at the contract wage \bar{w} , workers will put forward no more than the supported effort level \bar{e} .

This means that, due to the minimum wage characteristic of the contract wage \bar{w} , not all wage-effort combinations on the effort function are feasible. Because $w \geq \bar{w} \geq w_0$ the effective wage must be on the upper branch of the effort function, which reflects the no-shirking condition $V_N \geq V_M$. The lower branch $V_N \geq V_S$ is not achievable, because it requires a contract wage that would turn a sanctioned minimalist into a shirker. The effective effort function is therefore

$$e^d = \bar{e} + \frac{q(u(w) - u(\bar{w}))}{r + s + q}, \quad w \geq \bar{w}. \quad (5.10)$$

Notice that although the effort function suggests otherwise, it is not perfectly insulated from the unemployment utility V_U , because the contract wage and the supported effort still have to satisfy the feasibility condition (5.6). Depending on the values of \bar{w} and \bar{e} , an increase in V_U can, but need not, require either an increase of the contract wage or a decrease of the supported effort level.

5.3 THE FIRM

Consider now the problem of a representative firm with a profit function

$$\pi = R(eL) - wL.$$

The variable L is the number of workers employed by the firm, and $R(\cdot)$ is the firm's short run revenue function in effective labour eL , with $R(0) = 0$, $R'(eL) > 0$, and $R''(eL) < 0$. The firm maximizes its profit over employment and the wage, under the restriction of the contract (\bar{w}, \bar{e}, p, q) :

$$(w, L) = \arg \max R(eL) - wL \quad \text{s.t. } w \geq \bar{w}.$$

The first-order conditions for a profit maximum are⁷

$$w : \begin{cases} e_w R' = 1 & \text{if } w \geq \bar{w} \\ w = \bar{w} & \text{else} \end{cases}, \quad (5.11)$$

$$\text{and } L : eR' = w, \quad (5.12)$$

with e_w being the first partial derivative of the effort function (5.10) with respect to the wage.

Consider first the firm's wage-setting behaviour. If there is an interior solution for the wage, i. e., the contract wage is not binding, the two first-order conditions can be combined to the famous Solow condition: if employment is at its profit maximizing level, the profit maximizing ("efficiency") wage is found where the wage elasticity of effort is equal to one,

$$e_w \frac{w}{e} = 1. \quad (5.13)$$

Geometrically, the Solow condition (5.13) describes the point where the iso-profit line of the firm is tangent to the effort function.⁸ This is the point where the cost of a unit of effort is minimized. Notice that the efficiency wage only depends on the effort function, while the revenue function and with it the output market or the firm's production technology have no influence on the wage. Hence, the firm uses the wage only to solve the moral hazard problem, while all other factors are captured by the employment decision.

The contract wage is binding if it exceeds the wage that maximizes profits. This is the case when the effort elasticity at the contract wage is less than one,

$$e_w \frac{w}{e} \Big|_{w=\bar{w}} < 1. \quad (5.14)$$

The labour demand curve of the firm is described by equation (5.12). Its total differential with respect to employment and the wage,

$$\frac{dL}{dw} = -\frac{e_w R' - 1}{e^2 R''} - \frac{e_w}{e} L < 0,$$

⁷These conditions represent maxima because R is a concave function.

⁸The total differential of the profit function is $\frac{d\pi}{dw} = 1/R'$ at any level of employment. According to equation (5.12), the marginal revenue at the optimal employment level is $eR' = w$. Hence,

$$\frac{d\pi}{dw} = \frac{e}{w}.$$

shows that employment is a decreasing function of the effective wage. Notice that the denominator of the first fraction on the right-hand side is equivalent to the first-order condition of a profit maximum with respect to the wage. Consequently, the fraction vanishes if the minimum wage restriction is not binding. On the other hand, if the contract wage is binding, it must be greater than the profit maximizing wage, and $e_w R' - 1$ is negative. Therefore, there is a negative relation between wages and employment for *all* feasible wages.

5.4 THE CONTRACT, WAGES, AND EMPLOYMENT

This section analyses the effects of the contract wage \bar{w} , the supported effort level \bar{e} and the sanctioning probability q of a minimalist on wages and employment. In general, two scenarios have to be distinguished, depending on whether or not the minimum wage restriction $w \geq \bar{w}$ is binding. This is the case when the effort elasticity at the contract wage is less than one. In order to see which parameter restrictions result from this condition, substitute the expression (5.10) into (5.14) and rearrange to get

$$\bar{e} > \frac{v'}{r+s+q} \bar{w} u'(\bar{w}). \quad (5.15)$$

Because the right-hand side is an increasing function of the contract wage, this condition implies that, given the supported effort level, \bar{w} must be below a certain level to become binding. The greater \bar{e} is, the greater is the threshold level of \bar{w} . Notice that this implies that the firm will pay a mark-up on the contract wage (i. e., there will be a wage gap) for high rather than low levels of \bar{w} . This preliminary result is quite remarkable because it contradicts the model of Bulkeley and Miles (1996), who predict just the opposite relation between contract and effective wage.⁹

5.4.1 The contract wage

Assume first that the contract wage is indeed not binding. Applying the effort function (5.10) to equations (5.37), (5.38), and (5.41) from the appendix,¹⁰ the influence of the contract wage on the effective wage, effort and employment is calculated as

$$\frac{dw}{d\bar{w}} = -\frac{u'(\bar{w})}{wu''(w)} > \frac{1}{\rho} > 1, \quad (5.16)$$

$$\frac{de^d}{d\bar{w}} = \frac{1-\rho}{\rho} \cdot \frac{qu'(\bar{w})}{r+s+q} > 0, \quad (5.17)$$

$$\frac{dL}{d\bar{w}} = -\frac{L}{e} \left(\frac{1-\phi}{\phi} + \frac{1}{\rho} \right) \frac{qu'(\bar{w})}{r+s+q} < 0, \quad (5.18)$$

where $\rho := -wu''/u' \in (0, 1)$ is the relative risk aversion of the worker with respect to income, and $\phi := -eLR''/R' \in (0, 1)$ a corresponding measure of the firm's revenue

⁹Compare section 5.1.

¹⁰The equations in this subsection are derived in general terms in the appendix 5.A of this chapter for a general model.

function. Equation (5.16) states that the contract wage does not only have a positive effect on the effective wage, but that an increase in the contract wage leads to an even greater increase in the effective wage.¹¹ Combined with the previous result that a wage gap will occur only with contract wages *greater* than a certain level, the positive relation between the wage gap $w - \bar{w}$ and the contract wage is in strong contrast to other models of the wage gap, which usually predict that for great \bar{w} the gap is decreasing or even nihil.¹² Although this behaviour seems to be remarkable, it is perfectly in line with the standard shirking model, but with unemployment benefits replaced by the contract wage: if the contract wage increases, so does the utility of a minimalist. In order to motivate its workers to provide the required effort instead of just the supported level, the firm has to increase the wage more than proportionally, because it cannot monitor effort perfectly. Figure 5.2 provides a graphical illustration.

According to equation (5.17) a greater contract wage also leads to a higher required effort level. Again, this is just a standard result of shirking models. A greater utility in case of a sanction shifts the effort function downwards, but pushes the wage up. The second effect outweighs the first, because the firm is forced to increase the worker's rent from moral hazard. See also figure 5.2. Due to the moral hazard problem, also the instantaneous utility of an employed worker is positively related to the contract wage,

$$\frac{dv}{d\bar{w}} = u'(w) \frac{dw}{dx} - \frac{dc^d}{dx} = \frac{r+s+q\rho}{\rho(r+s+q)} u'(\bar{w}) > u'(\bar{w}) > 0$$

(see equation 5.39).

As a result of the behaviour of effective wage and effort the cost per effort unit, $c := w/e$, increases with \bar{w} ,

$$\frac{dc}{d\bar{w}} = \frac{q}{r+s+q} \frac{w}{e^2} u'(\bar{w}) > 0$$

(compare equation (5.40)). Because effort and employment are imperfect substitutes in production, the increased cost of effort also leads to a lower labour demand (equation (5.18)) and consequently to lower profits,

$$\frac{d\pi}{d\bar{w}} = -\frac{qu'(\bar{w})}{r+s+q} LR' = -\frac{u'(\bar{w})}{u'(w)} L < -L < 0$$

(compare equations (5.42) and (5.11)).

¹¹In order to prove that $dw/d\bar{w} > 1$, consider first the term $-u'(w)/wu''(w)$, which is the inverse of the measure of relative risk aversion, ρ , with respect to the wage. As u'' is negative, the expression $-u'(w)/wu''(w) > 1$ is equivalent to $u'(w) > -wu''(w)$. We integrate both sides to get $u(w) > u(w) - wu'(w) \Leftrightarrow 0 > -wu'(w)$, which is true because $u' > 0$. But if $-u'(w)/wu''(w) > 1$, then also $-u'(\bar{w})/w\bar{u}''(\bar{w}) > 1$, because $\bar{w} < w$ and u is a concave function. \square

¹²See the discussion in chapter 3.

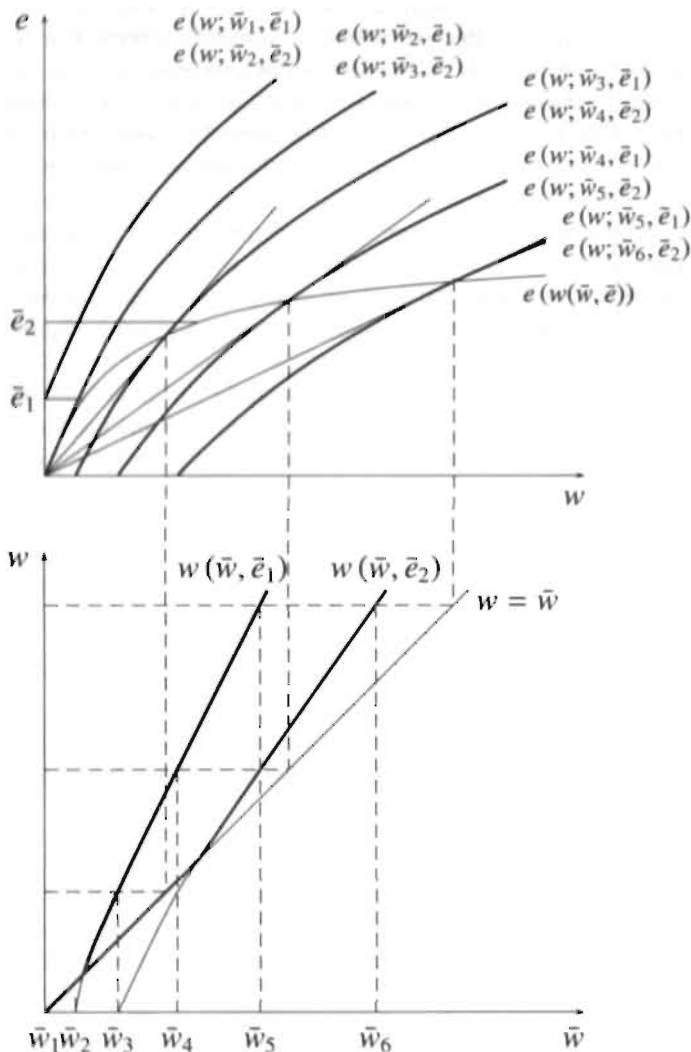


Figure 5.2: Effective effort e (upper graph) and effective wage w (lower graph) as functions of the supported effort level \bar{e} and the contract wage \bar{w} . In the upper graph, the efficiency wages are derived for effort functions at different levels of the contract wage and the supported effort level. The resulting relations between contract and effective wages are depicted in the lower graph. For clarity, the effort functions in the upper graph are drawn without their kink.

If the contract wage is below the threshold defined by equation (5.15), the effective wage is equal to the contract wage, $w = \bar{w}$: no wage gap occurs. This can, for example, be the case if the supported effort level is relatively high (see figure 5.2). Equations (5.16) to (5.18) then become

$$\begin{aligned}\frac{dw}{d\bar{w}}\Big|_{w=\bar{w}} &= 1, \\ \frac{de^d}{d\bar{w}}\Big|_{w=\bar{w}} &= -\frac{q}{r+s+q}u'(\bar{w}) + \frac{q}{r+s+q}u'(w)\Big|_{w=\bar{w}} = 0, \\ \frac{dL}{d\bar{w}}\Big|_{w=\bar{w}} &= \frac{1}{\bar{e}^2 R''} = -\frac{L}{\phi\bar{w}} < 0.\end{aligned}$$

As the effort function (5.10) suggests, the effective effort is always equal to the supported effort level at the contract wage. Consequently, the cost per effort unit increases proportionally with the contract wage, leading to a decreased labour demand. (Notice that $e_w R' - 1 < 0$ because the contract wage is greater than the profit maximizing wage.)

Because the required effort level is constant, the worker's rent of moral hazard is increasing by less than in the case where the contract wage is not binding. The reaction of the worker's instantaneous utility to a change of the contract wage is equal to

$$\frac{dv}{d\bar{w}}\Big|_{w=\bar{w}} = u'(\bar{w}) > 0.$$

Summarizing, a greater contract wage will always be beneficial for a worker who is protected from unemployment.

5.4.2 The supported effort level

If the contract wage is not binding, the required effort level \bar{e} has the following influence on the effective wage, effective effort and employment:

$$\begin{aligned}\frac{dw}{d\bar{e}} &= \frac{r+s+q}{qwu''(w)} = -\frac{r+s+q}{q\rho u'(w)} < 0, \\ \frac{de^d}{d\bar{e}} &= 1 + \frac{u'(w)}{wu''(w)} = -\frac{1-\rho}{\rho} < 0, \\ \frac{dL}{d\bar{e}} &= \frac{L}{e} \left(\frac{1-\phi}{\phi} - \frac{u'(w)}{wu''(w)} \right) = \left(\frac{1-\phi}{\phi} + \frac{1}{\rho} \right) \frac{L}{e} > 0.\end{aligned}$$

(Again, see the appendix 5.A for the derivation of these equations). The direction of the effects is just the opposite of the effect of a greater contract wage. A greater supported effort level decreases the utility of a sanctioned worker and shifts the effort function upwards (see also figure 5.2 on page 100). As a result, the firm has to grant a lower rent in order to prevent its workers from becoming shirkers, i. e.,

$$\frac{dv}{d\bar{e}} = u'(w) \frac{r+s+q}{qwu''(w)} + \frac{1-\rho}{\rho} = -\frac{r+s+q\rho}{q\rho} < 0.$$

Moreover, the cost per effort unit goes down, which results in a positive scale effect and makes labour demand increase.

If the contract wage is binding, the effects of an increased supported effort level are

$$\begin{aligned}\frac{dw}{d\bar{e}} \Big|_{w=\bar{w}} &= 0, \\ \frac{de^d}{d\bar{e}} \Big|_{w=\bar{w}} &= 1, \\ \frac{dL}{d\bar{e}} \Big|_{w=\bar{w}} &= \frac{1-\phi}{\phi} \frac{L}{e} > 0.\end{aligned}$$

In contrast to the case where the contract wage is not binding, the supported effort level has a positive effect on the effective effort (which is just equal to the supported effort), while the effect on the effective wage is naturally nihil. Nevertheless, the cost per effort unit decreases in \bar{e} , too, resulting in a greater labour demand, although this effect is weaker than in the unrestricted case. The effect on the utility of an employed worker will naturally be negative:

$$\frac{dv}{d\bar{e}} \Big|_{w=\bar{w}} = (u'(w) - e_w) \frac{dw}{d\bar{e}} \Big|_{w=\bar{w}} - e_{\bar{w}} = -1$$

Notice that an increase of the supported effort level leads to different results, if it makes the feasibility condition (5.5) bind and forces the contract wage up. In order to keep the analysis simple, we ignore this possibility and assume that the contract wage shows a greater inertia than the supported effort level.

5.4.3 The sanction probability

In case the contract wage is not binding, the effects of the probability q , that a minimalist is sanctioned, are

$$\frac{dw}{dq} = -\frac{(r+s)\bar{e}}{q^2 w u''(w)} = \frac{(r+s)\bar{e}}{q^2 \rho u'(w)} > 0, \quad (5.19)$$

$$\frac{de^d}{dq} = \frac{r+s}{r+s+q} \left(\frac{w u'(w)}{r+s+q} + \frac{1-\rho\bar{e}}{\rho q} \right) > 0, \quad (5.20)$$

$$\frac{dL}{dq} = \frac{r+s}{r+s+q} \frac{L}{e} \left(\frac{1-\phi}{\phi} \frac{w u'(w)}{r+s+q} - \frac{1-\rho\bar{e}}{\rho q} \right) \geq 0. \quad (5.21)$$

With a greater q being a minimalist becomes less attractive, which induces workers to provide any given effort level at a lower wage. As effort becomes cheaper, the firm will require more of it. As equation (5.20) indicates, the net effect on e^d is even positive which ultimately leads to a greater effective wage. Because of that, the effect on labour demand is indeterminate. A lower boundary for the derivative of employment can be determined by using expression (5.15), which determines when

the contract wage is binding. As dL/dq has been calculated for a not-binding \bar{w} , we must have

$$\frac{dL}{dq} > \frac{r+s}{r+s+q} \frac{L}{e} \left(\frac{1-\phi}{\phi} \frac{wu'(w)}{r+s+q} - \frac{1-\rho}{\rho} \frac{\bar{w}u'(\bar{w})}{r+s+q} \right).$$

Consequently, labour demand is more likely to be an increasing function of q , the more the wage is raised above the contract wage (as the wage gap indicates by how much effort has become cheaper), and the more risk averse the worker is at the effective wage.

Similarly, also the reaction of the worker's instantaneous utility,

$$\frac{dv}{dq} = \frac{r+s}{r+s+q} \left(\frac{r+s+q\rho}{q^2\rho} \bar{e} - \frac{wu'(w)}{r+s+q} \right) \leq 0,$$

depends on her risk aversion. The more risk averse she is, the less the wages and effort will react, and the less likely is a positive effect on utility. Applying condition (5.15) and using $wu'(w) > \bar{w}u'(\bar{w})$, it can be shown that the derivative of v has indeed a (positive) upper boundary:

$$\frac{dv}{dq} < \frac{(r+s)^2}{(r+s+q)^2} \frac{\bar{w}u'(\bar{w})}{q\rho}$$

If the contract wage is binding, a change of the sanction probability q will have no effect at all, because with $w = \bar{w}$ and $e^d = \bar{e}$ there is no difference between a nonshirker and a minimalist, and therefore no worker will be sanctioned. However, a decrease in q can of course suddenly make condition (5.15) binding, but in this case the behaviour of w , e^d and L is governed by equations (5.19) to (5.21) above.

Table 5.1 summarizes the effects of the contract wage, the supported effort level and the sanction probability on the effective wage, required effort and labour demand and the payoffs of workers and the firm.¹³

5.5 UNIONS AND THE LABOUR MARKET

This section analyses the preferences about wages, effort and employment of firms and three different types of unions. In order to make the results comparable, it is assumed that firms have identical, linear homogenous revenue functions in labour and capital. In this case, it does not matter whether L denotes the employment level of a single firm or the whole economy.

5.5.1 No unions

First, we consider the decision of a firm in the absence of unions. We assume that the probability p that a shirker is sanctioned, is exogenously given. It is the sanction

¹³Note that the greater sanctioning probability p is not considered in table 5.1 because it determines the part of the effort function that is assumed not to be relevant (see page 96).

	w	e	$c = \frac{w}{e}$	L	v	π
\bar{w}	> 1 (= 1)	> 0 (= 0)	> 0 (> 0)	< 0 (< 0)	> 0 (> 0)	< 0 (< 0)
\bar{e}	< 0 (= 0)	< 0 (= 1)	< 0 (> 0)	> 0 (> 0)	< 0 (= -1)	> 0 (> 0)
q	> 0 (= 0)	> 0 (= 0)	< 0 (= 0)	≥ 0 (= 0)	≤ 0 (= 0)	> 0 (= 0)

Table 5.1: Effects of the contract wage \bar{w} , the supported effort level \bar{e} , and the sanction probability q on the effective wage w , the effective effort e , the wage cost per effort unit c , employment L , workers' utility v , and profits π . In each column the first entry indicates the total derivative when the contract wage is not binding, while the second entry (in brackets) shows the total derivative for a binding contract wage.

probability that applies when a shirker is not supported by a union in a case against her. As the firm has no interest in protecting workers from being sanctioned (see table 5.1), the no-shirking condition is (5.7). Consequently, the effort function is

$$e^d = \frac{p(u(w) - rV_U)}{r + s + p}, \quad (5.22)$$

which is the previously unfeasible branch of the effort function (5.9). The firm will set the wage such that the Solow condition (5.13) is satisfied. Consequently, the effective wage solves

$$u(w) - wu'(w) = rV_U. \quad (5.23)$$

5.5.2 The insider union

The objective of the union in the shirking model will strongly depend on its degree of centralization. In the following, three highly stylized types of unions are distinguished: an *insider union* which is typically organized on the level of individual firms and which only covers employed workers, a *local union* which is organized at the level of one or a few firms and also covers the unemployment pool generated by the firm(s), and finally a *central union*, which covers all workers in the economy.

Consider first the insider union. Because it covers only workers who are currently employed, it will—if at all—only care about employment if the firm threatens to decrease its labour demand. However, if firms follow a last-in-first-out rule, workers know beforehand whether they are affected by this threat. Unless the median worker is affected, the union will maximize the expected lifetime utility of an employed worker,

$$U_I = V_N = \frac{u(w) - e^d + sV_U}{r + s}, \quad V_U = \text{const}, \quad (5.24)$$

restricted only by the shutdown condition of the firm, the worker's participation constraint, and a certain threshold employment level, that must not be undercut because otherwise the union would lose the support of its members.

Given the behaviour of effective wages, effort, and employment (see table 5.1), the agenda of the insider union is clear. It will increase the contract wage \bar{w} , decrease the effort level \bar{e} it supports against its members, and decrease the probability of a sanction against a minimalist worker q until the mentioned restrictions become binding. Because the union's objective is a linear function of the instantaneous utility of an employed worker, $v = u(w) - e^d$, the union will manipulate the shape of the effort function such that the firm will maximize its profits at the combination of effective wage and effort that the union desires. Due to the functional form of the effort function (see equation (5.10)), the levels of \bar{w} , \bar{e} , and q that solve the union's problem will not be uniquely determined but depend on each other.

Summarizing, the policy of the insider union—compared to the case without unions—will lead to lower productivity at higher wages, thus aggravating the inefficiency of the labour market.

5.5.3 The local union

The local union will care about the level of employment because of its unemployed members. In order to simplify the analysis, it is assumed that the union's objective is equal to the average utility of its members,

$$U_L = \frac{L}{N}V_N + \frac{N-L}{N}V_U, \quad V_U = \text{const.}, \quad (5.25)$$

even though not all the members face the same ex ante unemployment probability. The parameter N denotes the exogenously given number of members. The local union accounts for the direct employment effects of its behaviour. However, because the union is organized on a local level, it neglects the effect of L on overall unemployment. Therefore, it will treat V_U as a constant and ignore the side-effects of employment.¹⁴ As a result, the union will evaluate the trade-off between wages and employment incorrectly, putting too much weight on wages.

Consider the case of a monopoly union that has to decide about the contract wage \bar{w} and the supported effort level \bar{e} (at a given sanction probability q). The first-order condition for \bar{w} and \bar{e} is

$$\frac{dL}{dx}(V_N - V_U) + L \frac{dV_N}{dx} = \frac{1}{r+s} \left(\frac{dL}{dx}(v - rV_U) + L \frac{dv}{dx} \right) = 0, \quad x = \bar{w}, \bar{e}.$$

Using the results from section 5.4 and rearranging, it turns out that both first-order conditions are identical:

$$e = \frac{q(\rho + \phi - \phi\rho)}{\phi(r+s+q) + q\rho} (u(w) - rV_U) \quad (5.26)$$

The reason for this identity is that the contract wage and the supported effort level affect the effort function in a similar way, i. e., they are considered to be substitutes

¹⁴When discussing the central union, we will see that the expected lifetime utility of an unemployed also depends on the utility of an employed worker. We assume that the local union ignores this effect, and behaves as if V_U were a "black box".

with respect to their influence on the effort function. It is important to notice that equation (5.26) does not determine the effective wage and effort uniquely, but that it represents the local union's utility maximizing effort function. This does not mean that there are combinations of \bar{w} and \bar{e} , such that the real effort function (5.10) is equal to this "effort function" for all possible wages. Rather, the union sets \bar{w} and \bar{e} such that the effort function mimics equation (5.26) in the firm's optimum. In other words, the union sets the contract wage and the supported effort level such that the effective wage is determined as if the firm solved the Solow condition using equation (5.26). Consequently, the utility maximizing effective wage of the union solves

$$u(w) - wu'(w) = rV_U, \quad (5.27)$$

which leads to the same effective wage as in the case without union (see equation (5.23)). However, the resulting levels of required effort will typically be different. A comparison of the effort functions (5.22) and (5.26) reveals that the union will prefer a greater required effort than the non-unionized firm if

$$\phi < \frac{\phi \rho (r + s)}{(r + s)(p - q) + q\rho(r + s + p)}.$$

Whether this condition is satisfied depends on the values of ϕ , which measures how sharply the marginal value product of effective labour is decreasing, and ρ , the relative risk aversion with respect to workers' income. In general, the union will prefer a higher effort level if labour demand is very sensitive to wages. Thus, due to employment considerations the union might choose a higher effort level than in the no-union case, at the same wage. The reason is that, in contrast to the firm, the local union cares about the fate of the unemployed.¹⁵

Suppose now that the union is confined to contract wages that are too low to let a wage gap occur. In this case, the two first-order conditions for the contract wage and the supported effort level are

$$\begin{aligned} \phi \bar{w} u'(\bar{w}) &= v - rV_U, \\ \text{and } \frac{\phi}{1 - \phi} \bar{e} &= v - rV_U. \end{aligned}$$

¹⁵This would be one of the rare cases that a union decreases the inefficiency of the labour market. The positive effect of a higher effort level can be derived by analysing the firm's first-order condition for the profit maximizing employment level (see equation (5.12)), which states that labour input eL depends negatively on the wage cost per unit of effort, $R'(eL) = w/e$. If the wage set by the local union is equal to the wage in the no-union case, but effort is greater, then employment L must be greater under the local union. To see this, we differentiate the first-order condition:

$$\frac{dL}{de} = -\frac{R' + eLR''}{e^2R''}$$

So employment increases with effort if

$$R' + eLR'' > 0 \quad \Leftrightarrow \quad 1 > -eL \frac{R''}{R'} = \phi.$$

This is true for e.g. a short-run Cobb-Douglas production function.

Resolving the dependencies between these two equations reveals that the union will attempt to reach the same wage level as in the unrestricted case:

$$u(\bar{w}) - \bar{w}u'(\bar{w}) = rV_U \quad (5.28)$$

$$(1 - \phi)(u(\bar{w}) - rV_U) = \bar{e} \quad (5.29)$$

However, the chosen effort level might violate condition (5.15), which states that the supported effort level must exceed a certain level in order to make the contract wage binding.

5.5.4 The central union

Finally, the central union also maximizes the average utility of its members, but it is able to internalize the indirect effect of employment. Consider the asset equation of the expected utility of an unemployed worker,

$$rV_U = a + b(V_N - V_U). \quad (5.30)$$

As long as she is unemployed she receives an unemployment benefit, which gives her a utility of a . With probability b she finds a new job which provides her with a discounted lifetime utility of V_N . Because also employed workers face a probability s to become unemployed, the equations (5.2) and (5.30) can be solved simultaneously:

$$V_N^* = \frac{(r+b)(u(w) - e^d) + sa}{r(r+b+s)} \quad (5.31)$$

$$V_U^* = \frac{b(u(w) - e^d) + (r+s)a}{r(r+b+s)} \quad (5.32)$$

The asterisks indicate that the interdependency between V_N and V_U has been resolved.

Let for the moment L and N be employment and the number of workers in the whole economy. At each point in time sL workers are laid off for structural reasons and enter the unemployment pool, while $N - L$ workers are searching for employment. If unemployment is at its steady state, the reemployment probability must be equal to

$$b = s \frac{L}{N - L}. \quad (5.33)$$

The utility function of the central union can now be written as

$$\begin{aligned} U_C &= \frac{L}{N} V_N^* + \frac{N-L}{N} V_U^*, \quad b = s \frac{L}{N-L} \\ &= \frac{L}{N} \frac{v-a}{r} + \frac{a}{r}. \end{aligned} \quad (5.34)$$

The main difference between this objective and that of the local union (5.25) is that V_U is replaced by the unemployment benefit $a < rV_U$. Therefore, we can simply

reuse the results for the local union. The first-order condition with respect to the contract wage and the supported effort level is then

$$e = \frac{q(\rho + \phi - \phi\rho)}{\phi(r + s + q) + q\rho} (u(w) - a).$$

This "effort equation" prescribes a higher effort level than the corresponding equation (5.26) at any given wage.

Applying the Solow condition to the first-order condition yields the wage equation

$$u(w) - wu'(w) = a,$$

which implies that the central union prefers a lower effective wage than the local union (compare equation (5.27)). However, although the central union's "effort equation" lies above the equation of the local union, the effect of the lower wage dominates and the resulting effective effort will be lower than in the local union case.¹⁶ Nevertheless, this combination of lower wage and effort leads to lower cost per effort unit and therefore to a higher employment level.¹⁷

As in the case where the contract wage is not binding, the results for a binding contract wage can simply be adopted from the local unions. Equations (5.28) and (5.29) then become

$$\begin{aligned} u(\bar{w}) - \bar{w}u'(\bar{w}) &= a, \\ (1 - \phi)(u(\bar{w}) - a) &= \bar{c}. \end{aligned}$$

These equations compare in the same way to the corresponding equations of the local union as the equations for the not-binding contract wage do.

5.6 CONCLUSIONS

The model discussed in this chapter strongly supports the mark-up hypothesis of the wage gap. Unions can create a wage gap by setting a reference wage and effort

¹⁶To see this, combine the "effort function" and the wage-setting equation for both types of unions to get in both cases

$$e = \frac{q(\rho + \phi - \phi\rho)}{\phi(r + s + q) + q\rho} wu'(w),$$

which is to be evaluated at the respective union's optimum. Because $wu'(w)$ is increasing in w , the greater utility-maximizing wage of the local union must coincide with a greater effort level. \square

¹⁷The cost per effort unit is $c = w/e$. If evaluated at the profit maximizing wage, it is

$$c = \frac{(\phi(r + s + q) + q\rho)w}{q(\rho + \phi - \phi\rho)wu'(w)} = \frac{\phi(r + s + q) + q\rho}{q(\rho + \phi - \phi\rho)u'(w)},$$

for both types of unions at their respective optimal wage (see the previous footnote). The derivative of the cost with respect to the wage is

$$\frac{dc}{dw} = -\frac{(\phi(r + s + q) + q\rho)u''(w)}{q(\rho + \phi - \phi\rho)[u'(w)]^2} > 0.$$

Therefore the greater wage of the local union will lead to higher cost per unit of effort. \square

level. By protecting those workers from unemployment who fail to provide the effort the firm demands, but who still keep to the reference effort, unions dissolve the link between the effective wage and any outside option, and create a new link with the contract wage. Via this link, the contract wage does not only serve as a minimum wage, as in the model of Bulkeley and Miles (1996), but it becomes a powerful tool to influence the effective wage. By raising the contract wage, a union can increase the utility of a sanctioned worker, which in turn forces the firm to increase the effective wage. Because in shirking models workers earn rents due to imperfect monitoring, the firm will usually have to raise the effective wage by even more than the initial increase in the contract wage. In other words, the occurrence of a wage gap is not only likely, but a greater contract wage also leads to an increased wage gap. Moreover, it is shown that a wage gap will occur with high contract wage levels. In principle, the greed of the union or its members is therefore only restricted by employment considerations.

However, a wage gap will not occur in any case. Depending on the value of the parameters, it is possible that the firm maximizes its profits at a lower wage and effort level than the reference values set by the union. In this case, the contract wage serves as a minimum wage, and the effective effort is set equal to the supported level. Only then, unions are forced to set effective wages and effort directly. However, as is shown in section 5.5, this scenario results in the same effective wage as the scenario with a wage gap.¹⁸ This comes at no surprise, because ultimately it is the effective values unions care about.

As was argued in section 5.1, one drawback of direct negotiations over wages and effort is that the firm might shift some responsibility for monitoring the performance of workers to the unions, which could undermine their reputation with the workers. The wage gap scenario avoids this problem, while the union still has the ultimate control over the effective wage and effort. Nevertheless, the shirking model of this chapter is able to reproduce the result of Bulkeley and Miles (1996), that unionization does not necessarily lead to a reduction of effort. Instead, it comes to the same conclusion, that as long as a union cares about employment, the effective effort level depends on the workers' risk aversion and the shape of the firm's revenue function. Actually, unionization might lead to a higher effort level (and even to lower wages), such that the wage cost per effort unit is lower and employment is higher than in the case without unions. This is all the more remarkable, because it qualifies the claim of Walther (1989) and Goerke (2000), that effort is raised if shirkers are stigmatized and receive a lower unemployment benefit than nonshirkers who are laid off for structural reasons. In this chapter, the relation is just reversed: a sanctioned minimalist must necessarily receive a greater utility than an unemployed worker, but effort can still be greater than in a situation without a union, where unemployed shirkers are not discriminated.

Given the simplistic framework of the model, it is an appropriate question whether the assumed behaviour of unions is realistic. Of course, unions bargain over wages,

¹⁸Notice that the no-wage gap scenario never applies for a monopoly insider union, because it will prefer to set the supported effort level to zero.

but is there really something like a "supported effort level"? The introduction of this chapter has already hinted at the information problem involved in dealing explicitly with effort in collective bargaining on a centralized level. However, on the level of individual firms the representatives of unions usually do influence decisions about work standards and personnel. It is also on this level, where unions decide whether to back a worker whom the management wants to discipline. Hence, it is likely that a reference value of effort does indeed exist.

This suggests an interesting extension of the model. Instead of bargaining simultaneously over the contract wage and the supported effort level, negotiations could take place on two different levels: on a centralized levels, unions and employer federations bargain over the contract wage, while the supported effort level is determined on the level of individual firms. If this involves an agency problem, because the local branch of the union (and the employer) pursue different agendas than their counterparts on the centralized level, a two-tier bargaining process also would be able to uniquely determine the contract wage and the supported effort level, which is not always possible in simultaneous bargaining.

5.A ANALYSIS OF A PARAMETERIZED EFFORT FUNCTION

In this appendix the effects of the parameters \bar{w} , \bar{e} and q on wages, effort, employment and utility are derived in general terms. For this purpose the effort function is written as

$$e^d = e(w, x),$$

where x is one of the parameters \bar{w} , \bar{e} , and q . If the values of these parameters guarantee that the minimum wage is *not* binding, the usual Solow condition (5.13) determines the wage:

$$e_w(w, x) \frac{w}{e(w, x)} = 1 \quad \Leftrightarrow \quad we_w(w, x) - e(w, x) = 0 \quad (5.35)$$

Otherwise, the effective wage is simply equal to the contract wage,

$$w = \bar{w}. \quad (5.36)$$

The total differential of the equations (5.35) and (5.36) with respect to the effective wage w and the parameter x is¹⁹

$$\frac{dw}{dx} = \begin{cases} -\frac{we_{wx} - e_x}{we_{ww}} & \text{if } w \geq \bar{w} \\ \frac{d\bar{w}}{dx} & \text{if } w = \bar{w} \end{cases} \quad (5.37)$$

The parameter x has two effects on the optimal required effort level. On the one hand, it directly influences the shape of the effort function. On the other hand, it determines—via its effect on the wage—which point on the effort function is chosen. The overall effect is

$$\frac{de^d}{dx} = e_x + e_w \frac{dw}{dx} = \begin{cases} e_x - \frac{e_w}{we_{ww}} (we_{wx} - e_x) & \text{if } w \geq \bar{w} \\ e_x + e_w \frac{d\bar{w}}{dx} & \text{if } w = \bar{w} \end{cases} \quad (5.38)$$

Consequently, the effect of x on the instantaneous utility of an employed worker, $v = u(w) - e^d$, is given by

$$\begin{aligned} \frac{dv}{dx} &= u'(w) \frac{dw}{dx} - \frac{de^d}{dx} = (u'(w) - e_w) \frac{dw}{dx} - e_x \\ &= \begin{cases} -(u'(w) - e_w) \frac{we_{wx} - e_x}{we_{ww}} - e_x & \text{if } w \geq \bar{w} \\ (u'(w) - e_w) \frac{d\bar{w}}{dx} - e_x & \text{if } w = \bar{w} \end{cases} \end{aligned} \quad (5.39)$$

The effect of x on the cost per effort unit $c := w/e$ is interesting because it helps to assess the effect of the union on efficiency:

$$\frac{dc}{dx} = \frac{\frac{dw}{dx}e - w \frac{dw}{dx}}{e^2} = \begin{cases} -\frac{we_x}{e^2} & \text{if } w \geq \bar{w} \\ \frac{(\bar{e} - we_w) \frac{d\bar{w}}{dx} - we_x}{e^2} & \text{if } w = \bar{w} \end{cases} \quad (5.40)$$

The total differential of equation (5.12) describes the effect of x on the firm's employment decision:

$$\frac{dL}{dx} = -\frac{R' + eLR''}{e^2R''} e_x - \frac{L}{e} e_w \frac{dw}{dx} - \frac{e_w R' - 1}{e^2 R''} \frac{dw}{dx}$$

¹⁹Notice that the case $w = \bar{w}$ is not trivial because \bar{w} has to satisfy the feasibility condition (5.6).

Notice that $-eLR''/R'$ is a measure of the relative curvature of the revenue function—similar to the measure of relative risk aversion. Substituting $\phi := -eLR''/R'$ simplifies the equation above to

$$\frac{dL}{dx} = \begin{cases} \frac{L}{e} \left(\frac{1-\phi}{\phi} e_x + \frac{we_{wx} - e_x}{we_{ww}} e_w \right) & \text{if } w \geq \bar{w} \\ \frac{L}{e} \left(\frac{1-\phi}{\phi} e_x - e_w \frac{d\bar{w}}{dx} \right) - \frac{e_w R' - 1}{e^2 R''} \frac{d\bar{w}}{dx} & \text{if } w = \bar{w} \end{cases} \quad (5.41)$$

Finally, using equations (5.11), (5.12) and (5.37), the effect of x on the firm's profits can be calculated as

$$\begin{aligned} \frac{d\pi}{dx} &= (e_w R' - 1) L \frac{dw}{dx} + (eR' - w) \frac{dL}{dx} + e_x LR' \\ &= \begin{cases} e_x LR' & \text{if } w \geq \bar{w} \\ e_x LR' + (e_w R' - 1) L \frac{d\bar{w}}{dx} & \text{if } w = \bar{w} \end{cases} \end{aligned} \quad (5.42)$$

Part II

Surplus sharing and the wage gap

Why do wage gaps occur? This question actually comprises two problems: Why do employers and workers have an interest to determine a (contract) wage, although they know beforehand that this wage will be altered afterwards? and, Why are employers and workers not content with the contract wage, but settle on a different effective wage?

The answer given in the two chapters 6 and 7 in this part of this thesis is that a wage-setting system which provides for a predetermined, "fixed" base wage, but on the same moment allows to settle on a different effective wage, can be efficient. In the two chapters, two criteria are used for the efficiency of a wage-setting system. The first is, whether it avoids inefficient separations: as long as an employment relation generates a joint payoff that is greater than the sum of the outside options of firm and worker(s), the wage-setting system should provide for a wage that both parties prefer to their outside options. The second criterion is whether or not the wage-setting system induces firm and worker(s) to make optimal investments in the employment relationship instead of wasting resources outside the relationship.

The two chapters provide different, partly competing reasons for the efficiency of a wage-setting system with a predetermined wage that can be changed afterwards. Both chapters emphasize the importance of investments in the framework of an employment relation between a single worker and her employer. Chapter 6 focuses on the role of specific investments, i. e., investments which generate returns only within in the relation, but are worthless outside. In this case, both employee and employer are, on the one hand, interested in securing their returns from expropriation by the other side. By this argument, a fixed wage is beneficial to both parties. On the other hand, both parties are interested in maintaining the employment relation, because in the case of a separation they have to write off their investments. Hence, this argument just calls for a flexible wage. In chapter 6, a (re)negotiation game is proposed which carefully weighs the competing needs for stability and flexibility.

Chapter 7 basically uses the same framework, but it focuses on a different type of investments, which is commonly given more attention in the literature about wage formation. In the model of this chapter, the worker has the choice between investing effort either in the firm or in her outside option. It is argued that in this case only a wage consisting of a fixed base rate and a performance dependent component is able to provide the correct (efficient) incentives.

CHAPTER 6

Wage gaps, labour contracts and surplus sharing

6.1 INTRODUCTION

Why do employer federations and unions, individual employers and workers make wage contracts, when they know beforehand that the wage might be altered during the contract's period of validity? The alternatives would be either to make a contract that does not determine a wage at all, or to determine a wage and to forbid any change. However, such contracts are seldom found in reality, and, taking a heuristic approach, this could teach us that they are inferior to a contract that determines a wage that can be altered later on. But if this is true, in which way can the parties influence the process of *a posteriori* wage changes when making the contract, and how does this affect the wage gap?

This chapter attempts to give an answer to these questions by focusing on the relation between a firm and a single employee. It is argued that when they start a long term employment relation it is optimal to determine the distribution of the surplus to be generated *a priori*. This can be due to hold-up if specific investments matter, or to turn-over costs that a party incurs when ending the employment relation, and that make this party vulnerable to attempts to appropriate parts of its surplus. However, when time passes, economic circumstances might change in a way that one party prefers to end the relation if the wage is not adapted to the new situation. If an alternative distribution of the surplus could prevent this party from separating, a separation would be inefficient.

Hence, both parties have an interest to alter the wage rate under certain circumstances, but at the same time, they might prefer the process of alteration to obey certain rules, in order to avoid extortion and hold-up. The model in this chapter discusses a renegotiation game that has the desired features by warranting necessary adjustments of the wage rate to exogenous shocks. The wage gap is then the resulting difference between the wage originally written in the contract and the renegotiated wage.

Admittedly, by confining to this simplistic framework, the scope of this model is limited. As collective bargaining is left out of consideration, it might be desirable to replace the term "contract wage" by "contractual wage". However, wage gaps can also be observed in countries where there is no collective bargaining. In any case, the difference between the contractual wage at the beginning of a (contract) period

and the wage effectively paid in the run of that period has to be explained.

The chapter is organized as follows. Section 6.2 motivates an employment contract that weighs the advantages of wage rigidity and flexibility by determining a wage level and setting rules for changing it during the period of validity of the contract. In section 6.3, after introducing the time structure of the model, a renegotiation game is proposed that is in line with the contract. After exploring some general characteristics of the game, its outcome in the cases of complete and asymmetric information are derived in section 6.4. Section 6.5 examines the implications of the different scenarios for the contract wage and the wage gap. Section 6.6 deals with possible drawbacks of the model as an explanation of the wage gap. Furthermore, the theoretical predictions are contrasted with some empirical findings about wage behaviour in non-corporatist countries like the US. Finally, suggestions for further research are made.

6.2 WAGE CONTRACTS AND WAGE FLEXIBILITY

When a firm and a worker are making a contract, they have to take several issues into account. Next to agency and controllability problems, which we will neglect in the following, the fact that employment contracts are in general long-termed, has important implications.

First, it is likely that a positive net surplus is generated in the employment relation; the worker's marginal product is greater than the sum of the outside options of the parties to the contract. This positive net surplus can, for example, be due to investments the parties make in order to increase their benefit from the employment relation. The worker might spend time and effort to become acquainted with her colleagues and the organization she works for, she might move to the vicinity of her employer, and acquire firm specific human capital. The firm on the other hand might invest in equipment, even gear the place of work and the organization of work routines to the employee, and it might provide vocational training. All these investments have in common, that they do not only affect the payoff of the investor (and possibly the other party) *in* the relation, but are often worthless outside or even decrease the investor's outside option.¹

Another reason for a positive net surplus can be found in costs the parties have to bear when separating from each other and choosing an outside opportunity. As a consequence, their outside options *after* they have started the employment relation might be lower than before: employers for example often have to obtain permission to lay off workers from works councils, courts or even governmental institutions. And both parties might have to bear search costs after having separated.

Obviously, the parties to the contract covet the net surplus of the employment relation. In the case of specific investments, this makes the investor susceptible to

¹From there the term *specific investment*: the returns are only generated as long as the relation is maintained. Notice that in the case of the worker this type of investment must not be confused with effort. If a worker "invests" effort in her job, the returns accrue directly to the firm in the form of a higher productivity. The worker participates only indirectly in these returns by receiving a wage from the firm. Investments in effort are analysed in chapter 7 of this thesis.

attempts of the other party to appropriate a part of the returns. If the contract itself does not provide means to prevent this, the effect will be inefficiently low investments: hold-up occurs. As in general both parties benefit from these investments, and as hold-up might affect both, they have an interest to sign a contract which determines a wage *a priori*, which makes it difficult, if not impossible, to change it.

However, though a predetermined and rigid wage protects the parties' returns on investments, the long-term characteristic of the employment relation, that is a prerequisite for these investments, also makes flexibility desirable. If the contract prohibits any change of the wage, a party might find that her outside opportunities are more attractive than maintaining the employment relation; her outside option is binding. If in this case the sum of the parties' outside options exceeds the worker's marginal product, it is efficient if firm and employee separate. However, as long as still a positive net surplus is created, a separation would be inefficient, as both parties would benefit from changing the distribution of the surplus by choosing a new wage that makes neither outside option binding.

Apparently, there is a dilemma. If firm and employee agree on a contract that determines a wage and prohibits changing it afterwards, they are safe from extortion, and the returns of their specific investments are protected. They pay for this security with accepting the risk of inefficient separations. If in contrast the contract does not specify a wage or if it permits that the contractual wage is changed, inefficient separation can be avoided at the cost of hold-up and possibly costly clashes about the wage.

This dilemma can at least partially be resolved by making a contract that does not only settle a wage but also provides rules for changing it. If hold-up matters, it would be ideal to have a rule that allows for changes only in case they are necessary to avoid inefficient separations, and only to the extent that is necessary to avoid that a party's outside option becomes binding.

It is, of course, impracticable to include such a rule literally into an employment contract, because outside options are private knowledge. Firms and employees therefore have to rely on other arrangements. They can, for example, delegate the task of deciding about wage changes to a third party. As Teulings (1995) argues, collective bargaining might serve this purpose. As it—depending on the level of bargaining—accounts for aggregate shocks that affect outside option and marginal product, collective bargaining can reduce conflicts between a firm and its employees. Another possible remedy is to sign a contract that fixes the wage but also determines a period after which it has to be renegotiated. Finally, the contract might demand mutual consent as a prerequisite for any wage change. The ban of industrial actions during a period in which a contract wage is valid, would enforce this.² This is the starting point for the model presented in the following section.

²Two points should be noted here: first, in some countries, such as the United States, employers are explicitly allowed to announce a wage change unilaterally. The change is considered to be accepted even if the worker protests but continues to work (see Malcomson 1997). Second, firms and employees have other means than industrial action to put pressure on each other.

6.3 THE MODEL

6.3.1 Time structure of the model

In this and the following section we discuss the possible occurrence of a wage gap within a fixed period of validity of an employment contract between a firm and a single employee. For simplicity it is assumed that the parties' time horizon ends with this period; decisions taken after expiration of the contract, like over continuation of the employment relation or bargaining about a new contractual wage and working conditions, do not influence the decisions taken in the period of consideration. Moreover, no collective bargaining will occur in this period. Hence, any alteration of the wage can only be achieved by the parties to the contract themselves. The wage gap is thus defined as the difference between the wage written in the employment contract and the wage effectively paid during the period of validity of the contract.

The model consists of two periods, labelled 0 and 1, with period 1 being divided in a fixed number T of subperiods of equal length. In period 0, a firm and a worker meet in order to start an employment relation by making a contract that determines the terms under which work will take place. However, no work occurs in period 0, and when fixing the contract neither the firm nor the worker have exact information about the worker's marginal product and what their outside options will be when the contract comes to apply in period 1.³ What is important for the further analysis is that a surplus over the outside options might be created in the employment relation. This surplus can be due to specific investments or to costs occurring when a party takes her outside option.⁴ For reasons mentioned before, both firm and worker might prefer to agree upon the employment conditions before the contract comes to apply in period 1, instead of bargaining about shares of the surplus after all information has become available. The factors creating the surplus will not be modelled explicitly. Moreover, possible differences in work quality and quantity are not taken into account. The contract is therefore assumed to solely settle a wage rate w_0 for a constant amount of labour.⁵

In period 1 the employment contract comes to apply and the worker can take up her work. At the very beginning of this period, firm and worker learn their alternative opportunities, a_F and a_W , and the worker's marginal product m . These can either be common or private knowledge. Under complete information, i. e., if both the firm and the worker know the exact values of m , a_F and a_W , they will decide to end the employment relation immediately if it creates a negative surplus, $m < a_F + a_W$. Otherwise, if they believe that the relation does create a non-negative surplus, they continue the employment relation. However, knowing their outside options, a party

³See Hall and Lazear (1984) and MacLeod and Malcomson (1993) for models with a similar structure.

⁴In the case of specific investments, the assumption of incomplete information is not necessary. The superiority of a predetermined contract then results from the hold-up problem that would occur when the parties bargain a wage after the value of the relevant variables has been revealed. A predetermined contract protects the parties from being "ripped off".

⁵Instead of \bar{w} as in the rest of this thesis, the notation w_0 is chosen for the contract wage, because in this context it indicates the contractual wage in period 0.

The structure of the renegotiation game is as follows. As long as no party has taken its outside option, in each subperiod $t = 1, \dots, T$ of period 1, the following game takes place:

1. The proposing agent is selected randomly with probabilities θ (firm) and $1 - \theta$ (worker), with $\theta \in (0, 1)$.⁶
2. The agent selected at stage 1 makes a wage proposal w_t^X to take the place of the previously valid wage w_{t-1} (or w_0 , if $t = 1$), where the superscript X indicates a wage proposal of the firm ($X = F$) or the worker ($X = W$).
3. The agent who was not selected responds by one of the following actions:
 - accepting the proposal, which then becomes the new valid wage, $w_t = w_t^X$,
or
 - rejecting it, then the previously valid wage remains valid in this subperiod, $w_t = w_{t-1}$, or
 - ending the employment relation. In this case the renegotiation game ends, and firm and worker will earn their outside options, a_F and a_W , in this and all $T - t$ following subperiods.
4. If the responding agent has not ended the game, the selected agent decides whether to continue the employment relation (and the game) or to end it. Again, in the latter case, the firm and the worker will earn their outside options in this and all following subperiods.

If the responding agent has decided to continue, the worker produces her marginal product m and receives the wage w_t determined at the stages 2 and 3 in this subperiod. Accordingly, the firm receives $m - w_t$.

The game continues at stage 1 in subperiod $t + 1$, unless the end of period 1 (i. e., subperiod T) was reached.

The renegotiation game in a subperiod t is depicted in figure 6.2. Its structure is similar to the bargaining model of MacLeod and Malcomson (1993) and Malcomson (1997). However, there are two changes: first, in the mentioned studies both parties can decide whether they let work occur and produce the worker's marginal product before continuing the renegotiations in the subsequent subperiod, whereas in this game it is assumed that the parties have to let work occur if they have decided to continue the game. The reason is that the parties are not allowed to take industrial actions as long as the contract applies. Allowing for a deliberate choice about the occurrence of work would mean that the parties can put pressure on each other in

⁶The frequency at which a party is able to make a proposal is often interpreted as an expression for the party's bargaining power. The argument is that the more often a party makes a proposal the greater her influence on the bargaining outcome is. It will turn out, however, that in this special sequential bargaining game, the value of θ is irrelevant.

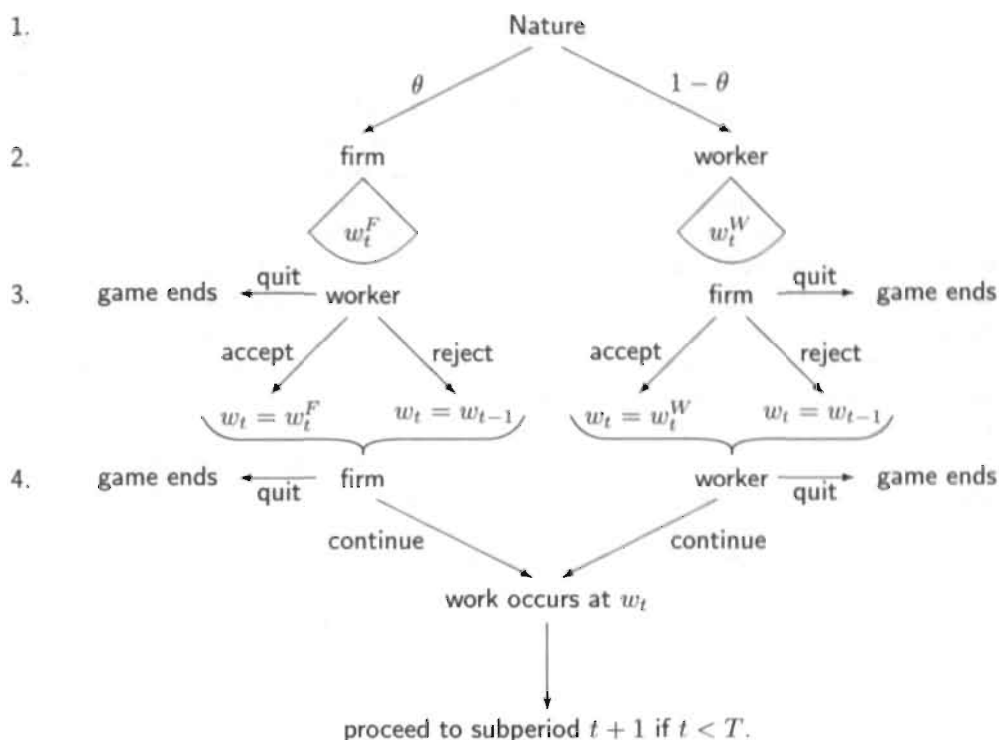


Figure 6.2: The stages of the renegotiation game in some subperiod t .

order to enforce a desired outcome.⁷ However, it was the intention of both parties to avoid this situation when making the contract.

Second, in the bargaining model of MacLeod and Malcomson (1993) and Malcomson (1997) there is nothing similar to stage 4. So in their model, a party cannot take her outside opportunities if she is selected to make a proposal. Malcomson (1997, p. 1928) argues that it is more plausible to assume that one party is always able to make a final offer before the other party quits for an outside option, instead of having one party making an offer and quitting for its outside opportunity if it is turned down without waiting for a counter offer.⁸ The reason why in the model of this chapter both parties—the one who makes an offer and the one who has to respond—are given the opportunity to end the game in each round is one of sym-

⁷Notice that industrial actions are in general forbidden during the period a wage contract applies. However, firm and worker might have other means to harm each other if there is a disagreement. The firm might, for example, announce short time work, and the worker might reduce her effort. Notice, that this gives the worker an advantage: reducing her effort will be relatively painless to her but harm the employer, whereas working short time harms both parties. Moreover, the firm often has to obtain permission for short time work from the works council.

⁸These alternatives reflect the two formulations of outside options in bargaining theory drawing back to Shaked and Sutton (1984) and Shaked (1994).

metry between the parties when they do not have the opportunity to refuse work. This argument can best be explained by example. Consider the renegotiation game without stage 4 and assume that one party is selected to make the wage proposals in all subperiods. Assume furthermore that the wage initially fixed in period 0 makes this party's outside option binding. Obviously, this party would be willing to quit if the wage is not renegotiated. However, without stage 4 the rules of renegotiations would not allow for this, and the (always) responding party can effectively resist any adaptation of the wage. The proposing party would be, as it were, taken hostage, and a high probability of being selected as proposing party—reflected by the parameter θ —would not be a sign of high bargaining power, as is the common interpretation, but turn out to be a sign of weakness. Moreover, it is not plausible to assume that there is a rule (implicitly) forbidding a party to quit at any time it wants to.

There are two characteristics which distinguish this renegotiation game (and that of MacLeod and Malcomson (1993)) from other non-cooperative bargaining models, drawing back to Rubinstein (1982). First, an explicit distinction is made between outside options and disagreement payoffs. As Binmore, Shaked, and Sutton (1989) show, both should not be confused if a bargaining situation is to be modelled adequately. Outside options become relevant for the game's outcome only if one party can credibly threaten to abandon the negotiations in favour of another opportunity, whereas disagreement payoffs accrue to the parties as long as no agreement is reached. Because they actually determine the cost of delaying an agreement they are by far more important for the outcome. The outside options, in contrast, just constitute restrictions to the bargaining outcome.

The second important characteristic of the renegotiation game is that when there is a disagreement the parties can always come back to the latest wage level they could agree upon in a previous subperiod (or the contract wage w_0 from period 0). Hence, this wage determines the disagreement payoffs in the game. As long as it leads to payoffs not less than the outside options, there is no cost of disagreement and the pressure to come to a mutual consent is virtually nihil.⁹

6.4 THE OUTCOMES OF RENEGOTIATIONS

6.4.1 Complete information

In the following the characteristics of the renegotiation game are presented verbally rather than in a formal way. Most of the results can be understood intuitively. The reader interested in the underlying mathematics can refer to the appendix of this chapter on page 135 ff. Throughout this section it is assumed, that the outside options a_W and a_F and the worker's marginal product do not change in period 1 (i. e., they are the same in all subperiods $t = 1, \dots, T$). However, as it is assumed that the renegotiation game continues even after an agreement has been reached,

⁹Only if it is assumed that worker and firm might refuse to let work occur, a disagreement is potentially painful.

it is straightforward to generalize the results derived in the following to multiple shocks (see also pp. 127).

In order to explore the outcome of the game the information available to the parties has to be taken into account. Assume first that they are fully informed about their own and their opponent's outside option and about the worker's marginal product.¹⁰ Obviously, they will end their relation at the earliest possible time when it becomes clear that the sum of the outside options exceeds the marginal product, $a_F + a_W > m$. In this case, no wage rate can be found that distributes the marginal product such that both parties prefer maintaining the employment relation to their outside opportunities (see scenario (b) in figure 6.3), and a separation is efficient. If however the employment relation generates a positive surplus, it is continued. The effective wage is then determined by the renegotiation game. As the contract wage w_0 determines the disagreement payoffs, it will co-determine the outcome of the game. Two scenarios can be distinguished.¹¹

First, if the contract wage is such that both firm and worker earn at least their outside options, $w_0 \geq a_W$ and $m - w_0 \geq a_F$, renegotiations cannot alter the wage. If in any subperiod a party makes a wage proposal in her favour, the responding party will reject at stage 3 of the game and the contract wage remains valid. Neither agent has an incentive to end the relation at stages 3 or 4. Thus, rejecting a proposal different from the contract wage is not costly, and the contract wage w_0 is retained throughout all subperiods in period 1. This outcome is depicted as scenario (a) in figure 6.3.

Second, if the contract wage makes one of the parties' outside options binding, i. e., either $0 \leq w_0 < a_W$ or $m - a_F < w_0 \leq m$, but still generates a positive surplus, $m \geq a_F + a_W$, both parties are interested in maintaining the employment relation given an adequate distribution of the surplus. However, if the party whose outside option is binding does not expect that the wage is altered, she would choose her outside opportunity at either stage 3 or 4 of the renegotiation game. As this would mean an inefficient separation, the other party has an incentive to give way to a wage change. The circumstance that the disagreement payoffs are still determined by the contract wage strongly forces the outcome of the renegotiation game to be close to the original wage. Whereas the party whose outside option is binding would effectively face a loss compared to his outside option during a disagreement, the other party receives more than she can expect after the agreement is reached. As shown in the appendix (p. 139) the outcome of the renegotiation game converges to the contract wage w_0 , restricted by the outside option under consideration, with increasing number of subperiods T . In this case, the other party becomes the residual claimant; the wage is set at the level of the binding outside option. This situation is depicted by the scenarios (c) and (d) in figure 6.3.¹² If the number of remaining

¹⁰The case of asymmetric information will be discussed in the following subsection.

¹¹The results referred to in the following are derived in the appendix. It should be noted they only hold if firm and worker have equal time preferences. If this is not the case, multiple equilibria are possible. Furthermore, it is assumed that the parties are risk neutral.

¹²Notice that this result does not hold if the parties can decide whether the marginal product is produced (and a wage is paid) during disagreement. Even if the decision about work is not

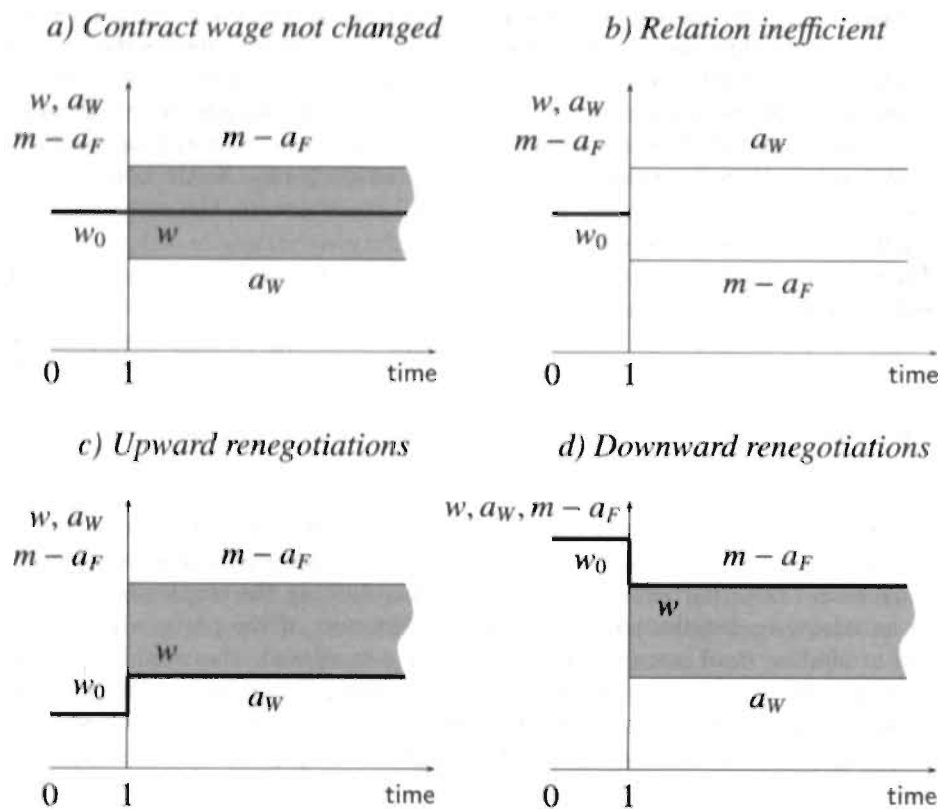


Figure 6.3: Contract and effective wage under complete information. w_0 is the contract wage, that was fixed in period 0 (before work can occur), w is the effective wage in period 1, $m - a_F$ is the wage that makes the firm's outside option binding, and a_W is the corresponding wage for the worker. The case where $w_0 < 0$ or $w_0 > m$, is not depicted.

subperiods is relatively small, the altered wage might provide a payoff well above the outside option to the respective party.

The results so far were derived for the case that the worker's marginal product m and the outside options a_F and a_W remain constant throughout the whole period 1. However, the longer period 1 lasts, the more likely the values of these variables will change over time. Apparently, in this case the outcome of the renegotiation game will be qualitatively the same if the parties do not anticipate future developments of m , a_F and a_W , and if after their values have changed the disagreement payoffs are no longer determined by the—now obsolete—contract wage, but by the latest wage rate both parties have agreed upon. Then, the effect is again only altered if it makes a party's outside option binding. However, even if the occurrence of future changes is anticipated, the predicted wage pattern is not necessarily different, if the number of subperiods between the anticipated shocks is sufficiently large. The reason is that in each subperiod before such a shock occurs, the responding party at stage 3 can successfully resist any change of the current wage rate as long as neither outside option is binding. If however the current wage makes an outside option binding the ~~wage proposal will still average out that the responding party just receives her~~ current outside option. The reason is again that the disagreement payoffs are always determined by the latest wage the parties could agree upon. However, after a shock has occurred, the wage rate will be adapted immediately.

Figure 6.4 shows how under complete information the effective wage might develop if the outside options change over time. As long as neither outside option is binding, the wage rate remains constant. If however one agent's share of the surplus drops below his outside option, the wage rate is changed in the way that this agent will earn precisely his outside option. It should be noted that this adaptation process is somewhat idealized. The depicted co-movement only results if the renegotiating parties believe the change of the outside option to be permanent and unique, and if there is a sufficiently large number of subsequent subperiods. Otherwise, the renegotiated wage might provide the party whose outside option is binding with a considerable share of the net surplus. However, once the wage is renegotiated, it will remain constant until an outside option becomes binding again. So the wage might not slightly glide with the outside options but perform jumps. But even when allowing for these implications the most striking implication of the model—the stickiness of the wage rate as long as no outside option is binding—remains intact.

(ab)used for putting pressure on the other party, a party might be worse off when work occurs at the contract wage than when no work occurs, i. e., either $w_0 < 0$ or $w_0 > m$. As no work will occur until an agreement is reached, the contract wage is irrelevant for the outcome of the renegotiation game. Hence, it becomes similar to a game of the Rubinstein type. The parties agree upon a wage that divides the net surplus according to the probabilities θ and $1 - \theta$ that either the firm or the worker is selected to make a wage proposal in a certain subperiod. The outcome is only restricted by the outside options. The parties can prevent such a situation by agreeing on penalties a party has to pay if she refuses to let work occur. If the penalty is sufficiently high, work is preferred to no work even in this case, and the outcome would be the same as if $w_0 \geq 0$ and $w_0 \leq m$.

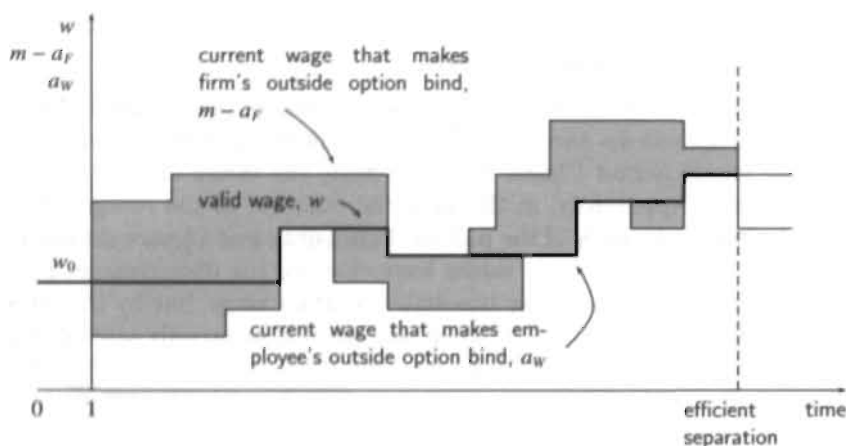


Figure 6.4: Development of the effective wage if the worker's marginal product m and the outside options a_F and a_W change over time. After MacLeod and Malcomson (1993).

6.4.2 Asymmetric information

The results of the previous subsection were derived under the assumption of complete information. Both agents know their own and their opponent's outside option as well as the worker's marginal product. As can be expected in this case, the resulting wage is always efficient in the sense that no inefficient layoffs or quits will occur. Though complete information is a common assumption in most models of bargaining, it is hardly found in real-life bargaining situations. The worker will not have full information about her value in the firm, m , and the firm's outside option, a_F , whereas the firm will lack full information about the worker's outside option, a_W . Rather, they will hold beliefs about these variables.

As the analysis of the renegotiation game with multiple periods is extremely difficult, only the wage proposals firm and worker would make in a one-period game are derived in the appendix on page 140. As usual, these proposals are similar to those that would occur with monopsonistic or monopolistic wage setting; both parties determine their wage proposal such that their expected payoff is maximized, taking into account the (subjective) probabilities that the opponent might reject the proposal or even end the employment relation and choose their outside opportunity. Under the most simplifying assumptions, the resulting proposals are

$$w^F = m - a_F - X_F, \quad w^W = a_W + X_W,$$

where X_F and X_W are non-negative terms reflecting the parties' beliefs about their opponent's outside option (and the marginal product) and about the opponent's belief about the parties' own outside option (and the marginal product). So both

parties propose wage rates that consist of their outside option plus a mark-up, taking into account that these proposals might induce their opponent to end the relationship.¹³ Formally, this result is equivalent to the two non-bargaining scenarios in the model of Hall and Lazear (1984), and same as in their model, this wage-setting behaviour can lead to inefficient quits and layoffs.

This inefficiency is not lessened when looking at the renegotiation game with more than one subperiod because the parties are not able to obtain more information and update their beliefs in the course of time. As work can occur at the contract wage or at a wage the parties agreed upon in the previous subperiod even if a proposal is rejected, provoking a disagreement by proposing a new wage or rejecting a proposal is costless and cannot serve as a credible signal that one party's outside option is really binding.¹⁴ The only situations where information is revealed are when a party accepts a proposal or when the employment relation is terminated by either a layoff or a quit. In both situations this information is of no use.

On this background the wage proposals in subperiod 1 of a multi-period renegotiation game will obey the same rules as those calculated for the one-shot game. It is important to note that these are no equilibrium proposals in the sense that they are always acceptable. Whether they will be accepted depends on the beliefs of the responding agent. It is possible that there is a continuous disagreement in all subperiods of the renegotiation game without one party separating. In addition, even if it was efficient to alter the wage due to a change of one party's outside option (or the worker's marginal product), an according proposal might be rejected because the responding party cannot verify the claim.

As a result, wages will be even more inert under asymmetric information than they are already under complete information. Furthermore, inefficient quits and layoffs can occur.

6.5 THE WAGE GAP AND WAGE CONTRACTS

The opportunity to alter a contractual wage during renegotiations and the information available during these renegotiations will in turn have an effect on the contract wage itself. The purpose of renegotiations is to adapt the wage level to unpredictable exogenous shocks that affect the parties' outside options and the marginal product. If there is complete information in period 1 this enables the parties of an employment relation to avoid inefficient quits and layoffs—the relation is continued as long as it yields a positive (net) surplus. Though determining the parties' payoffs, renegotiations are not a device for rent sharing.¹⁵ This happens when the wage contract

¹³Only if a party's wage proposal is equal to the wage that makes his outside option binding, this party can be sure that his opponent will not end the employment relation unless it is inefficient anyway.

¹⁴This characteristic distinguishes this game from the multistage bargaining game of Sobel and Takahashi (1983). In their model, rejecting a proposal is a costly signal because work does not occur in disagreement.

¹⁵Recall that the outcome of the renegotiation game is in general independent of the distribution of bargaining power, reflected by the probability that a party is selected to make a proposal in the

is fixed in period 0 before the employment relation actually starts.

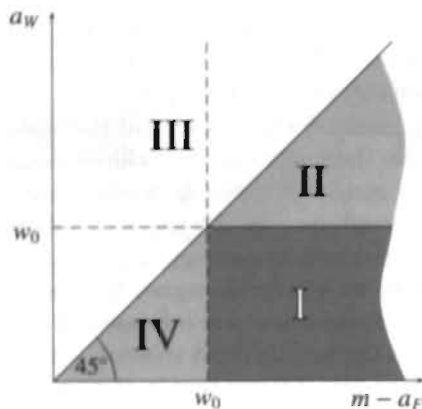


Figure 6.5: Outcomes of the renegotiation game in the $(a_W, m - a_F)$ space, depending on the level of the contract wage w_0 . The size of the different areas is correlated with the probability that the values of $(a_W, m - a_F)$ are within this area.

Source: Hall and Lazear (1984), Teulings (1995).

Suppose that the outside options a_F and a_W and the marginal product m do not change in the run of period 1. If their values are not known when firm and worker settle a wage contract, the parties are faced with the following problem, depicted in figure 6.5. For each possible level of the contract wage there are combinations of a_W and $m - a_F$ (area I in the figure), for which it does not have to be renegotiated, and combinations that make an adaptation of the wage rate necessary (areas II and IV). In area II upward renegotiations and in area IV downward renegotiations would be necessary in order to avoid an inefficient separation. Area III finally comprehends combinations for which continuing the relation is inefficient no matter at which level the contract wage is set.¹⁶

Firm and worker will be especially concerned about the occurrence of combinations of a_W and $m - a_F$ in the areas II and IV. If there is complete information in period 1, the firm will earn its outside option in area IV and the worker will earn her outside option in area II, and their respective opponent will become residual claimant. As the danger of inefficient separations does not exist (the size of area III does not depend on w_0), the firm strives at a contract wage as low as possible in

renegotiation game. There is, however, one special case where this is not true: if the contract wage makes a party's outside option binding ($w^C \notin [a_W, m - a_F]$) and if either $\theta = 0$ or $\theta = 1$ (one party can make all the proposals). As the contract wage cannot be maintained without making the party whose outside option is binding separate, the ever proposing party can set a wage such that the other party just receives her outside option, no matter whose outside option was binding at the contract wage.

¹⁶See also Teulings (1995).

order to avoid a situation as in area IV, whereas the worker has just the opposite aim. At the end, the influence the contract wage has on the probabilities to end up in area II or IV might be more important in bargaining than the actual level of the contract wage.

The parties' motives are different when there is asymmetric information in period 1. Because it is not warranted that the wage is adapted for combinations of a_W and $m - a_F$, represented by the areas II and IV, inefficient separations might occur. Therefore, next to getting a share of the surplus as big as possible, *both* parties have the common interest to find a contract wage that suffers no pressure from renegotiations. Thus, they will jointly strive at maximizing area I, and the wage contract will be less influenced by the distribution of bargaining power than with complete information.

These arguments still enforce the predictions concerning the wage gap that can be derived from the discussion in the previous section. With complete information, renegotiations ensure that the wage is adapted whenever a party's outside option becomes binding. This allows for rather extreme levels of the contract wage, which would just make a correction—possibly reckoned on—necessary. With asymmetric information, wages are more inert in period 1. As a consequence, firm and worker try to stabilize their relation by avoiding to end up in a situation where one of their outside options becomes binding. If they are successful, there will be even less pressure on the wage. The occurrence of a wage gap should therefore be the more likely the more information about his opponent a party has available.

6.6 CONCLUSIONS

This chapter attributes the occurrence of a wage gap to the need to adapt wage contracts to changing economic circumstances in order to avoid inefficient quits and layoffs. A renegotiation game is discussed that can provide the needed wage flexibility without unnecessarily undermining the original wage contract. However, to which extent renegotiations can serve this purpose depends on the information available to firm and worker. The outcome will be the more efficient the more accurate the agents' beliefs over their opponent's outside option are. As Hall and Lazear (1984) show, avoiding inefficient separations is impossible in the absence of full information. This has to be recognized when the wage contract is settled before the labour relation starts.

The model combines characteristics of the early Phillips-curve models of the wage gap in the tradition of Hansen and Rehn (1956), Gerfin (1969) and others, and of the bargaining models drawing back to Holden (1988, 1989, 1998), which have gained interest recently. As in the latter models, wage gaps are the result of deliberate renegotiations between a firm and its employees. In contrast to these models, wage contracts allow for renegotiations not in order to bargain about surplus sharing for a second time, but to avoid inefficient quits and layoffs. Therefore, the predicted wage pattern is more in line with the early Phillips-curve models of the wage drift and the adaptation hypothesis. By focusing solely on labour demand, these models however

only provide an explanation for the need to adapt wages, but not for the actual adaptation process. To this respect, this chapter has provided a microeconomic foundation for the Phillips-curve theory. However, its scope goes beyond this. It also focuses on the mutual dependency between contract and effective wage, which has been disregarded not only by the theories mentioned above.¹⁷

An obvious objection against the model as an adequate explanation for wage gaps is that it does not make a difference between effective wages above or below the contract wage on theoretical grounds. In other words, there is no downward wage rigidity. The wage gap, however, is *per definitionem* a positive difference between contract and effective wage.

However, there are arguments that tone down this objection. First, wage drift (and the wage gap) occupied economists mainly in the 1950s and 60s, a period that was characterized by rapid wage increases. Thereafter, when this development moderated, the interest waned remarkably. It is therefore possible that for historical reasons it has been overlooked that part of the wage gap might be due to mechanisms that can also work towards the other direction, an effective wage lower than the contract wage, too.

On this background it is worth noticing that effective wages below what has been initially agreed upon in a labour contract are not just an esoteric theoretical result. Empirical studies of Baker, Gibbs, and Holmstrom (1994) and McLaughlin (1994) of wage dynamics in the US suggest that there is a significant share of workers staying with the same employer while experiencing nominal wage decreases. When looking at real wages instead this share is naturally even greater. Moreover, the reported share of employees who receive nominal wage increases of *precisely* zero stands out. As is predicted by the model presented in this chapter, there obviously *are* forces that resist any wage change.

It should be noticed that such patterns might be specific for non-corporatist countries and especially for the US.¹⁸ However, wage increases are still more common than decreases. But taking into account the characteristics of employment contracts the model can account for that. Agreements about remuneration are normally settled in nominal terms, mostly for reasons of controllability and enforceability. Though indexing the wage is theoretically possible, it often turns out to be impracticable in reality. Next to control and information problems the parties would have to agree upon the relevant price index. As time passes, inflation does not only erode the purchasing power of wages, but also effects the agents' alternative opportunities and

¹⁷In the Phillips-curve model as well as in the efficiency wage models of Schlicht (1992) and Gahlen and Ramser (1987), or in the search model of Kleinhückelskotten and Spaetling (1980), the contract wage is taken as an exogenous variable. Moreover, the "contract wage" just serves as some kind of benchmark in these models and could easily be replaced by "previously valid wage", which could as well be the previously valid effective wage. In contrast, it is in principle possible to determine the contract wage in the models of local and central bargaining (see e. g. Holden 1998).

¹⁸Malcomson (1997) reports an asymmetry in labour relations in the US: if an employer announces a wage change, and an employee continues to work, be it under protest, the modification is assumed to be effective. If in contrast an employee demands a change, and the employer continues the employment relation, the claim is not supposed to be accepted. Consequently, the employer can act as a monopsonist. However, in long-term employment relations, employees are better protected.

the worker's marginal product in nominal terms. Employees are aware of rising wages elsewhere, and space for raising money wages is created by increasing product prices. It is therefore more likely that for a given wage rate the employee's outside option becomes binding rather than the outside option of the employer, resulting in a positive gap.

Moreover, there might be an information problem that causes downward rigidity. As was shown, the effective wage has a tendency to stick at the level prescribed in the contract if the parties are not fully informed about their opponent's outside option. As Hall and Lazear (1984) and the discussion of the asymmetric information case in this chapter have shown, it is in general impossible for a party to credibly reveal her outside option to the other party. Whether renegotiations can lead to necessary adaptations of the effective wage levels therefore depends on—possibly biased—information from other sources. It then seems plausible to assume that a firm's management has better information about the, at least *average*, outside option of its employees than a single employee has about her marginal product and the firm's outside option: a firm in general employs several workers, and information about the average wage levels for certain professions and branches are frequently published. Workers, on the other hand, have to deduce information about their job's value in the firm from rather vague indicators as the number of incoming orders and the general economic situation.

Other reasons for a downward rigidity of nominal wages might be found in psychological factors. Kahnemann, Knetsch, and Thaler (1986) report that most people consider it to be unfair if an employer takes advantage of an increased outside option. As, unlike the situation in the US, wages often can only be changed by mutual consent, employees might resist any attempts to cut wages even if such a behaviour is not optimal in view of the renegotiation game presented in section 6.3. Moreover, even an employer who can behave as a monopsonist cannot dismiss fairness considerations, because employees might express their discomfort with a lowered wage with a decrease in effort. If an efficiency wage argument plays a role in labour remuneration, this can be costly for the firm while causing virtually no cost for the employees. Both, the mentioned information problem and the psychological factors would lead to a situation in which the wage is raised (and a mark-up on the contract wage is paid) when upward renegotiations are necessary because the worker's outside option has become binding, whereas in the opposite case, when the firm's outside option becomes binding, downward renegotiations fail, and the employment relation is—inefficiently—dissolved by the firm.

Though the model of this chapter provides an explanation for the occurrence of a wage gap, it must not be overlooked that the assumed institutional setting is rather extreme, especially with respect to the renegotiation game. There are several factors that would flaw such renegotiations in reality. First, it is hardly possible for a firm to organize them, on a regular basis, for each employee. Furthermore, bargaining is rather costly as it consumes time and resources to obtain and process necessary information. Both parties would gain if at least the employees could delegate the task to central institutions. With renegotiations on the firm level between management

and a works council the predicted outcome will not change substantially, but might at least partially avoid the efficiency problems associated with insufficient information. The multistage wage formation process that has evolved in corporatist countries might be explained on this background. At the beginning of an employment relation, management and employee bargain about sharing a surplus. As time passes, the relation is subject to external shocks. These can be aggregate shocks affecting a whole business or even the whole economy, or idiosyncratic shocks on the firm level. Firm level renegotiations might fail to properly account for the former. Again, insufficient information might play a role. As Teulings (1995) argues, the gap could be filled by collective wage agreements.

Other points that deserve further analysis are the implications for employment and the investment decisions of both parties. In principle, both factors can be used to manipulate the net surplus or the outside options in order to achieve a greater payoff. The firm, for example, could manipulate the parameters decisive for the renegotiation outcome through its choice of the employment level. By hiring more workers, the firm could decrease its employees' marginal product and claim wage concessions. Because workers are interested in maintaining the employment relation as long as a positive net surplus is created, they might be willing to accept a lower wage in case the firm's outside option becomes binding. That this kind of behaviour on the part of firms is not observed might be due to the arguments put forward in favour of downward wage rigidity. Moreover, such a behaviour will undermine the firm's reputation and—if specific investments play a role—lead to hold-up on the side of the workers.

Workers, on the other hand, might have an incentive to direct parts of their investment, or effort, into increasing their outside option, instead of investing only *in* the employment relation, in order to get a greater share of the surplus. The implications of this possibility for the efficiency in an employment relation are analysed in the following chapter 7.

6.A DERIVATION OF THE OUTCOME OF THE RENEGOTIATION GAME

6.A.1 Complete information

We will solve for the outcome of the renegotiation game in several steps. First, we look at the last subperiod T and calculate the special solution for w_T . Then we turn to an arbitrarily chosen subperiod $t < T$ in order to derive some basic characteristics that help to analyse the game. Afterwards we state and prove some propositions about the game's outcome that will finally provide the general outcome of the game under complete information. If not mentioned the contrary, we will assume that the labour relation is efficient, i.e., that $m - a_F - a_W \geq 0$.

Subperiod T

Assume first that neither agent has ended the game in the subperiods before T . Denote the wage that was valid in subperiod $T - 1$ by w_{T-1} , and the wage that will be valid in subperiod T by w_T . We will solve for the outcome of the renegotiation game in subperiod T by backward induction:

At stage 4 (if reached), the agent selected to make a proposal in this subperiod decides whether to continue the game or to end it. If he decides to end the game, both agents will earn their outside option, $u_T|_{\text{end at st. 4}} = a_W$ and $\pi_T|_{\text{end at st. 4}} = a_F$. If he continues, both parties engage in production and share the marginal product m according to the wage w_T determined at stage 3. In this case, the payoffs are $u_T|_{\text{continue at st. 4}} = w_T$ and $\pi_T|_{\text{continue at st. 4}} = m - w_T$. Hence, if the worker is the selected agent, she will continue if $w_T \in [a_W, m]$, and if the firm is selected, it will continue if $w_T \in [0, m - a_F]$.

At stage 3, the responding agent has to decide between accepting the selected agent's wage proposal, rejecting it in favour of w_{T-1} , or to end the game. First, consider the case where the firm is the responding agent. It has to make a decision between w_{T-1} , the worker's proposal w_T^W and ending the game. Choosing the latter option yields again payoffs of $u_T|_{\text{end at st. 3}} = a_W$ and $\pi_T|_{\text{end at st. 3}} = a_F$. Furthermore, the firm knows that deciding for a wage which is smaller than a_W will make the worker end the game at the following stage. It will not decide for a wage greater than $m - a_F$, because this option cannot yield a higher payoff than ending the game. Hence, if $w_{T-1}, w_T^W \in [a_W, m - a_F]$ the firm will decide for the lower of the two wages, if $w_{T-1} \in [a_W, m - a_F]$ but $w_T^W \notin [a_W, m - a_F]$ it will decide for w_{T-1} and reject the worker's proposal, if $w_T^W \in [a_W, m - a_F]$ but $w_{T-1} \notin [a_W, m - a_F]$, it will decide for w_T^W and accept the worker's proposal, and if $w_{T-1}, w_T^W \notin [a_W, m - a_F]$, the firm will end the game. If the worker is the responding agent, her decision will be the same, except for the case $w_{T-1}, w_T^F \in [a_W, m - a_F]$ where she will decide for the higher of the two wages.

At stage 2 the selected agent has to make a wage proposal. Consider the decision of the worker when being selected. If $w_{T-1} \in [a_W, m - a_F]$, she knows that her proposal will only be accepted by the firm at stage 3 if her proposal w_T^W is less than w_{T-1} . However, the worker has no interest in making such a proposal because it would yield a lower payoff to her than w_{T-1} . So the worker has no interest in making a proposal.¹⁹ If however $w_{T-1} \notin [a_W, m - a_F]$, the firm will only end the game if also $w_T^W \notin [a_W, m - a_F]$. Hence,

¹⁹If the worker is forced to make a proposal by the rules of the game, she will optimally propose some wage greater or equal than w_{T-1} , knowing that the firm will reject it anyway. In order to simplify the notation, assume that the worker will propose $w_T^W = m - a_F$ if $w_{T-1} \notin [a_W, m - a_F]$.

the worker will set $w_T^W = \max[a_W, m - a_F] = m - a_F$, the firm will accept this proposal and trade will occur. In case the firm is selected to make a proposal, the argument is similar. The firm's optimal proposal will be $w_T^F = \min[a_W, m - a_F] = a_W$.

Hence, the expected wage in subperiod T before Nature makes her choice about the proposing agent is given by

$$E_{T-1}(w_T | w_{T-1}) = \begin{cases} w_{T-1} & \text{if } w_{T-1} \in [a_W, m - a_F] \\ \theta a_W + (1 - \theta)(m - a_F) & \text{if } w_{T-1} \notin [a_W, m - a_F] \end{cases}$$

Furthermore, both agents expect trade to occur in T and to earn a payoff at least as high as their outside options.

Some general characteristics

Having solved for the (conditional) outcome of the renegotiation game in subperiod T , we can now turn to some general characteristics of the game. First, we make some useful conventions.

For simplicity assume that both agents have the same time preference reflected by the discount factor $\delta \in]0, 1[$.²⁰ Denote the expected discounted payoff of agent i , $i = F, W$, as subgame-perfect outcomes of the subgame starting in subperiod t , given a wage w that was valid in some subperiod $t' \leq t$ by $V_t^i(w | t') = \sum_{z=t}^T \delta^{z-t} v^i(w_z | w | t')$, where $v^i(w_z | w | t')$ is the payoff agent i gets in subperiod z when w was the valid wage in t' . Let the total surplus generated in the labour relation be s_t in subperiod t and the total discounted surplus be $S_t = \sum_{z=t}^T \delta^{z-t} s_z$, then the maximal achievable surpluses are given by

$$\begin{aligned} s_t^* &= \max\{a_F + a_W, m\} \quad \forall t \leq T, \\ S_t^* &= \sum_{z=t}^T \delta^{z-t} s_z^* \\ &= \max\left\{ \frac{1 - \delta^{T+1-t}}{1 - \delta} (a_F + a_W), \frac{1 - \delta^{T+1-t}}{1 - \delta} m \right\} \quad \forall t \leq T, \end{aligned}$$

with $s_t^* = m$ if the labour relation is efficient, and $s_t^* = a_F + a_W$ if it is inefficient, i.e., if $m - a_F - a_W < 0$. We now turn to some basic characteristics of the payoff functions.

Proposition 1. (Range of the payoff functions)

If the labour relation is efficient, the minimum and the maximum discounted payoffs an agent i , $i = F, W$, can ever expect in a subperiod $t \leq T$ as subgame-perfect outcomes of the subgame starting in t are given by

$$\begin{aligned} V_t^{i, \min} &= \sum_{z=t}^T \delta^{z-t} a_i = \frac{1 - \delta^{T-t+1}}{1 - \delta} a_i, \\ V_t^{i, \max} &= \sum_{z=t}^T \delta^{z-t} (m - a_j) = \frac{1 - \delta^{T-t+1}}{1 - \delta} (m - a_j), \end{aligned}$$

where a_i and a_j are the outside options of the agents.

²⁰Notice that this is a crucial assumption. If the time preferences are not the same, the outcome of the renegotiation game might not be unique.

Proof. Consider an arbitrarily chosen subperiod $t \leq T$. Suppose that agent i is selected to make a wage proposal in this subperiod. At stage 4 (if reached), after agent j has decided whether to accept or reject the proposal (or to end the game) agent i must make a choice between continuing the game or ending it. If he ends the game his (certain) discounted payoff is given by $V_i^i(\text{end})$. If he continues the game with the wage w_t picked by the responding agent j at stage 3, he expects his discounted payoffs to be $V_i^i(w_t|t)$. Suppose that $V_i^i(w_t|t) < V_i^i(\text{end})$. Then agent i will optimally end the game and earn his outside option a_i henceforth. If however $V_i^i(w_t|t) \geq V_i^i(\text{end})$, he will optimally continue the game expecting a discounted payoff at least as high as $V_i^i(\text{end})$. Now $V_i^i(\text{end}) = \sum_{s=t}^T \delta^{s-t} a_i$.

Suppose now that agent i is the responding agent. At stage 3, after agent j has made his proposal w_t^j he must decide between accepting this proposal, rejecting it in favour of w_{t-1} , or ending the game. Denote the corresponding expected discounted payoffs by $V_i^i(w_{t-1}|t)$, $V_i^i(w_t^j|t)$ and $V_i^i(\text{end})$. The agent chooses to end the game if $V_i^i(\text{end}) > \max\{V_i^i(w_{t-1}|t), V_i^i(w_t^j|t)\}$. In this case, he will earn his outside option a_i henceforth. Consequently, $V_i^i(\text{end}) = \sum_{s=t}^T \delta^{s-t} a_i$.

Note that if agent i decides to end the game at stage 3 or 4 and earns his outside option henceforth, also agent j will earn his outside option henceforth. Consequently, each agent i , $i = F, W$, can expect a discounted payoff of at least $V_i^i, \min = \sum_{s=t}^T \delta^{s-t} a_i = \frac{1-\delta^{T-t+1}}{1-\delta} a_i$. \square

The maximum expected discounted payoff of agent i , $i = F, W$, can be calculated as the maximum surplus in case the labour relation is efficient minus the minimum expected discounted surplus of agent j . So $V_i^i, \max = S_t^* - V_j^j, \min = \sum_{s=t}^T \delta^{s-t} m - \sum_{s=t}^T \delta^{s-t} a_j = \frac{1-\delta^{T-t+1}}{1-\delta} (m - a_j)$. Agent i cannot expect a higher payoff, because this would imply that agent j earns less than his outside option. Agent j would end the game in this case and both agents would earn their outside option. If however $V_i^i = V_i^i, \max$, agent j is just indifferent between continuing and ending the game. \square

The analysis of the determination of the valid wage in a subperiod t can be done in a quite simple way without solving the outcome of the renegotiation game in t :

Proposition 2. (*Stability I*)

Suppose that the employment relation is efficient and that w_{t-1} is the valid wage in subperiod $t-1$. Then $\bar{w} \neq w_{t-1}$ can only become the valid wage in subperiod t if it is strictly preferred by both agents.

Proof. Denote the expected present value payoffs from having w_{t-1} as valid wage in t by $V_i^i(w_{t-1}|t)$ and $V_j^j(w_{t-1}|t)$, and the present value payoffs from having \bar{w} as valid wage in t by $V_i^i(\bar{w}|t)$ and $V_j^j(\bar{w}|t)$. Suppose that agent i is selected to make a proposal. Then, at stage 3, agent j must decide between accepting or rejecting i 's proposal (and ending the game). He will prefer accepting a proposal \bar{w} to rejecting it if $V_j^j(\bar{w}|t) > V_j^j(w_{t-1}|t)$. If however $V_j^j(w_{t-1}|t) \geq V_j^j(\bar{w}|t)$, he will prefer to reject the proposal. Consequently, at stage 2, agent i will optimally make a proposal $\bar{w} \neq w_{t-1}$ which satisfies

$$\bar{w} = \arg \max V_i^i(\bar{w}|t) \text{ s.t. } V_i^i(\bar{w}|t) > V_j^j(w_{t-1}|t).$$

If there is such a wage, agent j will accept it. Otherwise, w_{t-1} will be the valid wage in subperiod t , too. \square

Proposition 3. (Stability II)

Suppose that the employment relation is efficient and that w_{t-1} is the valid wage in subperiod $t-1$. Then it will also be the valid wage in subperiod t if

1. $S_t(w_{t-1}|t) = S^*$, and
2. $V_t^i(w_{t-1}|t) \geq V_t^i(a_i)$, $i = F, W$.

Proof. Agent i strictly prefers an alternative wage \bar{w} to w_{t-1} if $V_t^i(\bar{w}|t) > V_t^i(w_{t-1}|t)$, and agent j strictly prefers \bar{w} to w_{t-1} if $V_t^j(\bar{w}|t) > V_t^j(w_{t-1}|t)$. (Note that these conditions imply that $V_t^j(\bar{w}|t) \geq V_t^j(a_j)$ and $V_t^i(\bar{w}|t) \geq V_t^i(a_i)$, because otherwise, one of the agents would prefer to end the game.) Adding up the two inequations yields $V_t^i(\bar{w}|t) + V_t^j(\bar{w}|t) > V_t^i(w_{t-1}|t) + V_t^j(w_{t-1}|t) = S^*$. Hence, both conditions cannot be satisfied simultaneously, and by proposition 2a, w_{t-1} must also be the valid wage in subperiod t . \square

Note that proposition 3 implies that once a wage $w \in [a_W, m - a_F]$ has become the valid wage in a subperiod $t < T$, it will also be the valid wage in all subsequent subperiods until T (or until the agents' outside options change such that one of them becomes binding given the ongoing wage).

Outcome of the renegotiation game

So far we have shown that once a wage $w \in [a_W, m - a_F]$ has become the valid wage in a subperiod $t < T$, it will also be the valid wage in all subsequent periods. This implies that if the agents have bargained a wage $w_0 \in [a_W, m - a_F]$ in period 0, it will remain the valid wage throughout period 1; renegotiations cannot alter the wage level and the contract holds. However, w_0 might be set such that it makes either the worker's outside option bind, $w_0 < a_W$, or the outside option of the firm, $w_0 > m - a_F$. In these cases, the contract has to be renegotiated, because otherwise one of the agents would prefer to end the game.

Proposition 4. (Upward and downward renegotiation)

If the employment relation is efficient, $w_0 \in [0, m]$ and T sufficiently large, the outcome of the renegotiation game is as follows:

1. If $w_0 \in [a_W, m - a_F]$, w_0 will remain the valid wage in all subperiods $t \leq T$, and the payoffs to the agents are given by

$$u_t = w_0 \quad \forall t \leq T, \quad U_1 = \sum_{z=1}^T \delta^{z-1} u_t = \frac{1 - \delta^T}{1 - \delta} w_0,$$

$$\pi_t = m - w_0 \quad \forall t \leq T, \quad \Pi_1 = \sum_{z=1}^T \delta^{z-1} \pi_t = \frac{1 - \delta^T}{1 - \delta} (m - w_0).$$

2. If $w_0 \in [0, a_W[$ (the outside option of the worker is binding), $w = a_W$ will be the valid wage in all subperiods $t \leq T$, and the payoffs to the agents are given by

$$u_t = a_W \quad \forall t \leq T, \quad U_1 = \sum_{z=1}^T \delta^{z-1} u_t = \frac{1 - \delta^T}{1 - \delta} a_W,$$

$$\pi_t = m - a_W \quad \forall t \leq T, \quad \Pi_1 = \sum_{z=1}^T \delta^{z-1} \pi_t = \frac{1 - \delta^T}{1 - \delta} (m - a_W).$$

3. If $w_0 \in]m - a_F, m]$ (the outside option of the firm is binding), $w = m - a_F$ will be the valid wage in all subperiods $t \leq T$, and the payoffs to the agents are given by

$$u_t = m - a_F \quad \forall t \leq T, \quad U_1 = \sum_{z=1}^T \delta^{z-1} u_t = \frac{1 - \delta^T}{1 - \delta} (m - a_F),$$

$$\pi_t = a_F \quad \forall t \leq T, \quad \Pi_1 = \sum_{z=1}^T \delta^{z-1} \pi_t = \frac{1 - \delta^T}{1 - \delta} a_F.$$

Proof. Part (1) of the proposition has already been proven. Note that it is true for all T .

The proof of parts (2) and (3) can be done jointly. Consider subperiod $t < T$ and assume that there was no agreement so far, i.e., w_0 was the valid wage in all subperiods before t . Denote the subgame-perfect wage proposals in t by w_t^F and w_t^W . Note that $w_t^F, w_t^W \in [a_W, m - a_F]$ and that these proposals are acceptable. Since we have complete information, the agents also know what the subgame-perfect proposals in the following subperiod, w_{t+1}^F and w_{t+1}^W , will be. Ignore for the moment the possibility that an agent might end the game if the proposal is rejected. Then, if they reject w_t^F or w_t^W , they will earn $u(w_0) = w_0$ and $\pi(w_0) = m - w_0$ in subperiod t and expect to earn $u(\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W) = \theta w_{t+1}^F + (1 - \theta) w_{t+1}^W$ and $\pi(\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W) = m - \theta w_{t+1}^F - (1 - \theta) w_{t+1}^W$ henceforth. Therefore, the wage proposals have to satisfy

$$\sum_{z=t}^T \delta^{z-t} w_t^F \geq w_0 + \delta \sum_{z=t+1}^T \delta^{z-t-1} (\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W),$$

$$\sum_{z=t}^T \delta^{z-t} (m - w_t^W) \geq (m - w_0) + \delta \sum_{z=t+1}^T \delta^{z-t-1} (m - \theta w_{t+1}^F - (1 - \theta) w_{t+1}^W),$$

or equivalently,

$$w_t^F \geq \frac{1 - \delta}{1 - \delta^{T+1-t}} w_0 + \delta \frac{1 - \delta^{T-t}}{1 - \delta^{T+1-t}} [\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W],$$

$$w_t^W \leq \frac{1 - \delta}{1 - \delta^{T+1-t}} w_0 + \delta \frac{1 - \delta^{T-t}}{1 - \delta^{T+1-t}} [\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W].$$

Furthermore, the proposals have to satisfy $w_t^F \geq a_W$ and $w_t^W \leq m - a_F$, because otherwise an agent might end the game if a proposal is rejected. Note that both agents will optimally propose the same wage level if the right hand of the inequations is within the range $[a_W, m - a_F]$.²¹ So we have

$$\theta w_t^F + (1 - \theta) w_t^W = \frac{1 - \delta}{1 - \delta^{T+1-t}} w_0 + \delta \frac{1 - \delta^{T-t}}{1 - \delta^{T+1-t}} [\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W],$$

$$w_t^F \geq a_W, \quad w_t^W \leq m - a_F.$$

Rearranging we get a first order linear difference equation,

$$\left(1 - \delta^{T+1-t}\right) \left[\theta w_t^F + (1 - \theta) w_t^W\right]$$

$$= \left(1 - \delta\right) w_0 + \delta \left(1 - \delta^{T+1-(t+1)}\right) \left[\theta w_{t+1}^F + (1 - \theta) w_{t+1}^W\right],$$

²¹This means that the agents' optimal proposals are equal to the expected wage proposal in each subperiod (except of subperiod T). Hence, instead of solving for $\theta w_t^F + (1 - \theta) w_t^W$, $t < T$, we could replace this term by either w_t^F or w_t^W .

which has the general solution

$$\left[\theta w_t^F + (1 - \theta) w_t^W \right] = \frac{1 - \delta^{T-t}}{1 - \delta^{T+1-t}} w_0 + \delta^{T-t} \frac{1 - \delta}{1 - \delta^{T+1-t}} \left[\theta w_T^F + (1 - \theta) w_T^W \right].$$

The optimal wage proposals in subperiod T if $w_{T-1} \notin [a_W, m - a_F]$ are given by $w_T^F = a_W$ and $w_T^W = m - a_F$. Hence, the optimal proposals in subperiod 1 are $w_1^F = w_1^W = \Phi_1$, if $\Phi_1 \in [a_W, m - a_F]$, with $\Phi_1 = \frac{1 - \delta^{T-1}}{1 - \delta^T} w_0 + \delta^{T-1} \frac{1 - \delta}{1 - \delta^T} [\theta a_W + (1 - \theta)(m - a_F)]$. Note that $\lim_{T \rightarrow \infty} \Phi_1 = w_0$. This cannot be the valid wage, however. Therefore $w_1^F = w_1^W = a_W$ if the worker's outside option is binding, $w_0 < a_W$, and $w_1^F = w_1^W = m - a_F$ if the firm's outside option is binding, $w_0 > m - a_F$. The agent whose outside option is binding cannot expect a higher payoff than his outside option by either rejecting these proposal if he is the responding agent, or proposing an alternative wage if he is the agent to make a proposal. \square

The intuition behind this result is that before an agreement about a wage within the range $[a_W, m - a_F]$ is reached, the agent whose outside option is not binding, has a higher payoff in subperiods before the agreement is reached than he can expect to have afterwards. On the other hand, the agent whose outside option is binding, has payoffs lower than his outside option in these subperiods. Hence, the latter agent would like to reach an agreement as soon as possible, whereas the former would like to postpone an agreement, would not he have to take into account that the latter agent might end the game. Therefore, both agents come immediately to an agreement with the former agent receiving his outside option.

6.A.2 Asymmetric information

Suppose that the firm has private information about its outside option a_F and the worker's marginal product m , and that the worker is privately informed about her outside option a_W . Furthermore, both agents hold beliefs about their opponent's variables.

In order to illustrate the effects of this assumption on the outcome of the renegotiation game, consider a one-shot version of the renegotiation game²² and assume that the firm is selected to make a proposal.

At stage 4, the firm has to decide whether to end the relationship and earn its outside option a_F in this subperiod, or to continue and receive the marginal product minus the wage selected by the worker at stage 3. In the latter case, the firm earns $m - w$. Obviously, the firm will continue the relation if $w \leq m - a_F$. Suppose that the worker assigns a probability P^F to the case that the firm will continue the relation. It is reasonable to assume that $P^F = 1$ if $w = w^F$, i.e., that the worker believes that the firm is always willing to continue (and let work occur) if its proposal was accepted.

Upon this belief the worker has to choose between accepting or rejecting the firm's proposal w^F , or ending the relation. When accepting the worker knows that she will earn w^F , when ending the relationship she will receive her outside option a_W , and when rejecting the firm's proposal she expects her payoff to be $P^F w_0 + (1 - P^F) a_W$. So the worker will end the relationship only if both wage levels, w_0 and w^F , are less than her outside option. Moreover, she will prefer accepting the proposal to rejecting it if $w^F \geq P^F w_0 + (1 - P^F) a_W$. This, however, means that the firm, when determining its optimal proposal w^F , does not only have to hold a belief over the worker's outside option but also

²²I.e., the game is played only in a single subperiod. Equivalently, consider subperiod T .

over the worker's belief over $m - a_F$, reflected in the probability P^F . It can neglect the latter, if it makes a wage proposal greater than the previously valid wage. In this case, the worker can assume that the firm is willing to continue the relation under both wage levels and will assign a value of one to the probability P^F . For wage proposals less than the previously valid wage, the firm has to take the worker's belief into account, though.

Consequently, two cases have to be distinguished when analysing the firm's behaviour at stage 2. If it is committed to make a proposal not less than the previously valid wage, it sets

$$\begin{aligned} w^F &= \arg \max_{w^F \geq w_0} P^W (m - w^F) + (1 - P^W) a_F \\ &= m - a_F - \frac{P^W}{\frac{dP^W}{dw^F}}, \end{aligned}$$

where P^W is the probability the firm assigns to the case that the worker will continue the relation given this proposal, and $\frac{dP^W}{dw^F} > 0$. Note that at stage 3, rejecting the proposal is always weakly dominated by accepting it.

If there is no restriction on the firm's proposal, the structure of beliefs becomes much more complicated. In order to keep the notation simple, denote the probability that the firm assigns to the case that the worker will accept a proposal given the previously valid wage by Q_A^W , the probability that the worker will reject by Q_R^W , and the probability that the worker will end the relationship by Q_Q^W . Obviously, all these (subjective) probabilities depend on w^F and w . It is reasonable to assume that

$$\frac{\partial Q_A^W}{\partial w^F} > 0, \quad \frac{\partial Q_A^W}{\partial w_0} < 0, \quad \frac{\partial Q_R^W}{\partial w^F} < 0, \quad \frac{\partial Q_R^W}{\partial w_0} > 0, \quad \frac{\partial Q_Q^W}{\partial w^F} < 0, \quad \frac{\partial Q_Q^W}{\partial w_0} < 0.$$

The firm then sets

$$\begin{aligned} w^F &= \arg \max Q_A^W (m - w^F) + Q_R^W x + Q_Q^W a_F \\ &= m - \frac{Q_A^W - \frac{\partial Q_R^W}{\partial w^F} x - \frac{\partial Q_Q^W}{\partial w^F} a_F}{\frac{\partial Q_A^W}{\partial w^F}}, \end{aligned}$$

with $x = w_0$ if the firm continues the relation after the worker has rejected its proposal, and $x = a_F$ if the firm ends the relation. In the latter case, the optimal wage proposal simplifies to

$$w^F = m - a_F - \frac{Q_A^W}{\frac{\partial Q_A^W}{\partial w^F}},$$

which is similar to the equation when the firm was committed to set a wage not less than the previously valid wage.

The optimal wage proposal the worker would make can be calculated similarly.

CHAPTER 7

Wage bargaining under asymmetric information

7.1 INTRODUCTION

This chapter focuses again on the question, whether a wage-setting scheme in which a contract wage is determined that can be altered subsequently can be efficient. As in the previous chapter, two criteria are used for efficiency: first, does the wage-setting scheme avoid inefficient quits and layoffs, and second, does it induce firm and worker to make optimal investments in the employment relation? In answering these questions, we will use a set-up similar to that of the previous chapter 6. However, while that chapter stressed the importance of specific investments, which generate returns that accrue directly to the investor, we will now focus on the type of investment that is generally considered to be typical for employment relations: the worker exerts herself in the job and generates some output, which does not accrue to herself directly, but to her employer, who in turn pays a wage to the worker.

The question addressed in this chapter is therefore, whether there are contracts providing for a wage gap, which induce an efficient investment—or, as we will call it, allocation of effort—in the firm.

The indirect link between the investment (exertion or effort) and its returns (the wage) does not create a problem as long as the employer and the worker can make an enforceable contract about the investment. If, however, the worker can evade control by the employer and the investment decision becomes a hidden action, making a contract that induces a desired investment is a nontrivial task.

This problem has evoked an extensive literature that can be structured along two closely related ideas—efficiency wages, especially the shirking model of Shapiro and Stiglitz (1984), that is used at several places in this thesis,¹ and principal-agency. In the shirking model, a firm is usually perfectly able to control its worker's exertion (her effort) in the job, but because monitoring is costly, it confines itself to taking samples and threatens its workers with sanctions if caught shirking. Since monitoring is incomplete and the hardest possible sanction is usually a layoff, the firm will pay a so called efficiency wage that consists of an optimal mark-up on the worker's outside option to induce the worker to provide the desired effort in order to keep her job. The efficiency wage is announced *before* the worker decides about her effort.

¹See chapters 4 and 5.

Principal-agency is a broader approach that applies to a wide range of relations between two parties called principal and agent. The principal wants the agent to undertake a certain costly action, but he cannot monitor the agent. Instead, he receives information that is only indirectly related to the agent's action.² The problem of the principal (in our setting the firm) is to find a transfer (a wage) to the agent (the worker) to induce the desired action. Because the firm cannot simply tie the wage to the worker's unobservable input, the optimal transfer is usually related to the received information. The wage is therefore determined *after* the worker has decided about her action. Consequently, the contract between firm and worker must consist of a wage-setting scheme describing how the wage will be set *ex post*, after the firm has received the information. In the simplest setting, the optimal wage-setting scheme will reward favourable information and punish unfavourable information, in order to induce the worker to take actions that increase the probability of favourable information. In both efficiency-wage and principal-agent models, the wage, or wage-setting scheme, is used to reveal the worker's investment, thus resolving the problem of asymmetric information, and in both kinds of models this is costly to the firm. It faces a trade-off that usually results in inefficient allocations.

This chapter adopts the principal-agency approach. It assesses the performance of different wage-setting schemes in a highly stylized, yet reasonably realistic relation between a firm and a single worker, based on a model of Aghion, Compte, and Jehiel (1998). The worker's investment decision consists of the allocation of one unit of effort over two uses: in the firm, in order to increase the total surplus to be generated in the employment relation, or outside the firm, in order to increase her outside option. The firm cannot observe the effort allocation directly, but it can observe the generated surplus (the worker's marginal product), and it receives a signal about the worker's outside option. Based on the signal, the firm can assess the worker's real outside option only if it has a correct estimate of the worker's effort allocation.

As far as only the allocation of effort in the firm is concerned, this setting is a standard textbook principal-agent problem. The firm maximizes its profits subject to an incentive-compatibility constraint, which results from the worker's utility-maximization problem and describes how much effort the worker will allocate in the firm given a certain remuneration scheme. Usually, the incentive-compatibility constraint equalizes the marginal cost of effort allocated in the firm with its marginal return. As the returns accrue directly to the firm, the firm must promise to give the worker a certain share in the returns generated by the effort allocated in the firm, i. e., in her marginal product.³ The resulting effort allocation will typically be inefficient from a social point of view, as the firm will not transfer the whole surplus

²An alternative definition is given by Fudenberg and Tirole (1991, p. 243): "... [T]here is a 'principal' who would like to condition her actions on some information that is privately known by the other players, called 'agents.'" For a comprehensive overview of principal-agent models see also Hart and Holmström (1987).

³An even better solution for the firm would be to punish the worker for generating a low surplus. However, this is usually prevented *ex ante* by the participation constraint, which will be discussed in the next paragraph, and *ex post*, because either the worker can evade the punishment, or because it is not feasible or legal.

to the worker.

This problem could in principle be resolved by letting the worker enjoy the whole returns of effort allocated in the firm, minus a predetermined lump sum for the firm.⁴ Yet, the firm has to take another restriction into account, namely that the worker participates in the contract. If the firm fails to pay a wage that is not at least equal to the worker's outside option, the worker will quit. As long as the firm can observe the outside option, it can tailor the mentioned lump sum to avoid a quit. However, two features of our model complicate the situation: the outside option is private knowledge of the worker, and the worker can influence her outside option by allocating effort outside the firm.

In this chapter, we will not stress the information problem, but focus on the effect of the worker's ability to influence her outside option. If the worker can choose between allocating effort in the firm or outside, the following problem arises: one cannot simply tie the wage to the worker's marginal product only, in order to make the worker allocate all her effort in the firm and maximize the total surplus of the employment relation. Instead, the contract between firm and worker must also reward the worker's investment in the outside option, because otherwise the contract might jeopardize the employment relation, if the wage makes the outside option bind *ex post*.

The problem of finding an "optimal" contract that on the one hand provides the appropriate incentives for the worker to allocate all her effort in the firm, while on the other hand it respects the *ex post* participation constraint⁵, is virtually unsolvable in general terms. The worker is—using the terms of principal-agent theory—described by a two-dimensional type space, her marginal product, and her outside option, which makes it hard to derive meaningful results.⁶ Moreover, the incentive compatibility and participation constraints are nonconvex, and, as Rasmusen (1994, p. 176) points out, they will therefore "... rule out a very complicated set of possible wage functions."

In view of these difficulties, the strategy chosen in this chapter is based on a simplified version of Grossman and Hart's (1983) three-step procedure.⁷ Rather than directly attacking the optimization problem of the firm or a benevolent dictator, we investigate in the incentives different wage-setting schemes provide to both em-

⁴For example, the contract might provide for a piece rate reflecting the market value of the worker's output, but in order to earn that rate, the worker has to deliver a certain amount of output for free.

⁵Strictly speaking, the term participation constraint means that the worker has to agree to the contract *ex ante*, i. e., before she decides about her effort allocation, while we are concerned with an *ex post* constraint that avoids a quit of the worker after the consequences of her allocation have become visible.

⁶See for example Laffont, Maskin, and Rochet (1987), McAfee and McMillan (1988), and Rochet (1985).

⁷The term three-step procedure was not invented by Grossman and Hart themselves, but it was coined later on by Fudenberg and Tirole (1990). Related to the problem at hand, the three steps would be, (1) to find for each possible effort level the set of wage-setting schemes that induce this effort level, (2) to find the wage-setting scheme that induces this effort at lowest cost, and (3) to find the most efficient combination of effort and cost-minimizing wage-setting scheme.

ployees and employer. For each of these schemes we identify the effects of different parameters, concerning the distribution of bargaining power, the size of minimum wages, and the information structure, on the effort allocation, as well as on the payoffs of worker and firm. At this point, the goal is not only to find constellations of parameters that lead to an efficient effort allocation, but also to find out, whether the preferences of worker and firm concerning these parameters coincide. The underlying idea is, that an optimal wage-setting scheme would both, provide the worker with the incentive to allocate all her effort in the firm, and be preferred by firm and worker to other wage-setting schemes.

This chapter is organized as follows. Section 7.2 discusses the general structure of the employment relationship between worker and firm, and it explains the assumptions about the information structure of the model. The sections 7.3, 7.4, and 7.5 then focus on the characteristics of different wage-setting schemes. In the first part of section 7.3, the effective wage is solely determined by (Nash) bargaining between worker and firm. From this simple wage-setting scheme, some general conclusions about the behaviour of the two parties are derived. The second part of this section extends the one-tier to a two-tier wage-formation process by adding an exogenous minimum wage that can make the bargaining wage bind. Two-tier wage-formation processes are also discussed in sections 7.4 and 7.5. Section 7.4 analyses the performance of a wage-setting scheme similar to the one of the previous chapter 6, in which an exogenously given contract wage is only adjusted if it makes the outside option of either the firm or the worker bind. Section 7.5 discusses a wage-setting scheme, in which an exogenously given contract wage serves both as a minimum wage, and as a co-determinant of the threat points, when worker and firm negotiate over the distribution of the joint payoff. Finally, section 7.6 compares the outcomes of the different wage-setting schemes and discusses the implications for the occurrence of wage gaps.

7.2 ASSUMPTIONS

The time structure of the model is very similar to the model of Aghion, Compte, and Jehiel (1998). The model deals with the relation between an individual worker and her employer over two time periods, which are depicted in figure 7.1. In the first period, both parties meet and agree to start an employment relationship. No work occurs in this period. Instead it is used for decisions around the effort allocation of the worker.

First, firm and worker agree on a wage-setting scheme w . This can be a fixed wage rate, or a prescription how the wage is to be determined in the second period, after the worker's marginal product m and a verifiable signal s about her outside option have realized.

The worker has one unit of effort which she can allocate for two uses: in the firm, in order to achieve a greater marginal product, and outside the firm, in order to improve her outside option. We assume that firm and worker cannot include the effort allocation in their contract, because it is private knowledge of the worker.

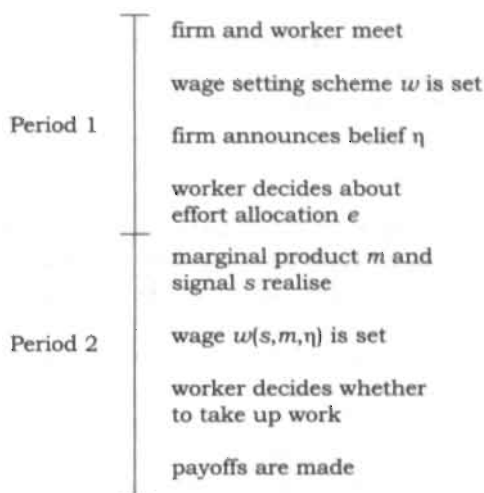


Figure 7.1: The time structure of the model

However, after the wage-setting scheme is set, the firm announces its belief η about how much effort the worker is going to allocate in the firm. This belief is assumed to be static in the sense that the firm will not update it after it has observed the signal. Knowing the wage-setting scheme and the firm's belief, the worker decides about her effort allocation, i. e., how much effort she allocates in the firm (e), and how much she allocates outside ($1 - e$). Notice that the assumption that e is a continuous variable is one of the important differences to the assumptions of Aghion, Compte, and Jehiel. As we will see, this difference can have severe implications for the outcome of the model.

The effort allocation influences the distribution of the signal s and the marginal product m of the worker. They realize at the beginning of the second period and are observed by both parties. We assume that in period 1 the functional forms of the probability-density functions of s and m are known by both the worker and the firm. However, only the worker knows her effort allocation, while the firm has an incorrect perception of the distributions, unless it holds a correct belief about the effort allocation.

The probability-density function and the distribution function of the marginal product are denoted by $g(m|e)$ and $G(m|e)$, with $g(m|e) \geq 0$ for $m \in [m_0, m_1]$ and all e , and $g(m|e) = 0$ for $m \notin [m_0, m_1]$ and all e . The more effort the worker has allocated in the firm, the greater is the probability that the realization of the marginal product exceeds a certain value $m^* \in [m_0, m_1]$. However, the more effort the worker already has allocated in the firm, the less is the marginal effect on this

probability:

$$\begin{aligned} \frac{\partial \Pr(m > m^*)}{\partial e} &= -\frac{\partial G(m^*|e)}{\partial e} \geq 0, & \frac{\partial^2 \Pr(m > m^*)}{\partial e^2} &< 0, \\ \lim_{e \rightarrow 1} \frac{\partial \Pr(m > m^*)}{\partial e} &= 0 \end{aligned} \quad (7.1)$$

A similar assumption is made for the signal s . It is some verifiable evidence about the worker's outside option, for example a concrete job offer or an advertisement for a job for which the worker qualifies. If the worker is allocating effort outside the firm, she is acquiring skills that can be useful in alternative jobs, and she is actively searching for job offers. These activities will increase her value for other employers, and they will produce evidence about her value outside the firm, e.g. in the form of letters of potential employers, or advertisements for jobs for which the worker is suited. Consequently, a lower e (more effort allocated outside the firm) increases the chance of a high signal. However, we assume that the more the worker allocates outside the firm, the less is its marginal effect. We denote the probability-density function and the distribution function by $h(s|e)$ and $H(s|e)$, with $h(s|e) \geq 0$ for $s \in [s_0, s_1]$ and all e , and $h(s|e) = 0$ for $s \notin [s_0, s_1]$ and all e . Let $s^* \in [s_0, s_1]$. Then,

$$\begin{aligned} \frac{\partial \Pr(s > s^*)}{\partial e} &= -\frac{\partial H(s^*|e)}{\partial e} \leq 0, & \frac{\partial^2 \Pr(s > s^*)}{\partial e^2} &< 0 \\ \lim_{e \rightarrow 0} \frac{\partial \Pr(s > s^*)}{\partial e} &= 0. \end{aligned} \quad (7.2)$$

Although the signal is verifiable evidence about the worker's outside option, it is typically not equal to the outside option. If, for instance, the worker presents a job offer as evidence, the relation between signal and outside option can be distorted in two ways: first, the job offer will not necessarily translate into a job, and second, an alternative employer might promise to pay wage supplements which are not contractible, and which the worker can therefore not present as evidence. Either way, the outside option will be positively related to the amount of effort the worker allocates outside the firm, because the worker then acquires more skills useful in other jobs. We assume that there is a positive relation between the signal and the outside option, which is the stronger, the more effort the worker has allocated outside the firm:

$$\frac{\partial a(s, e)}{\partial s} \geq 0, \quad \frac{\partial^2 a(s, e)}{\partial s^2} < 0 \quad \forall s, \quad \lim_{s \rightarrow s_1} \frac{\partial a(s, e)}{\partial s} = 0 \quad (7.3)$$

$$\frac{\partial^2 a(s, e)}{\partial s \partial e} \leq 0, \quad \frac{\partial^3 a(s, e)}{\partial s \partial e^2} < 0 \quad \forall s, e, \quad \lim_{e \rightarrow 0} \frac{\partial^2 a(s, e)}{\partial s \partial e} = 0 \quad (7.4)$$

Moreover, we assume that the worker's outside option increases in the amount of effort allocated outside the firm, but that the effect becomes smaller the more effort the worker has already outside the firm,

$$\frac{\partial a(s, e)}{\partial e} \leq 0, \quad \frac{\partial^2 a(s, e)}{\partial e^2} < 0, \quad \lim_{e \rightarrow 0} \frac{\partial a(s, e)}{\partial e} = 0, \quad \forall s. \quad (7.5)$$

We assume that both the firm and the worker know the functional form of the worker's outside option $a(s, e)$. Consequently, if the firm has a correct perception of the worker's effort allocation, $\eta = e$, it knows the distributions of the marginal product m and the signal s before they realize, and it knows the value of the worker's outside option after the signal has realized.

Recall that the effort allocation is private knowledge of the worker. We assume that the firm cannot infer the correct value of e from the realization of the signal or the marginal product. However, the firm knows the worker's objective, the wage-setting scheme, and the functional forms of the density functions and the worker's outside option. Therefore, it will typically be able to solve for the worker's optimal effort allocation.

A last set of assumptions concerns the firm's outside option a_F and the payoffs generated in the employment relation. Because we want to focus especially on the effects of the worker's outside option, we assume that the firm's outside option is a constant, that is known by both firm and worker already at the beginning of period 1. Moreover, same as in the previous chapter, we assume that the employment relation is efficient in the sense that a greater joint payoff is generated in the relation than outside, i. e., the worker's marginal product always exceeds the sum of the outside option of the firm (a_F) and the worker ($a(s, e)$):

$$\begin{aligned} m &\geq a(s, e) + a_F \quad \forall s, m, e \\ \Rightarrow m_0 &\geq a(s_1, 0) + a_F \end{aligned} \quad (7.6)$$

This assumption ensures that there is always a wage rate that makes both the firm and the worker prefer to maintain the employment relation to choosing their outside option. Hence, any quit or layoff is inefficient. The only possible reason for a separation of firm and worker is asymmetric information.

7.3 NASH BARGAINING

7.3.1 *The effort allocation of the worker*

The first wage-setting scheme that is analysed in this chapter is bargaining about the effective wage w_N , after the marginal product and the signal have realized. Same as Aghion, Compte, and Jehiel (1998), we will examine a slightly adapted Nash-bargaining solution that accounts for asymmetric information between the bargaining parties. Because the bargaining outcome w_N , if accepted by the worker, becomes the effective wage, the wage-setting scheme of this section does not make a provision for a wage gap. Nevertheless, it is useful to analyse this rather simple wage-setting scheme, for it provides some general insights in the behaviour of firm and worker under asymmetric information. This will prove to be useful when analysing the other wage-setting schemes. Another reason for dealing with Nash bargaining in detail is that we will augment the model with a minimum wage in section 7.3.3. The difference between minimum wage and bargained wage can then be interpreted as a wage gap.

In the course of bargaining, the firm will attempt to maximize its profits $\pi = m - w_N$, while the worker will try to maximize her wage w_N . Doing so, both parties make wage proposals such that the opponent's payoff be as close as possible to his outside option. In a noncooperative bargaining game à la Rubinstein (1982) the firm would propose a wage that is equal to the worker's outside option plus an amount that makes the worker just indifferent between accepting and rejecting the offer, while the worker would follow a similar strategy proposing a wage that is based on the firm's outside option.

However, because the firm does not know the worker's effort allocation, it can only base its proposals on its perception $a(s, \eta)$ rather than on the true outside option $a(s, e)$. Consequently, the bargaining outcome is described by

$$\begin{aligned} w_N &= \arg \max (m - w - a_F)^\theta (w - a(s, \eta))^{1-\theta} \\ &= \theta a(s, \eta) + (1 - \theta)(m - a_F), \end{aligned} \quad (7.7)$$

where $\theta \in [0, 1]$ is the bargaining power of the firm, and $1 - \theta$ the bargaining power of the worker. The bargained wage is a weighted average of the worker's outside option as perceived by the firm, and the marginal product minus the firm's outside option, with the weights being the bargaining power of the firm and the worker.⁸

At the moment the worker has to decide about her effort allocation in period 1, she knows the firm's belief, the distribution of bargaining power, and the distributions of the signal and the marginal product contingent on her decision. For the moment, we neglect the possibility that the worker wants to quit if the bargained wage turns out to be lower than her outside option. Then, the worker will choose to allocate that amount of effort in the firm that maximizes the expected wage

$$E[w_N|e] = \theta E[a(s, \eta)|e] + (1 - \theta) E[m - a_F|e].$$

The first-order condition,⁹

$$\theta \frac{\partial E[a(s, \eta)|e]}{\partial e} + (1 - \theta) \frac{\partial E[m - a_F|e]}{\partial e} = 0, \quad (7.8)$$

⁸The argument about the bargaining game follows Aghion, Compte, and Jehiel (1998). It should be noticed, however, that it is not entirely correct: if the firm overstates the amount of effort allocated in the firm, $\eta > e$, its estimate of the worker's outside option will be too low. This can—but need not—lead to a breakdown of the negotiations, namely if the firm fails to make acceptable wage proposals and if it has too much bargaining power. In order to avoid an unnecessary complication of the analysis, we neglect this case and assume instead that the wage w_N is the bargaining result, which the worker might reject by leaving the firm. In this subsection, we will moreover assume, that the worker is locked in the employment relation and has to accept w_N . Later on, this assumption will be relaxed.

⁹The second-order condition,

$$\theta \frac{\partial^2 E[a(s, \eta)|e]}{\partial e^2} + (1 - \theta) \frac{\partial^2 E[m - a_F|e]}{\partial e^2} < 0,$$

is strictly negative by equations (7.2), (7.3), (7.4), (7.5) and (7.1) in section 7.2. Hence, the first order condition (7.8) describes a maximum.

weighs the marginal effect of effort allocated in the firm on the expected perceived outside option and on the expected marginal product by the relative bargaining power of the firm and the worker. The greater the bargaining power of the firm is, the more important the firm's perception of the worker's outside option becomes for the determination of the bargained wage, and the more effort the worker will allocate outside the firm in order to produce better evidence about her outside option.¹⁰ At the same time, it becomes less important for the worker to achieve a greater marginal product. In the opposite case, the marginal product becomes more and more important, and the worker will allocate more effort in the firm.

By differentiating equation (7.8) with respect to the firm's bargaining power,

$$\frac{de}{d\theta} = \frac{\frac{\partial E[m-a_F|e]}{\partial e} - \frac{\partial E[a(s,\eta)|e]}{\partial e}}{\theta \frac{\partial^2 E[a(s,\eta)|e]}{\partial e^2} + (1-\theta) \frac{\partial^2 E[m-a_F|e]}{\partial e^2}} < 0, \quad (7.9)$$

we see that the wage maximizing amount of effort allocated in the firm is indeed a decreasing function of the firm's bargaining power.¹¹ The effect on the expected bargained wage is also negative, as the worker can only partially compensate for its lower share in the marginal product by allocating more effort to her outside option:

$$\frac{dE[w_N|e]}{d\theta} = \frac{\partial E[w_N|e]}{\partial \theta} + \frac{\partial E[w_N|e]}{\partial e} \frac{de}{d\theta} = \frac{\partial E[w_N|e]}{\partial \theta} < 0$$

The effect on the expected profits of the firm is ambiguous. On the one hand, the firm gets a greater share in the marginal product, because the wage decreases. On the other hand, the marginal product decreases because the worker allocates more effort outside the firm:

$$\frac{dE[\pi_N|e]}{d\theta} = \frac{dE[m|e]}{d\theta} - \frac{dE[w_N|e]}{d\theta} = \frac{\partial E[m|e]}{\partial e} \frac{de}{d\theta} - \frac{\partial E[w_N|e]}{\partial \theta} \geq 0$$

However, as the effect via the effort allocation is only indirect, it is likely that the positive effect of an increased share in m is prevailing. In the opposite case, the firm could increase its expected profits by keeping a lower profile in the negotiations with the worker. In any case, a higher bargaining power of the firm comes at the cost of a lower efficiency of the employment relation, as it diverts effort from the relation and leads to a lower expected marginal product.

Another parameter that influences the worker's effort allocation is the firm's belief η about the amount of effort e the worker allocates in the firm. The total differential of the first-order condition (7.8) with respect to the firm's belief is

$$\frac{de}{d\eta} = - \frac{\theta \frac{\partial^2 E[a(s,\eta)|e]}{\partial \eta \partial e}}{\theta \frac{\partial^2 E[a(s,\eta)|e]}{\partial e^2} + (1-\theta) \frac{\partial^2 E[m-a_F|e]}{\partial e^2}} > 0 \quad \text{for } e < 1. \quad (7.10)$$

¹⁰Recall that by allocating more effort to her outside option, the worker acquires more skills that are valuable outside the firm, and she increases her search activities on the labour market.

¹¹See equations (7.25), (7.26), (7.30) and (7.31) in the appendix for the signs of the derivatives in equation (7.9).

The positive relation between the belief η and the real effort allocation e is due to the effect of the belief on the wage proposals the firm makes when negotiating with the worker. After the firm has learned the signal s at the beginning of period 2, it forms its perception of the worker's outside option, $a(s, \eta)$. The greater its belief η is, the lower will be its estimate of the relation between a and s , and the lower will be its wage proposals, which are based on this estimate. Therefore, it pays less for the worker to allocate effort outside the firm, i.e., $1 - e$ is smaller. See also figure 7.2.

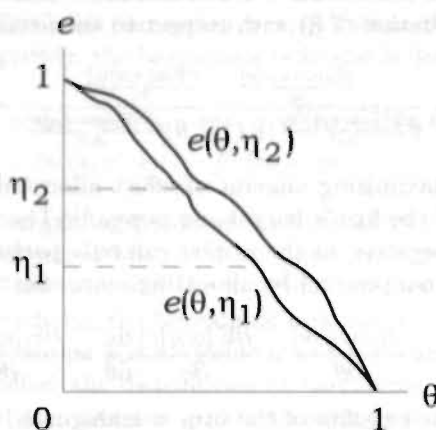


Figure 7.2: The worker's effort allocation as a function of the firm's bargaining power, at two different beliefs $\eta_2 > \eta_1$.

The effect of a greater belief on the expected profits of the firm is unambiguously positive:

$$\frac{dE[\pi_N|e]}{d\eta} = \frac{dE[m|e]}{d\eta} - \frac{dE[w_N|e]}{d\eta} = \frac{\partial E[m|e]}{\partial e} \frac{de}{d\eta} - \frac{\partial E[a(s, \eta)|e]}{\partial \eta} > 0$$

7.3.2 The belief of the firm

The belief η is peculiar because, on the one hand, a rational firm should form its belief on the basis of the behaviour of the worker (i.e., on the basis of the first order condition (7.8)), while on the other hand the behaviour of the worker is just influenced by the belief the firm announces. There are two possible consequences of this characteristic. First, the firm could announce its true belief, taking the worker's behaviour into account. This would result in fixed points $\eta = e(\theta, \eta)$. According to Brouwer's fixed point theorem at least one such fixed point for each distribution of bargaining power will exist. Unfortunately, it is impossible to determine whether a fixed point is unique or whether there is a multiplicity of points, without specifying

the functions.¹² In case there are multiple fixed points, the firm could in principle choose. As its expected profits increase in the amount of effort allocated in the firm, the firm would then optimally decide for the greatest fixed point.¹³

A second possibility is that the firm does not care about fixed points but uses its belief as a strategic variable. In this case, the firm would not announce its true belief but a greater value in order to achieve a greater expected profit. This possibility is problematic, because the worker can see through the manipulation, which would undermine the firm's bargaining position. During the wage negotiations the firm could not credibly assert that the worker's outside option is lower than it actually is. The consequence is that either the firm has to take back its claim, or that the negotiations break down. Yet, the firm will avoid a break down of the negotiations, as this would leave the firm with only its outside options. Hence, pretending that the worker's outside option is lower against better knowledge of the worker is not an optimal strategy. Moreover, even if the firm succeeds in reaching a bargained wage based on its false belief, it risks that the worker quits if her real outside option turns out to be greater than the wage.¹⁴ This possibility was up to this point ruled out by assumption.

We will now relax the assumption that the worker is locked in the employment relation, in order to provide another argument why the firm cannot deliberately announce an exaggerated belief. We will show that the fixed points $e(\theta, \eta)$ are not affected when the worker is allowed to quit, while it can be costly for the firm if it announces a belief that is too high against its better judgement.

If the worker is not locked in the employment relation she will leave the firm whenever at the beginning of period 2 the bargained wage turns out to be less than her outside option,

$$w_N = \theta a(s, \eta) + (1 - \theta)(m - a_F) < a(s, e). \quad (7.11)$$

Because the wage is never less than the outside option as perceived by the firm, a situation where the worker chooses her outside option can only occur if the belief is greater than the true effort allocation, $\eta > e$, and if the firm's bargaining power θ is high enough to make the wage sufficiently close to the perceived outside option.

In order to identify the combinations of e and θ at which the outside option can be binding *ex ante*, notice that the wage is increasing in the worker's marginal product m , and that both the wage and the outside option are increasing in the signal s . On the other hand, the outside option reacts more strongly to an increase in the signal if the firm overestimates the amount of effort allocated in the firm.

¹²The answer will depend on the size of $\frac{de}{d\eta}$ (see equation (7.10)). If it is either less or greater than one for all η , the fixed point is unique. Otherwise, multiple fixed points can exist.

¹³In the model of Aghion, Compte, and Jehiel (1998) there are indeed two fixed points (called equilibria) for a certain range of bargaining power θ . However, this result is due to their assumption that effort is a discrete variable, $e \in \{e^{\text{out}}, e^{\text{in}}\}$. Their dual fixed points capture the different—possibly unique—fixed points that occur if effort is a continuous variable.

¹⁴The firm might succeed if it is negotiating a wage not with only one worker, but with a group of several heterogeneous workers. However, in this case the firm's belief will most likely be wrong for every individual worker, anyway.

Hence, we can define two critical values e_N^* and θ_N^* such that

$$\theta_N^* a(s_1, \eta) + (1 - \theta_N^*) (m_0 - a_F) = a(s_1, e_N^*). \quad (7.12)$$

For values $\theta < \theta_N^*$ of the firm's bargaining power, and values $e > e_N^*$ of the worker's effort the outside option cannot be binding because there are no signals $s \in [s_0, s_1]$ and marginal products $m \in [m_0, m_1]$ that satisfy inequation (7.11). Notice that the greater the critical bargaining power is, the more effort the worker must allocate in the firm in order to avoid that her outside option can become binding: the first- and second-order total differentials of equation (7.12) are

$$\begin{aligned} \frac{de_N^*}{d\theta_N^*} &= -\frac{m_0 - a_F - a(s_1, \eta)}{\left. \frac{\partial a(s_1, e)}{\partial e} \right|_{e_N^*}} > 0, \\ \frac{d^2 e_N^*}{d\theta_N^{*2}} &= -\frac{\left. \frac{\partial^2 a(s_1, e)}{\partial e^2} \right|_{e_N^*}}{\left(\left. \frac{\partial a(s_1, e)}{\partial e} \right|_{e_N^*} \right)^3} (m_0 - a_F - a(s_1, \eta))^2 < 0. \end{aligned}$$

We can now consider the allocation decision of the worker. At the moment the worker has to decide about her effort allocation in period 1, she now expects her payoff in period 2 to be

$$E[u_N|e] = E[w_N|e] + A_N(e),$$

where A_N is the non-negative additional expected payoff due to the opportunity to leave the firm. In the appendix on p. 186 ff. its characteristics are analysed in detail. Because the gap between the firm's belief and the true effort allocation narrows when the worker allocates more effort $e < \eta$ in the firm, the additional expected payoff is decreasing in e . Therefore, for combinations $(\theta, e) \in ((\theta_N^*, 1], [0, e_N^*))$, where the outside option can turn out to be binding, the derivative of expected utility,

$$\frac{\partial E[u_N|e]}{\partial e} = \frac{\partial E[w_N|e]}{\partial e} + \frac{\partial A_N(e)}{\partial e} < \frac{\partial E[w_N|e]}{\partial e}, \quad \text{for } \theta > \theta_N^*, e < e_N^*,$$

is smaller than the derivative of the expected wage, and the worker will choose to allocate less effort in the firm than in the scenario where the worker is locked in the employment relation. However, because her outside option cannot become binding if the worker chooses the same effort allocation the firm has predicted, the fixed point(s) $e = \eta$ are not affected. See figure 7.3 for a comparison of the two scenarios.

In contrast to the scenario of section 7.3.1, the reaction of the worker now makes it costly for the firm to—at least massively—overstate its belief. If the firm exaggerates, the worker will not only choose to allocate less effort in the firm, leading to reduced expected profits, but she might ultimately decide to quit, leaving the firm with an *ex post* profit just equal to its outside option.

Due to the inability of the firm to succeed in bargaining with a false exaggerated belief, as well as due the potential cost involved if the firm succeeds with a false

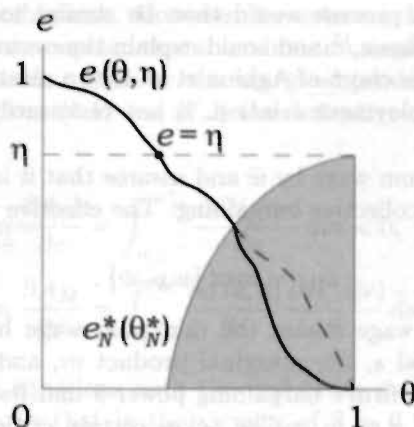


Figure 7.3: The worker's effort allocation with the opportunity to quit. The shaded area indicates combinations of e and θ where the worker's outside option can become binding. The probability that the outside option becomes binding is equal to zero on the curve $e_N^*(\theta_N^*)$ and equal to one on the lines $e = 0$ and $\theta = 1$ within the area. The dashed line is added for comparison with the scenario where the worker cannot quit.

belief, the firm is highly restricted in using its belief as a strategic variable. For the performance of the employment relation this is good as well as bad: on the one hand, the firm is prevented from jeopardizing the employment relationship by massively exaggerating the belief. On the other hand, a mild exaggeration which does not endanger the relationship is also prevented, although it would induce the worker to allocate more effort in the firm, and lead to a higher joint payoff.

The main reason for the failure of the Nash-bargaining scenario to produce the efficient effort allocation $e = 1$ is that the wage depends on both the worker's performance in the firm and on her outside option. The more important the outside option becomes for the determination of the worker's payoff, the less effort she allocates in the firm. Unless the firm sacrifices its share in the joint payoff, Nash bargaining leads to inefficient outcomes.

7.3.3 Nash bargaining with a minimum wage

In order to establish an efficient fixed point in the Nash-bargaining scenario Aghion, Compte, and Jehiel (1998) propose the introduction of a minimum wage that replaces the bargained wage if the bargained wage is less than the minimum wage. They argue that such a wage-setting scheme with an appropriate minimum wage can (re)establish an efficient fixed point $e = \eta$ if the Nash-bargaining wage fails to. According to Aghion et al., the minimum wage could be enforced by a union. The

resulting wage-formation process would then be similar to the two-tier processes discussed earlier in this thesis,¹⁵ and could explain the occurrence of wage gaps. As we will show, however, the claim of Aghion et al. that a minimum wage can enhance the efficiency of the employment relation, is not necessarily true when effort is a continuous variable.¹⁶

We denote the minimum wage by \bar{w} and assume that it is exogenously given, for example as the result of collective bargaining. The effective wage w_M in period 2 is now equal to

$$w_M = \max \{w_N, \bar{w}\}.$$

Whether the bargained wage makes the minimum wage binding depends on the constellation of the signal s , the marginal product m , and the parameters θ and η . At a given level of the firm's bargaining power θ and its belief η , the minimum wage becomes binding if s or m realize below certain critical values $s_M^* \in (s_0, s_1]$ and $m_M^* \in (m_0, m_1]$. The critical signal just makes the bargained wage equal to the minimum wage at given m , θ , and η ,

$$\theta a(s_M^*, \eta) + (1 - \theta)(m - a_F) = \bar{w}, \quad (7.13)$$

while at the critical marginal product the same relation holds at s_1 ,

$$\theta a(s_0, \eta) + (1 - \theta)(m_M^* - a_F) = \bar{w}.$$

In period 1, the worker expects her payoff in period 2 to be equal to the expected bargained wage plus a term A_M reflecting the situations where the minimum wage becomes binding,

$$E[w_M|e] = E[w_N|e] + A_M.$$

Notice that the expected payoff does not depend on the real outside option of the worker. This is justified because the arguments developed in the previous section about the firm's belief still apply: the firm announces a belief that leads to a fixed point $e = \eta$, and the worker's outside option can never become binding.

The expected additional payoff is related to the expected difference between minimum and bargained wage at signals and marginal products that make the minimum wage bind:

$$A_M = \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} [\bar{w} - w_N] h(s|e) g(m|e) ds de \quad (7.14)$$

Its derivative with respect to the effort allocation is calculated in the appendix on page 188, equation (7.40). It is equal to

$$\frac{\partial A_M}{\partial e} = (1 - \theta) \int_{m_0}^{m_M^*} \left(\frac{H(s_M^*|e)}{\partial e} G(m|e) + H(s_M^*|e) \frac{\partial G(m|e)}{\partial e} \right) dm.$$

¹⁵ See the models in the chapters 3.5, 4, and 5.

¹⁶ Recall that Aghion et al. assume effort to be a binary variable.

Its sign is in general ambiguous. However, for extreme distributions of bargaining power—i. e., under the assumption that the minimum wage can be binding for all θ —either the worker's outside option or the marginal product vanish from the wage equation, and by equations (7.1) and (7.2) the derivative of the expected additional payoff becomes

$$\lim_{\theta \rightarrow 0} \frac{\partial A_M}{\partial e} = \int_{m_0}^{m_M} \frac{\partial G(m|e)}{\partial e} dm < 0, \quad (7.15)$$

$$\lim_{\theta \rightarrow 1} \frac{\partial A_M}{\partial e} = \int_{m_0}^{s_M} \frac{\partial a(s, \eta)}{\partial s} \frac{\partial H(s|e)}{\partial e} ds > 0. \quad (7.16)$$

A comparison of these equations with the first-order condition (7.8) in the Nash-bargaining scenario shows that in the vicinity of $\theta = 0$ (full bargaining power of the worker), the amount of effort allocated in the firm tends to be lower due to the minimum wage, while in the vicinity of $\theta = 1$ (full bargaining of the firm), the amount of effort allocated in the firm just tends to be lower than without a minimum wage. Although the optimal effort allocations at $\theta = 0$ and $\theta = 1$ cannot be altered, the equations (7.15) and (7.16) indicate that for $\theta \in (0, 1)$ the minimum wage induces the worker to choose less extreme effort allocations than in the Nash-bargaining scenario. If the derivative of the expected additional payoff is continuous, the optimal effort allocation will behave as depicted in figure 7.4.

The left part of figure 7.4 depicts the effort allocation when the minimum wage is sufficiently high, such that it can be binding for all distributions of bargaining power. If the firm's bargaining power is low, the first-order derivative of the expected wage is smaller than in the Nash-bargaining scenario, which means that the worker will choose to allocate less effort in the firm, depicted by the switch from point A to A' . In the opposite case, with the firm's bargaining power being high, the worker decides to allocate more effort in the firm, resulting in a switch from point B to B' . The right part of the figure depicts two possible effort allocations in case the minimum wage cannot be binding for all distributions of bargaining power. The worker will still allocate more effort in the firm if the firm's bargaining power is high, but for lower θ , the effect of the minimum wage on the effort allocation is ambiguous. It cannot be solved for without specifying the functions.

The described effect of the minimum wage qualifies the claim of Aghion et al., that a minimum wage can increase the amount of effort allocated in the firm. The minimum wage insures the worker against low outcomes of the wage negotiations, but it does not make a distinction between the case in which this low outcome is due to a low realization of the signal and the case in which it is due to a low realization of the marginal product. Therefore, the minimum wage encourages moral hazard. If, for example, the firm's bargaining power is low, the bargained wage is mainly determined by the worker's marginal product, and without a minimum wage, the worker would allocate most of her effort in the firm in order to achieve a high marginal product. Although the marginal return of effort allocated in the outside option is relatively high in this situation, the worker prefers to allocate most effort

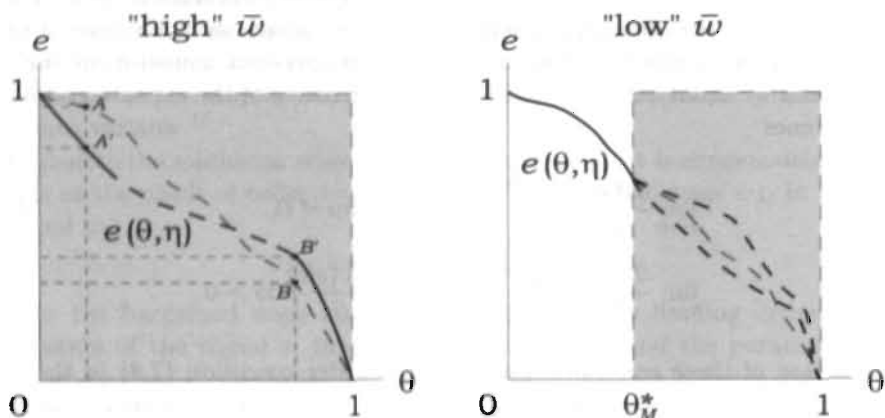


Figure 7.4: The worker's effort allocation in the minimum wage scenario. The shaded areas indicate the combinations of effort and the firm's bargaining power at which the minimum wage can be binding. The bold line depicts effort as a function of the firm's bargaining power. Its dashed part indicates an area for which the shape of the effort function can only be deduced from its behaviour in the vicinity of $\theta = 0$ and $\theta = 1$ (points A , A' , B , and B'). For comparison, the thin dashed line depicts the effort allocation of the Nash-bargaining scenario.

in the firm because also the probability that the signal realizes at an unfavourably low level is high. The minimum wage, however, protects the worker from this risk, and she can allocate more effort in her outside option. For the same reason, the minimum wage leads to more effort allocated in the firm if the firm's bargaining power is high.

Despite its inefficiencies, the worker will always favour the introduction of a minimum wage. Its effect on her expected value function is given by

$$\frac{dE[w_M|e]}{d\bar{w}} = \frac{\partial E[w_M|e]}{\partial e} \frac{de}{d\bar{w}} + \frac{\partial E[w_M|e]}{\partial \bar{w}} = \frac{\partial E[w_M|e]}{\partial \bar{w}} \geq 0.$$

The reaction of the expected profits at the worker's optimum to a change in the minimum wage is described by

$$\begin{aligned} \frac{dE[\pi_M|e]}{d\bar{w}} &= \left(\frac{\partial E[\pi_M|e]}{\partial e} - \frac{\partial E[w_M|e]}{\partial e} \right) \frac{de}{d\bar{w}} - \frac{\partial E[w_M|e]}{\partial \bar{w}} \\ &= \frac{\partial E[\pi_M|e]}{\partial e} \frac{de}{d\bar{w}} - \frac{\partial E[w_M|e]}{\partial \bar{w}}. \end{aligned}$$

Since the expected effective wage $E[w_M|e]$ is always an increasing function in \bar{w} , the effect on expected profits depends on the worker's reaction to a change in the

minimum wage. If the firm has high bargaining power, the worker will choose to allocate more effort in the firm, if the minimum wage increases. In this case, the firm gains from a higher expected marginal product, but it loses from a higher expected effective wage. Most likely, the overall effect is negative, as the changed effort allocation is a second-order effect. If, in contrast, the firm has low bargaining power, the effect on expected profits is unambiguously negative. A higher minimum wage will still lead to a higher expected effective wage, but the worker will choose to allocate less effort in the firm, which leads to a lower expected marginal product. Consequently, the firm will most likely oppose the introduction of a minimum wage in any case.

The results of this section indicate that a minimum wage that is imposed as a restriction on the bargained wage, is in general not the best way to achieve an efficient effort allocation, because it fails to give the adequate incentives. Still, it is questionable whether such a wage-setting scheme is realistic, anyway. If the minimum wage is the result of collective bargaining, as Aghion et al. argue, it should not just serve as a restriction on the outcome of negotiations between firm and worker, but influence the whole bargaining process. It is likely that in the course of bargaining, the parties' threat points are determined by the minimum (or, contract) wage, and not by their outside options. Models that recognize this are discussed in the chapters 4, 5, 6, and in chapter 3, section 5 of this thesis. In the following, we will investigate whether the application of two of these models, the one of Holden (1988, 1989), and that of MacLeod and Malcomson (1993), results in more efficient effort allocations.

7.4 FIXED WAGES AND RENEGOTIATIONS

This section is related to the asymmetric information case of the previous chapter 6. We consider a wage-setting scheme in which firm and worker still bargain about the effective wage, but in which their threat points (or, disagreement payoffs) are now determined by an exogenously given contract wage \bar{w} , which is the result of collective bargaining between unions and employer federations. A related bargaining game is proposed by MacLeod and Malcomson (1993) and a modified version of it is analysed in detail in chapter 6 of this thesis. It is based on the most simple assumption that the disagreement payoffs are just equal to \bar{w} for the worker, and $m - \bar{w}$ for the firm. This assumption can be justified by the common obligation binding on employers and unions to avoid industrial actions after an agreement has been reached in collective bargaining.¹⁷ Due to these disagreement payoffs it is not costly for either party to reject a wage proposal that deviates from the contract wage, unless the contract wage makes one of the outside options bind. In the latter case, rejecting a wage offer that secures for the party with the binding outside option a payoff that is at

¹⁷Such an obligation is known in virtually all countries that have a system of collective bargaining. Yet, apart from industrial action, both workers and firms have other means to make a disagreement about the effective wage costly. See, for example, the bargaining model of Holden (1988), which is discussed in the next section.

least equal to its outside option, will jeopardize the employment relation. Hence, the wage will be adapted to the outside option, giving the affected party a payoff exactly equal to the outside option, while the other party becomes residual claimant. It is important to notice that the adjustment will occur in mutual consent, because no party has an interest to put the employment relation at risk, and because the wage is only altered by an amount that is just necessary to keep the relation intact.

If we apply this bargaining structure to the model of this chapter, the *ex post* effective wage w_F in period 2 solves

$$w_F = \arg \max ((m - w) - (m - \bar{w}))^\theta (w - \bar{w})^{1-\theta}, \theta \in (0, 1)$$

s.t. $w \geq \bar{w}, a(s, \eta) \leq w \leq m - a_F.$

The first factor in the Nash maximand is equal to the difference between the firm's *ex post* payoff after an agreement has been reached, $m - w$, and its payoff during a disagreement, $m - \bar{w}$. This factor can be simplified to $-(w - \bar{w})$. The second factor is equal to the worker's agreement payoff w minus her disagreement payoff \bar{w} . Both factors are weighed by the respective bargaining power of the firm and the worker. As the two factors are just the negative of each other, the Nash maximand has an unrestricted maximum only at $w_F = \bar{w}$, irrespective of the distribution of bargaining power.¹⁸ The reason for the irrelevance of bargaining power is that, during a disagreement, neither the firm nor the worker can inflict bargaining cost on their respective opponent, so their joint payoff after an agreement is reached cannot be greater than their joint payoff in a disagreement. Consequently, there is no Pareto-efficient wage other than the contract wage, unless the contract wage makes one of the outside options bind. Hence,

$$w_F = \begin{cases} w_{F1} = \bar{w} & \text{if } a(s, \eta) \leq \bar{w} \leq m - a_F \\ w_{F2} = m - a_F & \text{if } m - \bar{w} < a_F \\ w_{F3} = a(s, \eta) & \text{if } \bar{w} < a(s, \eta) \end{cases}$$

As was pointed out in chapter 6, this effective wage can lead to efficient specific investments, i.e., investments that generate returns which accrue directly to the investing party, because hold up is avoided. In this chapter, however, a different kind of investment is considered: the effort the worker allocates (or, invests) in the firm does not accrue directly to herself but first of all to the firm, which then pays a wage to the worker. In the following, we will show that this change in the assumption about the kind of investment will in general lead to inefficient outcomes.

In order to derive the worker's effort allocation, three cases have to be distinguished. In the first case, the contract wage $\bar{w} = \bar{w}_1$ cannot make any outside option bind *a priori*, i.e., $a(s_1, \eta) \leq \bar{w}_1 \leq m_0 - a_F$. When deciding about her effort allocation in period 1, the worker then knows that the expected wage is equal to the

¹⁸This is not true in case either party has full bargaining power ($\theta = 0$ or $\theta = 1$). This case is excluded by assumption. However, in the original noncooperative bargaining game, which is discussed in chapter 6, a party with full bargaining power cannot carry through a wage proposal different from the contract wage, because the other party can still reject it.

contract wage,

$$E[w_{F1}|e] = \bar{w}_1, \quad (7.17)$$

and that her decision has no influence on her payoff. Consequently, the worker is indifferent between all possible effort allocations, unless additional factors play a role. Such factors, all beyond the scope of this model, could be the duration of the employment relation, or loyalty issues. For example, the worker might allocate all her effort in the firm expecting that in return, if economic circumstances worsen, the firm will not lay her off. If, in contrast, the worker faces a certain layoff probability that is independent of her performance, she will prefer to allocate all her effort in her outside option.

The second case comprises all contract wages \bar{w}_2 that can make the firm's outside option bind, i.e., $m_0 - \bar{w}_2 < a_F$. Since by equation (7.6) only one of the outside options can be binding, the worker's expected payoff in this case does not depend on her outside option. Her objective is

$$\begin{aligned} \max_e E[w_2|e] &= \int_{m_0}^{m_{F2}^*} (m - a_F) g(m|e) dm + \int_{m_{F2}^*}^{m_1} \bar{w}_2 g(m|e) dm \\ &= \bar{w}_2 - \int_{m_0}^{m_{F2}^*} G(m|e) dm, \end{aligned} \quad (7.18)$$

where m_{F2}^* is the critical level of the marginal product that makes the firm's *ex post* payoff exactly equal to its outside option, $m_{F2}^* - \bar{w}_2 = a_F$. If the marginal product realizes at levels $m \geq m_{F2}^*$, the firm's outside option will not be binding, and the worker will receive the contract wage. But for lower realizations of the marginal product, the effective wage has to be adapted to the firm's outside option, and the worker will only receive a wage of $w_{F2} = m - a_F$. The derivative of the worker's objective,

$$\frac{\partial E[w_{F2}|e]}{\partial e} = - \int_{m_0}^{m_{F2}^*} \frac{\partial G(m|e)}{\partial e} dm \geq 0,$$

is nonnegative by equation (7.1). This means, that the worker will allocate all her effort in the firm, in order to avoid that the firm's outside option becomes binding.

In the last case, the contract wage \bar{w}_3 is set such that only the worker's outside option can become binding, i.e., $\bar{w}_3 < a(s_1, 0)$. The expected utility of the worker now depends on the relation between the worker's true effort allocation and the firm's belief about it. Depending on this relation, two subcases have to be distinguished.

In the first subcase, the worker chooses to allocate more effort in the employment relation than the firm predicts, $e > \eta$. The firm's perception of the worker's outside option is always greater than the worker's true outside option, $a(s, \eta) > a(s, e) \forall s$, and the firm will voluntarily adjust the wage at signals s for which an adjustment is not necessary, and it will adjust the wage to an extent that is not necessary in order to maintain the employment relation.

We denote the critical signal that makes the perceived outside option exactly equal to contract wage, $a(s_{F3a}^*, \eta) = \bar{w}_3$, by s_{F3a}^* . The worker's expected effective

wage (and payoff) can now be written as

$$\begin{aligned} E[w_{F3a}|e] &= \int_{s_0}^{s_{F3a}^*} \bar{w}_3 h(s, e) ds + \int_{s_{F3a}^*}^{s_1} a(s, \eta) h(s|e) ds \\ &= a(s_1, \eta) - \int_{s_{F3a}^*}^{s_1} \frac{\partial a(s, \eta)}{\partial s} H(s|e) ds. \end{aligned} \quad (7.19)$$

The worker maximizes its expected payoff over the amount of effort allocated in the firm. Because the derivative of the objective,

$$\frac{\partial E[w_{F3a}|e]}{\partial e} = - \int_{s_{F3a}^*}^{s_1} \frac{\partial a(s, e)}{\partial s} \frac{\partial H(s|e)}{\partial e} ds \leq 0,$$

is nonpositive for all e , the worker will allocate all her effort in her outside option. Because this is a contradiction to the original assumption that $e > \eta$, this first case can never occur if the worker acts rationally.

In the second subcase, the firm's belief is greater than the real amount of effort allocated in the firm, $e < \eta$, i. e., the firm will underestimate the worker's outside option for all realizations of the signal, $a(s, \eta) < a(s, e) \forall s$. If the worker's outside option now becomes binding, the firm will reject a wage adjustment that is actually necessary to save the employment relation. Either it will reject any adaptation of the wage, or it will only grant a wage increase that is insufficient to make the worker prefer to stay in the firm. In either case, the worker chooses her outside option. Therefore, her expected payoff is

$$\begin{aligned} E[u_{F3b}|e] &= \int_{s_0}^{s_{F3b}^*} \bar{w}_3 h(s|e) ds + \int_{s_{F3b}^*}^{s_1} a(s, e) h(s|e) ds \\ &= a(s_1, e) - \int_{s_{F3b}^*}^{s_1} \frac{\partial a(s, e)}{\partial s} H(s|e) ds, \end{aligned}$$

where s_{F3b}^* is the critical signal at which the contract wage is exactly equal to the worker's outside option, $\bar{w} = a(s_{F3b}^*, e)$. The derivative of the expected payoff is calculated as

$$\begin{aligned} \frac{\partial E[u_{F3b}|e]}{\partial e} &= \frac{ds_{F3b}^*}{de} (\bar{w} - a(s_{F3b}^*, e)) h(s_{F3b}^*|e) + \int_{s_0}^{s_{F3b}^*} \bar{w} \frac{\partial h(s|e)}{\partial e} ds \\ &\quad + \int_{s_{F3b}^*}^{s_1} \left(\frac{\partial a(s, e)}{\partial e} h(s|e) + a(s, e) \frac{\partial h(s|e)}{\partial e} \right) ds \\ &= \int_{s_{F3b}^*}^{s_1} \frac{\partial a(s, e)}{\partial e} h(s|e) ds - \int_{s_{F3b}^*}^{s_1} \frac{\partial a(s, e)}{\partial s} \frac{\partial H(s|e)}{\partial e} ds < 0. \end{aligned}$$

Because it is less than zero, the worker will again allocate all her effort in her outside option. Obviously, the firm's belief does not matter for the worker's decision, because both the real outside option and the firm's perception of the outside option are

increasing in the signal. With the chosen effort allocation, the worker maximizes the expected signal and thereby the probability that her outside option is greater than the contract wage, without having to fear any repercussions from neglecting her marginal product. Thus, a low contract wage results in a moral hazard problem that is similar to the one that occurs under the wage-setting scheme with Nash bargaining augmented by a minimum wage.¹⁹

Although a contract wage less than the maximal outside option $a(s_1, 0)$ leads to an inefficient effort allocation, it will not put the employment relation at risk, because a rational firm will anticipate the worker's decision and form the belief $\eta = 0$. Therefore, there will always be a mutual consent about a necessary wage adjustment, and the equilibrium expected payoff is equal to $E[w_{F3a}|0]$.

In contrast to the wage-setting schemes we have discussed so far, an infinitesimally small change in the contract (or, fixed) wage \bar{w} can now lead to a discrete change in the effort allocation, namely if this change results in the switch from one case to another. This may have consequences for the firm's and the worker's preferences about the fixed wage. As an inspection of the relevant expected payoffs of the worker in the three cases, (7.17), (7.18), and (7.19), shows, the worker will benefit from a higher fixed wage in any case, similar to the minimum wage scenario of the previous section. The expected payoffs of the firm in the three cases, given the choices of the worker, are equal to

$$E[\pi|e] = \begin{cases} E[\pi_1|e] = E[m|e] - \bar{w}, & e \in [0, 1] & \text{if } a(s_1, 0) \leq \bar{w} \leq m_0 - a_F \\ E[\pi_2|1] = E[m|1] - E[w_{F2}|1] & & \text{if } m_0 - \bar{w} < a_F \\ E[\pi_3|0] = E[m|0] - E[w_{F3a}|0] & & \text{if } \bar{w} < a(s_1, 0) \end{cases}.$$

Within each of the three cases, the firm will prefer a contract wage as low as possible, because the effort allocation does not change and the expected marginal product is constant, while the expected wages are increasing in the contract wage. However, if a change in the contract wage leads to a switch between the cases, the firm might prefer a higher contract wage.

In order to show that such a situation can (but need not) exist, suppose that the minimal joint payoff that can be generated in the employment relation, $m_0 - a_F - a(s_1, 0)$, is very small such that $a(s_1, 0) \approx m_0 - a_F$. Compare two contract wages \bar{w}_2 and \bar{w}_3 , and assume that the former is infinitesimally greater than $m_0 - a_F$ and can make the firm's outside option bind, while the latter is infinitesimally smaller than $a(s_1, 0)$ and can make the worker's outside option bind. Now compare the expected effective wages related to \bar{w}_2 and \bar{w}_3 ,

$$\lim_{\bar{w}_2 \rightarrow m_0 - a_F} E[w_{F2}|1] = \lim_{\bar{w}_2 \rightarrow m_0 - a_F} \bar{w}_2 - \int_{m_0}^{m_{F2}^*} G(m|1) dm = m_0 - a_F,$$

$$\lim_{\bar{w}_3 \rightarrow a(s_1, 0)} E[w_{F3a}|0] = \lim_{\bar{w}_3 \rightarrow a(s_1, 0)} a(s_1, 0) - \int_{s_{F3a}^*}^{s_1} \frac{\partial a(s, 0)}{\partial s} H(s|0) ds = a(s_1, 0).$$

¹⁹See section 7.3.3, p. 157.

Because $a(s_1, 0) \approx m_0 - a_F$, the firm expects that it has to pay virtually the same effective wage in both cases. At the same time, the expected marginal product is greater for \bar{w}_3 , because the worker allocates all her effort in the firm.

In general, the firm will approve a contract-wage increase that leads to a switch from case 3 to case 2, if the resulting increase in the expected marginal product is great enough to compensate for the higher expected wage due to the switch. The described wage-setting scheme can then lead to an efficient outcome, even if the contract wage was initially so low that it could make the worker's outside option bind.

7.5 LOCAL WAGE BARGAINING

Holden (1988) proposes a negotiation game which he claims to be suitable for negotiations between firms and workers on the level of individual firms ("local" wage bargaining), after the contract wage has been set by collective bargaining. Similar to the previously discussed bargaining game, Holden starts out from the idea that during local wage bargaining, both parties are obliged to avoid industrial actions, such that the disagreement payoffs are determined by the contract wage rather than the outside options. However, different from the former model, he assumes that workers can express their dissatisfaction by exerting themselves less in the job—for example by working to rule, which is not considered to be industrial action. By lowering their marginal product during local wage bargaining, workers make a disagreement costly to the firm and can obtain an effective wage that is greater than the contract wage set by collective bargaining, even if no outside option is binding.

The Holden model is discussed in detail in chapter 3.5 of this thesis. At this place, we use a slightly different version adapted to bargaining between a single worker and the firm, in order to investigate whether it can lead to an efficient effort allocation of the worker.²⁰ During a disagreement, the worker can decrease her marginal product by a factor $\alpha \in (0, 1)$, while the firm has to pay the contract wage \bar{w} . Consequently, the disagreement payoff of the firm is equal to $\alpha m - \bar{w}$, while the worker receives \bar{w} . Different from the original Holden model, we make a provision for the case that the bargaining outcome makes one of the (perceived) outside options bind. In order to avoid an inefficient quit or layoff, we assume that in this case, firm and worker mutually agree on adapting the effective wage such that the party with the binding outside option receives exactly its (perceived) outside option. The effective wage is then

$$\begin{aligned} w_H &= \arg \max ((m - w) - (\alpha m - \bar{w}))^\theta (w - \bar{w})^{1-\theta} \\ &= \arg \max ((1 - \alpha)m - (w - \bar{w}))^\theta (w - \bar{w})^{1-\theta} \\ \text{s.t. } & w \geq \bar{w}, a(s, \eta) \leq w \leq m - a_F. \end{aligned}$$

²⁰The term "local wage bargaining" for the wage-formation process described in this section is chosen in order to refer to Holden's model. Notice, however, that *all* the wage-setting schemes discussed in this chapter are in fact based on some kind of local bargaining on the firm's level.

The first factor is the difference between the firm's payoffs during and after the negotiations, weighed by its bargaining power, and the second term is the related difference of payoffs of the worker, weighed by her bargaining power. Notice that, in contrast to the renegotiation scenario considered in the previous section, a disagreement is now costly to the bargaining parties. Due to the reduced effort during a disagreement, reaching an agreement will create a surplus, and the negotiations are just about sharing out this surplus between the parties. Therefore, bargaining power matters in this scenario. The resulting effective wage is equal to

$$w_H = \begin{cases} w_{H1} = \bar{w} + (1 - \theta)(1 - \alpha)m & \text{if } a(s, \eta) \leq w_{H1} \leq m - a_F \\ w_{H2} = m - a_F & \text{if } m - w_{H1} < a_F \\ w_{H3} = a(s, \eta) & \text{if } w_{H1} < a(s, \eta) \end{cases}$$

If the bargained wage makes neither outside option bind, the effective wage is equal to the contract wage plus a mark-up, which positively depends on the marginal product, the worker's bargaining power, and her ability to reduce the marginal product during the negotiations. The previously discussed scenario of a fixed wage with renegotiations can be interpreted as a special case of the local wage-bargaining scenario, in which the firm has full bargaining power ($\theta = 1$), or in which the worker has no ability to reduce the marginal product during a disagreement ($\alpha = 1$). As this case was investigated before, we confine the analysis to situations where $(1 - \theta)(1 - \alpha) > 0$.

To determine the worker's decision about her effort allocation in period 1 we must again distinguish three cases, one in which the bargained wage makes neither outside option binding (case 1), another in which only the firm's outside option can become binding (case 2), and a third, in which only the worker's outside option can become binding (cases 3a and 3b). In the latter, we must again make a provision for the case that the firm underestimates the worker's real outside option.

The bargaining outcome can make neither outside option bind *a priori* if $a(s_1, \eta) \leq \bar{w} + (1 - \theta)(1 - \alpha)m_0 \leq m_0 - a_F$. In this case, the worker expects the effective wage to be

$$\begin{aligned} E[w_{H1}|e] &= \bar{w} + (1 - \theta)(1 - \alpha)E[m|e] \\ &= \bar{w} + (1 - \theta)(1 - \alpha) \int_{m_0}^{m_1} mg(m|e) dm. \end{aligned}$$

The expected wage is an increasing function of the expected marginal product. Therefore, it depends positively on the amount of effort allocated in the firm: the wage provides an incentive to the worker to allocate all her effort in the firm.

In the second case, the bargained wage can make the firm's outside option bind, $\bar{w} + (1 - \theta)(1 - \alpha)m_0 > m_0 - a_F$. We denote the critical marginal product that makes the bargained wage exactly equal to the outside option by m_{H2}^* , i. e., $\bar{w} + (1 - \theta)(1 - \alpha)m_{H2}^* = m_{H2}^* - a_F$, or equivalently,

$$m_{H2}^* = \frac{a_F + \bar{w}}{1 - (1 - \theta)(1 - \alpha)}.$$

The bargaining wage w_{H1} makes the firm's outside option bind, if the marginal product realizes at a level lower than m_{H2}^* . Then, the worker becomes residual claimant and earns $m - a_F$ instead of the bargaining wage. Her expected effective wage is now

$$\begin{aligned} E[w_{H2}|e] &= \int_{m_0}^{m_{H2}^*} (m - a_F) g(m|e) dm \\ &\quad + \int_{m_{H2}^*}^{m_1} [\bar{w} + (1 - \theta)(1 - \alpha)m] g(m|e) dm \\ &= \bar{w} + (1 - \theta)(1 - \alpha)m_1 - \int_{m_0}^{m_{H2}^*} G(m|e) dm \\ &\quad - (1 - \theta)(1 - \alpha) \int_{m_{H2}^*}^{m_1} G(m|e) dm. \end{aligned}$$

Its derivative with respect to the amount of effort allocated in the firm is nonnegative,

$$\frac{\partial E[w_{H2}|e]}{\partial e} = - \int_{m_0}^{m_{H2}^*} \frac{\partial G(m|e)}{\partial e} dm - (1 - \theta)(1 - \alpha) \int_{m_{H2}^*}^{m_1} \frac{\partial G(m|e)}{\partial e} dm \geq 0.$$

Because the effective wage depends on the realization of the marginal product, but not on the worker's outside option, the worker will allocate all her effort in the firm.

In contrast to the two previously discussed cases, the worker's expected payoff depends on her (perceived) outside option in the third case. The bargaining wage can make her outside option bind if $\bar{w} + (1 - \theta)(1 - \alpha)m_0 \leq a(s_1, 0)$. Because the wage can only be adjusted by mutual consent, the effective wage depends on the firm's perception of the outside option. Same as in the related case in the previous section, we will first assume that the firm either holds a correct belief, or that it holds a belief that does not lead to an underestimation of the worker's outside option. Consequently, the worker will never quit for her outside option. This subcase, in which $e \geq \eta$, will be indicated by the index $H3a$.

If the signal exceeds a critical level s_{H3a}^* , at which the bargained wage w_{H1} and the perceived outside option are exactly equal,

$$\bar{w} + (1 - \theta)(1 - \alpha)m = a(s_{H3a}^*, \eta),$$

the effective wage will be adjusted to the outside option as it is perceived by the firm. Moreover, there might be a critical m_{H3a}^* , defined by the equation

$$\bar{w} + (1 - \theta)(1 - \alpha)m_{H3a}^* := a(s_1, \eta),$$

such that for realizations of the marginal product greater than this critical level there are no signals that can make the outside option bind. The expected effective wage is in this case equal to the expected bargained wage plus an additional payoff A_{H3a} , that reflects the situations in which the perceived outside option is binding:

$$E[w_{H3a}|e] = E[w_{H1}|e] + A_{H3a}$$

The expected additional payoff A_{H3a} is derived in general terms in the appendix of this chapter on the pages 179 ff. Applying equation (7.35) to the local wage-bargaining scenario yields

$$\begin{aligned} A_{H3a} &= \int_{m_0}^{m_{H3a}^*} \int_{s_{H3a}^*}^{s_1} [a(s, \eta) - \bar{w} - (1 - \theta)(1 - \alpha)m] h(s|e) g(m|e) ds dm \\ &= (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3a}^*} [1 - H(s_{H3a}^*|e)] G(m|e) dm. \end{aligned}$$

By equation (7.36) its derivative with respect to e is negative,

$$\begin{aligned} \frac{\partial A_{H3a}}{\partial e} &= \int_{m_0}^{m_{H3a}^*} \int_{s_{H3a}^*}^{s_1} [a(s, \eta) - \bar{w} - (1 - \theta)(1 - \alpha)m] \\ &\quad \times \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\ &= (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3a}^*} [1 - H(s_{H3a}^*|e)] \frac{\partial G(m|e)}{\partial e} dm \\ &\quad - (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3a}^*} \frac{\partial H(s_{H3a}^*|e)}{\partial e} G(m|e) dm < 0, \quad (7.20) \end{aligned}$$

because the expected (positive) difference between outside option and bargained wage becomes smaller the more effort the worker allocates in the firm. Without specifying the functional forms of the worker's outside option and the density functions of the signal and the marginal product, it is impossible to solve for the worker's optimal effort allocation. However, we can evaluate the derivative of the effective wage in the vicinities of $e = 1$, and $e = 0$. By equations (7.1) and (7.2) the derivative of the worker's objective converges to

$$\begin{aligned} \lim_{e \rightarrow 0} \frac{\partial E[w_{H3a}|e]}{\partial e} &= -(1 - \theta)(1 - \alpha) \int_{m_{H3a}^*}^{m_1} \frac{\partial G(m|e)}{\partial e} dm \\ &\quad - (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3a}^*} H(s_{H3a}^*|e) \frac{\partial G(m|e)}{\partial e} dm > 0, \\ \lim_{e \rightarrow 1} \frac{\partial E[w_{H3a}|e]}{\partial e} &= -(1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3a}^*} \frac{\partial H(s_{H3a}^*|e)}{\partial e} G(m|e) dm < 0. \end{aligned}$$

The signs imply that the effort allocation that maximizes the expected wage has an interior solution $e \in (0, 1)$. The worker will *not* allocate all her effort in the firm, because at $e = 1$, the expected marginal return of effort allocated in the outside option is greater than the expected marginal return of effort allocated in the firm (which is equal to zero). Different from the fixed-wage scenario of section 7.4, however, the worker also will not allocate all her effort in her outside option.

According to the efficiency criterion introduced in the introduction to this chapter, local wage bargaining performs better than the fixed wage with renegotiations.

Which effort allocation the worker will ultimately choose, depends on the probability that the worker's outside option becomes binding, and on the size of the expected additional payoff. Both factors are in turn determined by the contract wage, the worker's share in her marginal product, and the firm's belief about the worker's effort allocation. These three parameters influence the derivative of the expected additional payoff. We apply equation (7.37) from the appendix to get

$$\begin{aligned} \frac{\partial^2 A_{H3a}}{\partial e \partial \bar{w}} &= \int_{m_0}^{m_{H3a}^*} \frac{\partial H(s_{H3a}^*|e)}{\partial e} g(m|e) dm \\ &\quad - \int_{m_0}^{m_{H3a}^*} h(s_{H3a}^*|e) \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_{H3a}^*} \frac{(1-\theta)(1-\alpha)}{\partial e} \frac{\partial G(m|e)}{\partial e} dm > 0, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 A_{H3a}}{\partial e \partial (1-\theta)(1-\alpha)} &= \int_{m_0}^{m_{H3a}^*} m \frac{\partial H(s_{H3a}^*|e)}{\partial e} g(m|e) dm \\ &\quad - \int_{m_0}^{m_{H3a}^*} mh(s_{H3a}^*|e) \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_{H3a}^*} \frac{(1-\theta)(1-\alpha)}{\partial e} \frac{\partial G(m|e)}{\partial e} dm > 0, \end{aligned}$$

and finally

$$\begin{aligned} \frac{\partial^2 A_{H3a}}{\partial e \partial \eta} &= - \int_{m_0}^{m_{H3a}^*} \frac{\partial a(s_{H3a}^*, \eta)}{\partial \eta} \frac{\partial H(s_{H3a}^*|e)}{\partial e} g(m|e) dm \\ &\quad - \int_{m_0}^{m_{H3a}^*} \int_{s_{H3a}^*}^{s_1} \frac{\partial^2 a(s, \eta)}{\partial s \partial \eta} \frac{\partial H(s|e)}{\partial e} g(m|e) ds dm \\ &\quad + (1-\theta)(1-\alpha) \int_{m_0}^{m_{H3a}^*} \left. \frac{\partial a(s, \eta)}{\partial \eta} \frac{\partial a(s, \eta)}{\partial s} \right|_{s_{H3a}^*} h(s_{H3a}^*|e) \frac{\partial G(m|e)}{\partial e} dm > 0. \end{aligned}$$

A greater contract wage, higher bargaining power of the worker, a greater ability to lower the marginal product during a disagreement, and a greater belief of the firm raise the bargained wage w_{H1} , and thereby lead to a lower expected additional payoff. Therefore, it becomes less attractive for the worker to allocate effort in her outside option. As the total differential of the first order condition $\frac{\partial E[w_{H3a}|e]}{\partial e} = 0$

shows, the worker will in these cases allocate more effort in the firm,²¹

$$\begin{aligned} \frac{de}{d\bar{w}} &= -\frac{\frac{\partial^2 E[w_{H3a}|e]}{\partial e \partial \bar{w}}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} = -\frac{\frac{\partial^2 A_{H3a}}{\partial e \partial \bar{w}}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} > 0, \\ \frac{de}{d(1-\theta)(1-\alpha)} &= -\frac{\frac{\partial^2 E[w_{H3a}|e]}{\partial e \partial (1-\theta)(1-\alpha)}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} = -\frac{\frac{\partial E[m|e]}{\partial e} + \frac{\partial^2 A_{H3a}}{\partial e \partial (1-\theta)(1-\alpha)}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} > 0, \\ \frac{de}{d\eta} &= -\frac{\frac{\partial^2 E[w_{H3a}|e]}{\partial e \partial \eta}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} = -\frac{\frac{\partial^2 A_{H3a}}{\partial e \partial \eta}} {\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}} > 0. \end{aligned}$$

Notice that the effort allocation is more sensitive to a change of the parameters describing the worker's bargaining power, than to a change of the contract wage. The reason is that the factor $(1-\theta)(1-\alpha)$ determines how strongly the bargained wage w_{H1} depends on the worker's marginal product, which in turn determines the incentive effect of the bargained wage. The contract wage, in contrast, only co-determines the level of the bargained wage.

The described effort allocation holds only if the worker accepts the effective wage w_{H3} , and if she does not leave the firm because her real outside option is greater than the firm's perception of it. If, however, the firm holds a belief $\eta > e$, the worker will be forced to quit if her real outside option is binding, because the firm will not agree to the necessary adjustment of the wage. In the last part of this section we will discuss how the worker's behaviour changes, if the firm overestimates the amount of effort allocated in the firm, and thereby underestimates the worker's outside option. This subcase will be indicated by the index $H3b$.

When we replace the outside option as perceived by the firm with the worker's real outside option, the effort allocation does not only influence the probability that the outside option becomes binding, but also the difference between outside option and bargained wage. In the equation of the expected additional payoff,

$$\begin{aligned} A_{H3b} &= \int_{m_0}^{m_{H3b}^*} \int_{s_{H3b}^*}^{s_1} [a(s, \eta) - \bar{w} - (1-\theta)(1-\alpha)m] h(s|e) g(m|e) ds dm \\ &= (1-\theta)(1-\alpha) \int_{m_0}^{m_{H3b}^*} [1 - H(s_{H3b}^*|e)] G(m|e) dm, \end{aligned}$$

the critical values of the signal and the marginal product, s_{H3b}^* and m_{H3b}^* , are now functions of e . Applying equation (7.36) from the appendix to this case, the derivative of the expected additional payoff with respect to the amount of effort allocated in

²¹ Since we evaluate the total differential at the utility maximizing level of e , the second derivative $\frac{\partial^2 E[w_{H3a}|e]}{\partial e^2}$ must be less than zero.

the firm can be written as

$$\begin{aligned} \frac{\partial A_{H3b}}{\partial e} &= \int_{m_0}^{m_{H3b}^*} \int_{s_{H3b}^*}^{s^1} \frac{\partial a(s, e)}{\partial e} h(s|e) g(m|e) ds dm \\ &+ (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3b}^*} [1 - H(s_{H3b}^*|e)] \frac{\partial G(m|e)}{\partial e} dm \\ &- (1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3b}^*} \frac{\partial H(s_{H3b}^*|e)}{\partial e} G(m|e) dm < 0. \end{aligned}$$

The first summand describes the effect of a change in e on the size of the additional payoff, while the last two summands describe its effect on the probability that the outside option becomes binding. The latter are similar to the corresponding equation (7.20) for the case 3a, in which the worker accepts the effective wage w_{H3} . With all three summands being negative, the worker has an even greater incentive to allocate effort in her outside option, than in the case where the effective wage w_{H3} is acceptable. In contrast to that case, the derivative of the expected additional payoff can now also be negative in the vicinity of $e = 0$:

$$\begin{aligned} \lim_{e \rightarrow 0} \frac{\partial E[u_{H3a}|e]}{\partial e} &= -(1 - \theta)(1 - \alpha) \int_{m_0}^{m_{H3b}^*} H(s_{H3b}^*|e) \frac{\partial G(m|e)}{\partial e} dm \\ &+ \int_{m_0}^{m_{H3b}^*} \int_{s_{H3b}^*}^{s^1} \frac{\partial a(s, e)}{\partial e} h(s|e) g(m|e) ds dm \geq 0 \end{aligned}$$

As the derivative in the vicinity of $e = 1$ is negative anyway,²² there is now a possibility—depending on the parameters and the distributions of the signal and the marginal product—that the worker will choose $e = 0$ and allocate all her effort in her outside option. However, the greater the contract wage, the worker's bargaining power, or her ability to lower the marginal product during a disagreement are, the smaller is the difference between the outside option and the effective wage, and the lower is the incentive to allocate effort in the outside option. The derivative of the

²²This is true in *general* if the worker's outside option can be binding. See also equation (7.38) on page 185 in the appendix.

expected additional payoff is increasing in these parameters:

$$\begin{aligned} \frac{\partial^2 A_{H3b}}{\partial e \partial \bar{w}} &= \int_{m_0}^{m_{H3b}^*} \frac{\partial H(s_{H3b}^*|e)}{\partial e} g(m|e) dm \\ &\quad - \int_{m_0}^{m_{H3b}^*} \left. \frac{\frac{\partial a(s,e)}{\partial e}}{\frac{\partial a(s_{H3b}^*,e)}{\partial s}} \right|_{s_{H3b}^*} h(s_{H3b}^*|e) g(m|e) dm \\ &\quad - (1-\theta)(1-\alpha) \int_{m_0}^{m_{H3b}^*} \left. \frac{h(s_{H3b}^*|e)}{\frac{\partial a(s,e)}{\partial s}} \right|_{s_{H3b}^*} \frac{\partial G(m|e)}{\partial e} dm > 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 A_{H3b}}{\partial e \partial (1-\theta)(1-\alpha)} &= \int_{m_0}^{m_{H3b}^*} m \frac{\partial H(s_{H3b}^*|e)}{\partial e} g(m|e) dm \\ &\quad - \int_{m_0}^{m_{H3b}^*} m \left. \frac{\frac{\partial a(s,e)}{\partial e}}{\frac{\partial a(s_{H3b}^*,e)}{\partial s}} \right|_{s_{H3b}^*} h(s_{H3b}^*|e) g(m|e) dm \\ &\quad - (1-\theta)(1-\alpha) \int_{m_0}^{m_{H3b}^*} m \left. \frac{h(s_{H3b}^*|e)}{\frac{\partial a(s,e)}{\partial s}} \right|_{s_{H3b}^*} \frac{\partial G(m|e)}{\partial e} dm > 0 \end{aligned}$$

Again, the effect of a change of $(1-\theta)(1-\alpha)$ will be stronger than the effect of a change of the contract wage \bar{w} .

When forming its belief, a rational firm will anticipate the worker's behaviour. As we argued in section 7.3.3, the firm has no interest to provoke the worker into quitting, and will announce its belief truthfully. As long as there are multiple fixed points $e = \eta$, the firm will naturally pick the greatest one, which yields the greatest expected marginal product and the lowest expected effective wage. Otherwise, if only a single fixed point exists, the firm can achieve a more efficient effort allocation only via the bargaining outcome—by conceding some of its bargaining power (lower θ) or by allowing a greater reduction of the marginal product (lower α) during a disagreement. Most likely, however, the firm has no interest to do so, as both a reduced θ and a reduced α have a direct negative impact on the wage cost, while the resulting increase in effort allocated in the firm yields only an indirect positive effect on the expected marginal product. This is different from the related case of the scenario with fixed wages and renegotiations (see section 7.4), where a change in parameters can possibly induce a switch from $e = 0$ to $e = 1$, and lead to an increase in expected profits, while the wage cost go up, too. In general, however, the local wage-bargaining scenario leads to more efficient effort allocations, because for any constellation of the contract wage, there is a *strong* incentive to allocate effort in the firm, whereas in the fixed-wage scenario, strong incentives are only provided if the firm's outside option threatens to become binding.

7.6 CONCLUSIONS

This chapter has analysed an employment relation between an individual worker and her employer, in which the worker can decide how to allocate one unit of effort over two uses: in the firm, in order to increase her marginal product, or outside the firm, in order to increase her outside option. The firm can observe neither the effort allocation nor the worker's outside option, but has to form a belief about the two variables. Using this simple structure, we have discussed the efficiency aspects of different wage-setting schemes under asymmetric information. Two yardsticks were used for efficiency: first, whether a wage-setting scheme provides the necessary incentives to the worker to allocate all her effort in the firm, thereby maximizing the joint payoff, and second, whether the wage-setting scheme is able to maintain the employment relationship, which in principle generates a greater joint payoff than the sum of the outside options of firm and worker. Table 7.1 on page 176 provides an overview of the wage-setting schemes discussed in this chapter with respect to their efficiency.

The result concerning the second problem is that in none of the discussed wage-setting schemes the asymmetric information about the worker's effort allocation puts the employment relation at risk. In all cases, the wage-setting schemes revealed enough information to the firm to deduce the effort allocation of the worker. Therefore, the only possible reason for an (inefficient) separation would be that the firm acts dumb and, in order to achieve a lower wage, pretends to believe that the worker will allocate less effort in her outside option than she actually will. We have shown, however, that such an attempt to deceive will either not be feasible, because the worker looks through it, or it can even backfire on the firm: a lower wage increases the probability that the worker's outside option can become binding, which induces the worker to allocate less effort in the firm. At the end, the firm will see its expected profit decrease.

The first problem about the incentives of the different wage-setting schemes has proved to be more complicated. All the wage-setting schemes under consideration have in common that the effective wage is determined by bargaining between a single worker and her employer. In section 7.3, we have analysed simple Nash bargaining between firm and employer about the effective wage. The bargaining outcome turns out to be inefficient because it depends positively on both, the worker's marginal product and (the firm's perception of) her outside option. The more bargaining power the firm has, the smaller is the share of the joint payoff the worker receives, and the more the effective wage depends on the outside option, which makes it increasingly unattractive to allocate effort in the firm. The firm, on the other hand, can in principle achieve a more efficient effort allocation, either by misreporting its perception of the worker's outside option, or by giving up some bargaining power and letting the worker enjoy a greater share of her marginal product. As turns out, however, the former is not feasible or even jeopardizes the employment relation, whereas the latter involves cost that the firm is most likely not willing to bear.

In the last part of section 7.3, a minimum wage was introduced in the Nash-

bargaining scenario, in order to check the argument of Aghion, Compte, and Jehiel (1998), that it could induce the worker to allocate more effort in the firm. Aghion, Compte, and Jehiel identify the minimum wage with the outcome of collective bargaining between a union and employers, and reason that, due to the claimed effect on the worker's effort allocation, unions can enhance efficiency. Our results qualify this claim because the minimum wage encourages moral hazard. It provides insurance against unfavourable realizations of the marginal product *and* the outside option, which induces the worker to allocate her effort more evenly over its two uses. While this leads to more effort allocated in the firm if the firm's bargaining power is high and the bargaining wage is mainly determined by the worker's outside option, it has the contrary effect if the firm's bargaining power is low.

Aghion, Compte, and Jehiel's (1998) claim about the efficiency enhancing effect of unions would even turn out to be wrong if, as is likely, the bargaining power of employees in the firm is positively related to the bargaining power of their union. If the position of the employees vis-à-vis the employer is weak, the Nash-bargaining wage will be low and most effort is allocated outside the firm. But if the union is also weak, it will typically achieve a low contract (minimum) wage, which only provides little insurance and will not change the effort allocation much. On the other hand, if the bargaining position of both the union and the individual employees is strong, the resulting high contract wage will even decrease the amount of effort allocated in the firm.

Sections 7.4 and 7.5 analyse two wage-setting schemes that are also discussed in another context elsewhere in this thesis. Section 7.4 discusses again the renegotiation-game model of chapter 6. In that chapter, it was argued that a wage-setting scheme that provides for a fixed contract wage, which can only be adjusted if the outside option of either the firm or the worker becomes binding, avoids hold-up. It leads to efficient specific investments (with returns accruing directly to the investing party), while avoiding inefficient quits and layoffs. As turns out, however, this wage-setting scheme in most cases fails to lead to an efficient effort allocation. With a fixed wage, the worker is neither encouraged to exert herself more in her job, nor to invest in her outside option. Strong incentives are only given if the contract wage can make one of the outside options bind. If this is the firm's outside option, then the worker will allocate all her effort in the firm, in order to avoid that her wage must be reduced to prevent the firm from laying her off. If, however, the contract wage is low enough to possibly make the worker's outside option bind, the worker has an incentive to maximize her outside option by allocating all her effort outside the firm. Consequently, the effort allocation is extremely sensitive to the size of the contract wage: efficient incentives are given only in case the contract wage is sufficiently high to possibly make the firm's outside option bind. In the other cases, either no incentives are given at all, or they lead to the least efficient effort allocation.

The fact that the fixed wage performs well in providing efficient incentives in only one of three possible constellations, while it can even lead to a totally inefficient effort allocation in another case, might give the impression that this wage-setting scheme is in general inferior to the others. It should be noticed, however, that this

impression *can* be deceptive. As we have shown, just because the effort allocation is very sensitive to the contract wage, the preferences of the firm and the worker may coincide, with both parties favouring contract wages that are sufficiently high to make only the firm's outside option bind. If an increase in the contract wage induces the worker to switch her effort allocation, the firm's labour cost goes up, but on the other hand, it benefits from a higher marginal product. At the end, the firm might see its profits increase. Therefore, high contract wages that provide efficient incentives might be more likely, than we have predicted in the previous chapter.²³

However, if this is true, another problem about the interpretation of wage gaps is raised. Already in the previous chapter 6 we have seen that, depending on the size of the contract wage, both positive (if the worker's outside option is binding) and negative gaps (if the firm's outside option is binding) can occur. In that chapter, we argued that positive gaps might be more likely, because first, firm and worker have conflicting preferences about the size of the contract wage, and second, the worker's outside option might increase over time. If however, as we predict in this chapter, also the firm can have an interest in raising the contract wage, negative gaps would become more likely. Moreover, the interpretation of positive wage gaps would become most peculiar: the results of this chapter imply that positive wage gaps will only occur in a situation where the contract wage is low and workers allocate all their effort outside the firm. While the efficiency-wage hypothesis of wage gaps claims that firms raise the effective wage over the contract wage in order to motivate their workers, positive wage gaps would instead be a result of the workers' frustration about low contract wages.

Section 7.5 finally examines the incentive effects of local wage bargaining à la Holden (1988, 1989). Same as in the Nash-bargaining scenario considered in section 7.3, firm and worker negotiate about the division of the worker's marginal product. Yet this time, the threat points are not determined by the parties' outside options, but by the contract wage. The resulting effective wage is equal to the contract wage plus a mark-up (a wage gap), which is positively related to the worker's marginal product and her bargaining power. Because the bargained wage does not depend on the worker's outside option, it only provides an incentive to allocate effort in the firm, but no incentive to allocate effort in the outside option. Unlike Holden, however, we make a provision for binding outside options and allow for necessary adjustments of the locally bargained wage. Due to this possibility, the worker can be induced to allocate some effort outside the firm, if the locally bargained wage is low enough to make the worker's outside option bind. Yet, the incentive to allocate effort outside the firm is weaker than in the comparable situation in the fixed-wage scenario, because the worker's expected payoff still depends on her share in her marginal product, which she receives via the bargained wage.

The local wage-bargaining scenario also leads to different predictions about the occurrence and the characteristics of wage gaps. Because the mark-up on the con-

²³In fact, the argument here is the same as for efficiency wages. The employer announces *ex ante* to pay a fixed wage greater than the worker's outside option. The worker then exerts herself in the job in order to avoid that her high earnings are reduced *ex post*.

tract wage is just the object of local wage bargaining, positive wage gaps are inherent in this wage-setting scheme. Negative gaps can only occur if already the contract wage alone makes the firm's outside option bind. Moreover, different from the fixed-wage scenario, positive wage gaps are not related to an inefficient effort allocation, because the bargained wage gap just provides *the* incentive to the worker to allocate effort in the firm.

The discussion has shown, that out of the four wage-setting schemes that have been analysed in this chapter, only the fixed-wage and the local wage-bargaining scenario can lead to an efficient effort allocation without requiring, that the worker becomes *ex ante* residual claimant over the surplus generated in the employment relation. However, in order to provide positive incentives to the worker, also the fixed-wage scenario requires a relatively high wage that *can* make the firm's outside option bind, such that the worker still can become residual claimant *ex post*. Local wage bargaining, in contrast, leads to an efficient effort allocation at any but the lowest levels of the effective wage. In fact, the local wage-bargaining scenario owes this favourable characteristic to the structural similarity of its effective wage to the classical incentive scheme of principal-agency models: the wage of the worker depends on a performance-dependent component (the wage gap), which makes the worker put forth the desired effort in the job, and on a lump-sum component (the contract wage), which distributes the surplus between firm and worker. This similarity suggests that the firm should in principle not be concerned about local wage bargaining. Although employees (or a local union) force local wage bargaining on the firm by the threat that they will otherwise reduce their effort, the resulting wage gap just leads to an efficient effort allocation, which will *ceteris paribus* also benefit the firm. Instead, the firm should mostly be concerned about the size of the contract wage. This might explain why the public discussion about wages mainly focuses on collective bargaining.

Scheme	Effective Wage	Efficient Separation	Efficient Effort Allocation
Nash bargaining over the effective wage (section 7.3.1)	The wage is a weighed average of the worker's marginal product minus the firm's outside option, and the worker's outside options as perceived by the firm. Higher bargaining power of the firm results in a lower wage.	The firm can perceive the worker's outside option correctly: the wage will never make the worker's outside option bind.	The amount of effort allocated in the firm depends negatively on the firm's bargaining power. Only if the worker has full bargaining power, she will allocate all her effort in the firm. However, the firm has no interest in giving up some of its bargaining power to achieve a more efficient effort allocation.
Nash bargaining with a minimum wage (section 7.3.3)	The wage is the maximum of the Nash-bargaining outcome and an exogenously given wage rate (the <i>minimum wage</i>)	The firm can perceive the worker's outside option correctly: the wage will never make the worker's outside option bind.	The minimum wage provides insurance to the worker against bad outcomes. The effort allocation still depends on the distribution of bargaining power, but the worker tends to spread her effort more evenly over its two uses. The firm has no interest in giving up some of its bargaining power.
Fixed wage and renegotiations (section 7.4)	The wage is equal to a predetermined wage. It is only altered if it makes one of the parties' outside options bind. In that case, the respective other party becomes residual claimant.	The firm can perceive the worker's outside option correctly: the wage will never make the worker's outside option bind.	If the fixed wage can make neither outside option bind, the worker's effort allocation is undetermined. If it can make the worker's outside option bind, she will allocate all her effort outside the firm. The worker will only allocate all her effort in the firm, if the fixed wage can make the firm's outside option bind. The firm may have an interest in a higher fixed wage to achieve this outcome.
Local wage bargaining (section 7.5)	The wage is equal to an exogenously given wage rate plus a bargained mark-up depending on the worker's marginal product. The effective wage may be adapted to the parties' outside options.	The firm can perceive the worker's outside option correctly: if the wage is adapted to the outside options if necessary, no inefficient separation will occur.	The worker allocates all her effort in the firm. Only in case that the bargained wage can make the worker's outside option bind, she will allocate some (but not necessarily all her) effort in her outside option.

Table 7.1: Characteristics of the different wage-setting schemes

7.A APPENDIX

7.A.1 *The expected outside option and marginal product*

The worker's expectation about her outside option, contingent on her effort allocation e , is

$$E[a(s, e) | e] = \int_{s_0}^{s_1} a(s, e) h(s|e) ds = a(s_1, e) - \int_{s_0}^{s_1} \frac{\partial a(s, e)}{\partial s} H(s|e) ds. \quad (7.21)$$

Its derivatives with respect to the effort allocated in the firm are

$$\begin{aligned} \frac{dE[a(s, e) | e]}{de} &= \int_{s_0}^{s_1} \frac{\partial a(s, e)}{\partial e} h(s|e) ds + a(s, e) \frac{\partial h(s|e)}{\partial e} ds \\ &= \int_{s_0}^{s_1} \frac{\partial a(s, e)}{\partial e} h(s|e) ds - \int_{s_0}^{s_1} \frac{\partial a(s, e)}{\partial s} \frac{\partial H(s|e)}{\partial e} ds < 0, \end{aligned} \quad (7.22)$$

$$\begin{aligned} \frac{d^2E[a(s, e) | e]}{de^2} &= \int_{s_0}^{s_1} \frac{\partial^2 a(s, e)}{\partial e^2} h(s|e) ds + 2 \int_{s_0}^{s_1} \frac{\partial a(s, e)}{\partial e} \frac{\partial h(s|e)}{\partial e} ds \\ &\quad + \int_{s_0}^{s_1} a(s, e) \frac{\partial^2 h(s|e)}{\partial e^2} ds \geq 0. \end{aligned} \quad (7.23)$$

(See assumptions (7.2), (7.3) and (7.5).)

The firm's expectation about the worker's outside option is similar, but the true effort allocation e has to be replaced by the firm's belief η .

The expected value of the worker's outside option as perceived by the firm, but contingent on the true effort allocation, is

$$\begin{aligned} E[a(s, \eta) | e] &= \int_{s_0}^{s_1} a(s, \eta) h(s|e) ds \\ &= a(s_1, \eta) - \int_{s_0}^{s_1} \frac{\partial a(s, \eta)}{\partial s} H(s|e) ds. \end{aligned} \quad (7.24)$$

The signs of its derivatives with respect to the true effort allocation and the firm's belief are

$$\frac{\partial E[a(s, \eta) | e]}{\partial e} = \int_{s_0}^{s_1} a(s, \eta) \frac{\partial h(s|e)}{\partial e} ds = - \int_{s_0}^{s_1} \frac{\partial a(s, \eta)}{\partial s} \frac{\partial H(s|e)}{\partial e} ds \leq 0, \quad (7.25)$$

$$\frac{\partial^2 E[a(s, \eta) | e]}{\partial e^2} = \int_{s_0}^{s_1} a(s, \eta) \frac{\partial^2 h(s|e)}{\partial e^2} ds = - \int_{s_0}^{s_1} \frac{\partial a(s, \eta)}{\partial s} \frac{\partial^2 H(s|e)}{\partial e^2} ds \leq 0, \quad (7.26)$$

$$\frac{\partial E[a(s, \eta) | e]}{\partial \eta} = \int_{s_0}^{s_1} \frac{\partial a(s, \eta)}{\partial \eta} h(s|e) ds < 0, \quad (7.27)$$

$$\frac{\partial E[a(s, \eta) | e]}{\partial \eta \partial e} = \int_{s_0}^{s_1} \frac{\partial a(s, \eta)}{\partial \eta} \frac{\partial h(s|e)}{\partial e} ds = - \int_{s_0}^{s_1} \frac{\partial^2 a(s, \eta)}{\partial s \partial \eta} \frac{\partial H(s|e)}{\partial e} ds \geq 0. \quad (7.28)$$

(See assumptions (7.2), (7.3), (7.4) and (7.5).)

The expected value of the worker's marginal product is

$$E[m|e] = \int_{m_0}^{m_1} mg(m|e) dm = m_1 - \int_{m_0}^{m_1} G(m|e) dm. \quad (7.29)$$

The signs of its derivatives with respect to the effort allocation are

$$\frac{\partial E[m|e]}{\partial e} = - \int_{m_0}^{m_1} \frac{\partial G(m|e)}{\partial e} dm \geq 0, \quad (7.30)$$

$$\frac{\partial^2 E[m|e]}{\partial e^2} = - \int_{m_0}^{m_1} \frac{\partial^2 G(m|e)}{\partial e^2} dm \leq 0. \quad (7.31)$$

(See assumption (7.1).)

7.A.2 General characteristics of the expected payoff

In this section we derive the expected payoff of the worker and its relevant derivatives in case the worker's real or perceived outside option can become binding. We do this in rather general terms in order to be able to apply the results to the different wage setting schemes, especially Nash bargaining with binding outside options (section 7.3.2), fixed wage with renegotiations (section 7.4), and local wage bargaining (section 7.5). The scenario of Nash bargaining with a minimum wage (section 7.3.3) has a different constellation of parameters. (See also appendix 7.A.4.)

The expected bargaining wage

In most general terms the *ex post* bargaining wage w can be written as a function of the worker's marginal product m , the signal s , and a vector of parameters ξ , which—depending on the considered wage-setting scheme—can include the firm's bargaining power, its belief, or a minimum wage:

$$w := w(m, s, \xi), \quad \frac{\partial w}{\partial m} > 0, \quad \frac{\partial w}{\partial s} \geq 0$$

Consider the expected bargained wage,

$$E[w|e] = \int_{m_0}^{m_1} \int_{s_0}^{s_1} wh(s|e)g(m|e) ds dm.$$

It can be developed by repeatedly integrating by parts. The inner integral is calculated as

$$\begin{aligned} \int_{s_0}^{s_1} wh(s|e) ds &= w(s_1)H(s_1|e) - w(s_0)H(s_0|e) - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds \\ &= w(s_1) - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds, \end{aligned}$$

for $H(s_0|e) = 0$, and $H(s_1|e) = 1$. The expected bargained wage can now be written as the sum of two integrals,

$$E[w|e] = \int_{m_0}^{m_1} w(s_1)g(m|e) dm - \int_{m_0}^{m_1} \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e)g(m|e) ds dm.$$

The first summand is calculated as

$$\begin{aligned} \int_{m_0}^{m_1} w(s_1)g(m|e) dm &= w(m_1, s_1)G(m_1|e) - w(m_0, s_1)G(m_0|e) \\ &\quad - \int_{m_0}^{m_1} \frac{\partial w(s_1)}{\partial m} G(m|e) dm \\ &= w(m_1, s_1) - \int_{m_0}^{m_1} \frac{\partial w(s_1)}{\partial m} G(m|e) dm, \end{aligned}$$

because $G(m_0|e) = 0$, and $G(m_1|e) = 1$. The second summand is also integrated by parts, with $\int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds$ being one part and $g(m|e)$ the other. Then,

$$\begin{aligned} & \int_{m_0}^{m_1} \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) g(m|e) ds dm \\ &= G(m_1|e) \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds - G(m_0|e) \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds \\ & \quad - \int_{m_0}^{m_1} \int_{s_0}^{s_1} \frac{\partial^2 w}{\partial s \partial m} H(s|e) G(m|e) ds dm \\ &= \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds - \int_{m_0}^{m_1} \int_{s_0}^{s_1} \frac{\partial^2 w}{\partial s \partial m} H(s|e) G(m|e) ds dm. \end{aligned}$$

As all wage-setting schemes considered in this chapter are linear in the marginal product and the (perceived) outside option, we have $\frac{\partial w(s_1)}{\partial m} = \frac{\partial w}{\partial m}$ and $\frac{\partial^2 w}{\partial s \partial m} = 0$. Adding up the two summands then yields

$$E[w|e] = w(m_1, s_1) - \int_{m_0}^{m_1} \frac{\partial w}{\partial m} G(m|e) dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} H(s|e) ds.$$

Its derivatives with respect to the amount of effort allocated in the firm are equal to

$$\begin{aligned} \frac{\partial E[w|e]}{\partial e} &= - \int_{m_0}^{m_1} \frac{\partial w}{\partial m} \frac{\partial G(m|e)}{\partial e} dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} \frac{\partial H(s|e)}{\partial e} ds, & (7.32) \\ \frac{\partial^2 E[w|e]}{\partial e^2} &= - \int_{m_0}^{m_1} \frac{\partial w}{\partial m} \frac{\partial^2 G(m|e)}{\partial e^2} dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} \frac{\partial^2 H(s|e)}{\partial e^2} ds < 0, \\ \frac{\partial^2 E[w|e]}{\partial e \partial \xi} &= - \int_{m_0}^{m_1} \frac{\partial^2 w}{\partial m \partial \xi} \frac{\partial G(m|e)}{\partial e} dm - \int_{s_0}^{s_1} \frac{\partial^2 w}{\partial s \partial \xi} \frac{\partial H(s|e)}{\partial e} ds. \end{aligned}$$

The sign of the second derivative is due to the assumptions about the distributions of m and s (equations (7.1) and (7.2)).

The expected additional payoff

The *ex post* additional payoff over the wage is equal to the difference between (perceived) outside option and the bargained wage. If the firm perceives the worker's outside option correctly, or if it overestimates it, the additional payoff is equal to

$$a(s, \eta) - w(m, s, \xi),$$

while it is equal to

$$a(s, e) - w(m, s, \xi)$$

if the firm underestimates the worker's outside option. In order to capture both cases, we denote the *ex post* additional payoff by

$$f := f(m, s, e, \xi), \quad \frac{\partial f}{\partial m} < 0, \quad \frac{\partial f}{\partial s} \geq 0, \quad \frac{\partial f}{\partial e} \leq 0.$$

If, given the vector of parameters ξ , there is a positive probability that the worker's (perceived) outside option can become binding, there are critical levels of the signal and

the marginal product, denoted s^* and m^* , such that

$$\begin{aligned} f(m, s^*, e, \xi) &= 0, \\ f(m^*, s_1, e, \xi) &= 0, \end{aligned}$$

with

$$\begin{aligned} \frac{ds^*}{dm} &= - \frac{\partial f / \partial m}{\partial f / \partial s} \Big|_{s^*} = \frac{\partial w / \partial m}{\partial f / \partial s} \Big|_{s^*} > 0, & \frac{ds^*}{de} &= - \frac{\partial f / \partial e}{\partial f / \partial s} \Big|_{s^*} = - \frac{\partial a / \partial e}{\partial f / \partial s} \Big|_{s^*} \geq 0, \\ \frac{ds^*}{d\xi} &= - \frac{\partial f / \partial \xi}{\partial f / \partial s} \Big|_{s^*} = \frac{\partial w / \partial \xi}{\partial f / \partial s} \Big|_{s^*}, \\ \frac{dm^*}{de} &= - \frac{\partial f / \partial e}{\partial f / \partial m} \Big|_{m^*} = \frac{\partial a / \partial e}{\partial w / \partial m} \Big|_{m^*} \leq 0, & \frac{dm^*}{d\xi} &= - \frac{\partial f / \partial \xi}{\partial f / \partial m} \Big|_{m^*} = - \frac{\partial w / \partial \xi}{\partial w / \partial m} \Big|_{m^*}. \end{aligned}$$

The worker's outside option is binding for signals $s > s^*$ and marginal products $m < m^*$.

After the firm has announced its belief, the worker chooses the effort allocation that maximizes her expected payoff. Because she receives her (perceived) outside option if $s > s^*$ and $m < m^*$, her objective is

$$\begin{aligned} E[u|e] &= \int_{m_0}^{m^*} \int_{s_0}^{s^*} wh(s|e)g(m|e) ds dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} ah(s|e)g(m|e) ds dm \\ &\quad + \int_{m^*}^{m_1} \int_{s_0}^{s_1} wh(s|e)g(m|e) ds dm \\ &= \int_{m_0}^{m_1} \int_{s_0}^{s_1} wh(s|e)g(m|e) ds dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} (a-w)h(s|e)g(m|e) ds dm \\ &= \int_{m_0}^{m_1} \int_{s_0}^{s_1} wh(s|e)g(m|e) ds dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} fh(s|e)g(m|e) ds dm \\ &= E[w|e] + A, \end{aligned}$$

where a is either equal to $a(s, \eta)$ or equal to $a(s, e)$, depending on the case under consideration.

Consider the second summand A of the expected payoff. In the following, we call

$$A = \int_{m_0}^{m^*} \int_{s^*}^{s_1} fh(s|e)g(m|e) ds dm \quad (7.33)$$

the expected additional payoff. Its term can also be developed by repeatedly integrating by parts. Applying this rule for the inner integral we get

$$\begin{aligned} \int_{s^*}^{s_1} fh(s|e) ds &= f(s_1)H(s_1|e) - f(s^*)H(s^*|e) - \int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds \\ &= f(s_1)H(s_1|e) - \int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds, \end{aligned}$$

because by definition, $f(s^*) = 0$ at all $m \in [m_0, m^*]$.

The expected additional payoff can now be written as

$$A = H(s_1|e) \int_{m_0}^{m^*} f(s_1)g(m|e) dm - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e)g(m|e) ds dm.$$

We integrate the two summands by parts. The first becomes

$$\begin{aligned} & H(s_1|e) \int_{m_0}^{m^*} f(s_1) g(m|e) dm \\ &= H(s_1|e) [f(m_1, s^*) G(m^*|e) - f(m_0, s_1) G(m_0|e)] \\ &\quad - H(s_1|e) \int_{m_0}^{m^*} \frac{\partial f(s_1)}{\partial m} G(m|e) dm \\ &= -f(m_0, s_1) H(s_1|e) G(m_0|e) - H(s_1|e) \int_{m_0}^{m^*} \frac{\partial f(s_1)}{\partial m} G(m|e) dm, \end{aligned}$$

because by definition, $f(m^*, s_1) = 0$. In order to develop the second summand, we take the inner integral $\int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds$ as one part, and $g(m|e)$ as the other. Notice that $s^*(m^*) = s_1$, $\frac{ds^*}{dm} = -\frac{\partial f/\partial m}{\partial f/\partial s} \Big|_{s^*} = \frac{\partial w/\partial m}{\partial f/\partial s} \Big|_{s^*}$, and that $\frac{d}{dm} \int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds = -\frac{ds^*}{dm} \frac{\partial f}{\partial s} \Big|_{s^*} H(s^*|e) + \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial m} H(s|e) ds$. Then,

$$\begin{aligned} & \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial s} H(s|e) g(m|e) ds dm \\ &= G(m^*|e) \int_{s^*(m^*)}^{s_1} \frac{\partial f(m^*)}{\partial s} H(s|e) ds - G(m_0|e) \int_{s^*(m_0)}^{s_1} \frac{\partial f(m_0)}{\partial s} H(s|e) ds \\ &\quad - \int_{m_0}^{m^*} \left[-\frac{ds^*}{dm} \frac{\partial f}{\partial s} \Big|_{s^*} H(s^*|e) + \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial m} H(s|e) \right] G(m|e) dm \\ &= -G(m_0|e) \int_{s^*(m_0)}^{s_1} \frac{\partial f(m_0)}{\partial s} H(s|e) ds - \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial m} H(s^*|e) G(m|e) dm \\ &\quad - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial m} H(s|e) G(m|e) dm. \end{aligned}$$

Adding up the two summands yields

$$\begin{aligned} A &= -f(m_0, s_1) H(s_1|e) G(m_0|e) + G(m_0|e) \int_{s^*(m_0)}^{s_1} \frac{\partial f(m_0)}{\partial s} H(s|e) ds \\ &\quad + \int_{m_0}^{m^*} \left[\frac{\partial f(s^*)}{\partial m} H(s^*|e) - \frac{\partial f(s_1)}{\partial m} H(s_1|e) \right] G(m|e) dm \\ &\quad + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial m} H(s|e) G(m|e) dm. \end{aligned}$$

This term can be further simplified by using the fact that all wage-setting schemes under consideration imply a bargained wage that is linear in the marginal product and the outside option as perceived by the firm. Consequently, $\frac{\partial f(s_1)}{\partial m} = \frac{\partial f(s^*)}{\partial m} = \frac{\partial f}{\partial m}$, $\frac{\partial f(m_0)}{\partial s} = \frac{\partial f}{\partial s}$, and $\frac{\partial^2 f}{\partial s \partial m} = 0$. Then,

$$\begin{aligned} A &= -f(m_0, s_1) H(s_1|e) G(m_0|e) + G(m_0|e) \int_{s^*(m_0)}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds \\ &\quad + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [H(s^*|e) - H(s_1|e)] G(m|e) dm. \end{aligned} \quad (7.34)$$

Moreover, $H(s_0|e) = G(m_0|e) = 0$, and $H(s_1|e) = G(m_1|e) = 1$. Hence,

$$A = - \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [1 - H(s^*|e)] G(m|e) dm. \quad (7.35)$$

In order to calculate the first derivative of the expected additional payoff with respect to the amount of effort allocated in the firm, we must take into account that the boundaries of the integral are functions of e if the real outside option can become binding. In general, we get

$$\begin{aligned} \frac{\partial A}{\partial e} &= \frac{\partial}{\partial e} \int_{m_0}^{m^*} \int_{s^*}^{s_1} fh(s|e)g(m|e) ds dm \\ &= \frac{dm^*}{de} \int_{s^*(m^*)}^{s_1} f(m^*)h(s|e)g(m^*|e) ds \\ &\quad + \int_{m_0}^{m^*} -\frac{ds^*}{de} f(s^*)h(s^*|e)g(m|e) dm \\ &\quad + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \left[\frac{\partial f}{\partial e} h(s|e)g(m|e) + f \frac{\partial h(s|e)}{\partial e} g(m|e) + fh(s|e) \frac{\partial g(m|e)}{\partial e} \right] ds dm. \end{aligned}$$

Because $s^*(m^*) = s_1$ and $f(s^*) = 0$, the first two summands are equal to zero. Consequently,

$$\begin{aligned} \frac{\partial A}{\partial e} &= \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e)g(m|e) ds dm \\ &\quad + \int_{m_0}^{m^*} \int_{s^*}^{s_1} f \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm. \end{aligned}$$

The first summand is nonpositive, because $\frac{\partial f}{\partial e} \leq 0$. If the firm does not underestimate the worker's outside option it even vanishes because $\frac{\partial f}{\partial e} = 0$. The second summand of $\frac{\partial A}{\partial e}$ is similar to the term of A in equation (7.33), with either $h(s|e)$ or $g(m|e)$ being replaced by their derivative with respect to e . Using equation (7.34), the second summand of $\frac{\partial A}{\partial e}$ can be written as

$$\begin{aligned} &\int_{m_0}^{m^*} \int_{s^*}^{s_1} f \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\ &= \int_{m_0}^{m^*} \int_{s^*}^{s_1} f \frac{\partial h(s|e)}{\partial e} g(m|e) ds dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} fh(s|e) \frac{\partial g(m|e)}{\partial e} ds dm \\ &= -f(m_0, s_1) \frac{\partial H(s_1|e)}{\partial e} G(m_0|e) - f(m_0, s_1) H(s_1|e) \frac{\partial G(m_0|e)}{\partial e} \\ &\quad + G(m_0|e) \int_{s^*(m_0)}^{s_1} \frac{\partial f}{\partial s} \frac{\partial H(s|e)}{\partial e} ds + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} \left[\frac{\partial H(s^*|e)}{\partial e} - \frac{\partial H(s_1|e)}{\partial e} \right] G(m|e) dm \\ &\quad + \frac{\partial G(m_0|e)}{\partial e} \int_{s^*(m_0)}^{s_1} \frac{\partial f}{\partial s} H(s|e) ds + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [H(s^*|e) - H(s_1|e)] \frac{\partial G(m|e)}{\partial e} dm. \end{aligned}$$

Since $H(s_0|e) = G(m_0|e) = 0$, and $H(s_1|e) = G(m_1|e) = 1$, irrespective of the value of e , their derivatives are $\frac{\partial H(s_1|e)}{\partial e} = \frac{\partial G(m_0|e)}{\partial e} = 0$ for all e . Therefore, the second summand can

be further simplified to

$$\begin{aligned} & \int_{m_0}^{m^*} \int_{s^*}^{s_1} f \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\ &= \int_{m_0}^{m^*} \frac{\partial f}{\partial m} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [H(s^*|e) - 1] \frac{\partial G(m|e)}{\partial e} dm, \end{aligned}$$

and the derivative of A becomes

$$\begin{aligned} \frac{\partial A}{\partial e} &= \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm \\ &+ \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [H(s^*|e) - 1] \frac{\partial G(m|e)}{\partial e} dm < 0. \end{aligned} \quad (7.36)$$

It is strictly negative, because $\frac{\partial f}{\partial e} \leq 0$, $\frac{\partial f}{\partial m} < 0$, $\frac{\partial H(s^*|e)}{\partial e} > 0$, $\frac{\partial G(m|e)}{\partial e} \leq 0$, and $H(s^*|e) < 1$.

Unfortunately, the sign of the second derivative of A cannot be determined without specifying the distribution functions.

The cross derivative with respect to a parameter ξ is given by

$$\begin{aligned} \frac{\partial^2 A}{\partial e \partial \xi} &= \frac{\partial}{\partial \xi} \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm \\ &+ \frac{\partial}{\partial \xi} \int_{m_0}^{m^*} \int_{s^*}^{s_1} f \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm. \end{aligned}$$

Consider the first summand. It is

$$\begin{aligned} & \frac{\partial}{\partial \xi} \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm \\ &= \frac{dm^*}{d\xi} \int_{s^*(m^*)}^{s_1} \frac{\partial f(m^*)}{\partial e} h(s|e) g(m^*|e) ds \\ &+ \int_{m_0}^{m^*} -\frac{ds^*}{d\xi} \frac{\partial f(s^*)}{\partial e} h(s^*|e) g(m|e) dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial e \partial \xi} h(s|e) g(m|e) ds dm \\ &= - \int_{m_0}^{m^*} \frac{ds^*}{d\xi} \frac{\partial f(s^*)}{\partial e} h(s^*|e) g(m|e) dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial e \partial \xi} h(s|e) g(m|e) ds dm. \end{aligned}$$

The second summand is calculated as

$$\begin{aligned} & \frac{\partial}{\partial \xi} \int_{m_0}^{m^*} \int_{s^*}^{s_1} f \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\ &= \frac{dm^*}{d\xi} \int_{s^*(m^*)}^{s_1} f(m^*) \left(\frac{\partial h(s|e)}{\partial e} g(m^*|e) + h(s|e) \frac{\partial g(m^*|e)}{\partial e} \right) ds \\ &+ \int_{m_0}^{m^*} -\frac{ds^*}{d\xi} f(s^*) \left(\frac{\partial h(s^*|e)}{\partial e} g(m|e) + h(s^*|e) \frac{\partial g(m|e)}{\partial e} \right) dm \\ &+ \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial \xi} \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\ &= \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial \xi} \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm. \end{aligned}$$

Repeatedly integrating by parts this simplifies to

$$\begin{aligned}
 & \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial \xi} \left(\frac{\partial h(s|e)}{\partial e} g(m|e) + h(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\
 = & \int_{m_0}^{m^*} \left[\frac{\partial f(s_1)}{\partial \xi} \left(\frac{\partial H(s_1|e)}{\partial e} g(m|e) + H(s_1|e) \frac{\partial g(m|e)}{\partial e} \right) \right] dm \\
 & - \int_{m_0}^{m^*} \left[\frac{\partial f(s^*)}{\partial \xi} \left(\frac{\partial H(s^*|e)}{\partial e} g(m|e) + H(s^*|e) \frac{\partial g(m|e)}{\partial e} \right) \right] dm \\
 & - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial \xi} \left(\frac{\partial H(s|e)}{\partial e} g(m|e) + H(s|e) \frac{\partial g(m|e)}{\partial e} \right) ds dm \\
 = & - \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial \xi} \frac{\partial H(s^*|e)}{\partial e} g(m|e) dm - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial \xi} \frac{\partial H(s|e)}{\partial e} g(m|e) ds dm \\
 & + \int_{m_0}^{m^*} \frac{\partial f(s_1)}{\partial \xi} \frac{\partial g(m|e)}{\partial e} dm - \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial \xi} H(s^*|e) \frac{\partial g(m|e)}{\partial e} dm \\
 & - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial \xi} H(s|e) \frac{\partial g(m|e)}{\partial e} ds dm,
 \end{aligned}$$

with

$$\begin{aligned}
 \int_{m_0}^{m^*} \frac{\partial f(s_1)}{\partial \xi} \frac{\partial g(m|e)}{\partial e} dm &= \frac{\partial f(m^*, s_1)}{\partial \xi} \frac{\partial G(m^*|e)}{\partial e} - \frac{\partial f(m_0, s_1)}{\partial \xi} \frac{\partial G(m_0|e)}{\partial e} \\
 &= \frac{\partial f(m^*, s_1)}{\partial \xi} \frac{\partial G(m^*|e)}{\partial e},
 \end{aligned}$$

and

$$\begin{aligned}
 \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial \xi} H(s^*|e) \frac{\partial g(m|e)}{\partial e} dm \\
 &= \frac{\partial f(m^*, s_1)}{\partial \xi} \frac{\partial G(m^*|e)}{\partial e} \\
 &\quad - \int_{m_0}^{m^*} \left(\frac{\partial^2 f(s^*)}{\partial s^* \partial \xi} H(s^*|e) + \frac{\partial f(s^*)}{\partial \xi} h(s^*|e) \right) \frac{ds^*}{dm} \frac{\partial G(m|e)}{\partial e} dm,
 \end{aligned}$$

and

$$\begin{aligned}
 & \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial \xi} H(s|e) \frac{\partial g(m|e)}{\partial e} ds dm \\
 = & \int_{s^*(m^*)}^{s_1} \frac{\partial^2 f(m^*)}{\partial s \partial \xi} H(s|e) \frac{\partial G(m^*|e)}{\partial e} ds - \int_{s^*(m_0)}^{s_1} \frac{\partial^2 f(m_0)}{\partial s \partial \xi} H(s|e) \frac{\partial G(m_0|e)}{\partial e} ds \\
 & - \int_{m_0}^{m^*} \frac{ds^*}{dm} \frac{\partial^2 f(s^*)}{\partial s^* \partial \xi} H(s^*|e) \frac{\partial G(m|e)}{\partial e} dm \\
 = & \int_{m_0}^{m^*} \frac{ds^*}{dm} \frac{\partial^2 f(s^*)}{\partial s^* \partial \xi} H(s^*|e) \frac{\partial G(m|e)}{\partial e} dm.
 \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\partial^2 A}{\partial e \partial \xi} = & - \int_{m_0}^{m^*} \frac{ds^*}{d\xi} \frac{\partial f(s^*)}{\partial e} h(s^*|e) g(m|e) dm + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial e \partial \xi} h(s|e) g(m|e) ds dm \\ & - \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial \xi} \frac{\partial H(s^*|e)}{\partial e} g(m|e) dm - \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial^2 f}{\partial s \partial \xi} \frac{\partial H(s|e)}{\partial e} g(m|e) ds dm \\ & + \int_{m_0}^{m^*} \frac{\partial f(s^*)}{\partial \xi} h(s^*|e) \frac{ds^*}{dm} \frac{\partial G(m|e)}{\partial e} dm. \end{aligned} \quad (7.37)$$

The optimal effort allocation if the worker's outside option can be binding

We are now ready to derive some general results about the effort allocation if the real or perceived outside option can become binding. Using the equations (7.32) and (7.36), we can write the first derivative of the worker's expected payoff with respect to the amount of effort allocated in the firm as

$$\begin{aligned} \frac{\partial E[u|e]}{\partial e} = & \frac{\partial E[w|e]}{\partial e} + \frac{\partial A}{\partial e} \\ = & - \int_{m_0}^{m_1} \frac{\partial w}{\partial m} \frac{\partial G(m|e)}{\partial e} dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} \frac{\partial H(s|e)}{\partial e} ds \\ & + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm \\ & + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm + \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [H(s^*|e) - 1] \frac{\partial G(m|e)}{\partial e} dm. \end{aligned}$$

After simplifying and arranging the terms, this becomes

$$\begin{aligned} \frac{\partial E[u|e]}{\partial e} = & - \int_{m_0}^{m^*} \frac{\partial w}{\partial m} H(s^*|e) \frac{\partial G(m|e)}{\partial e} dm - \int_{m^*}^{m_1} \frac{\partial w}{\partial m} \frac{\partial G(m|e)}{\partial e} dm \\ & - \int_{m_0}^{m^*} \frac{\partial w}{\partial m} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} \frac{\partial H(s|e)}{\partial e} ds \\ & + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm. \end{aligned}$$

In general, its sign cannot be determined. However, in the vicinity of the two extreme effort allocations we get

$$\begin{aligned} \lim_{e \rightarrow 0} \frac{\partial E u}{\partial e} = & - \int_{m_0}^{m^*} \frac{\partial w}{\partial m} H(s^*|e) \frac{\partial G(m|e)}{\partial e} dm - \int_{m^*}^{m_1} \frac{\partial w}{\partial m} \frac{\partial G(m|e)}{\partial e} dm \\ & + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm \geq 0, \\ \lim_{e \rightarrow 1} \frac{\partial E u}{\partial e} = & - \int_{m_0}^{m^*} \frac{\partial w}{\partial m} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm - \int_{s_0}^{s_1} \frac{\partial w}{\partial s} \frac{\partial H(s|e)}{\partial e} ds \\ & + \int_{m_0}^{m^*} \int_{s^*}^{s_1} \frac{\partial f}{\partial e} h(s|e) g(m|e) ds dm < 0. \end{aligned} \quad (7.38)$$

This implies that the worker will never allocate all her effort in the firm if either her real outside option or her outside option as perceived by the firm can become binding.

7.A.3 The expected additional payoff in the Nash bargaining scenario

The worker will quit if the bargained wage is less than her outside option, $w_N < a(s, e)$. This situation can only occur if the firm overestimates the amount of effort allocated in the firm, $\eta > e$, and if the worker has not the full bargaining power.

In order to identify the values of the signal and the marginal product for which the worker's outside option can be binding, notice first that both the wage and the outside option are functions of the signal. We define a critical signal $s_N^*(m, \theta, e, \eta)$, such that

$$a(s_N^*, e) = w_N(s_N^*, m, \theta) = \theta a(s_N^*, \eta) + (1 - \theta)(m - a_F). \quad (7.39)$$

Because of assumption (7.4), the outside option is binding for all signals $s > s_N^*(m, \theta)$, if such a signal exists in $[s_0, s_1]$. Differentiating equation (7.39) shows that the critical signal is an increasing function in the marginal product m and the amount of effort allocated in the firm e , and a decreasing function of the firm's bargaining power θ and its belief η :

$$\begin{aligned} \frac{ds_N^*}{dm} &= \frac{1 - \theta}{\left. \frac{\partial a(s, e)}{\partial s} \right|_{s_N^*} - \theta \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_N^*}} \geq 0, & \frac{ds_N^*}{d\theta} &= -(m - a_F - a(s_N^*, \eta)) \leq 0, \\ \frac{ds_N^*}{de} &= -\frac{\left. \frac{\partial a(s_N^*, e)}{\partial e} \right|_{s_N^*}}{\left. \frac{\partial a(s, e)}{\partial s} \right|_{s_N^*} - \theta \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_N^*}} > 0, & \frac{ds_N^*}{d\eta} &= \frac{\theta \left. \frac{\partial a(s_N^*, \eta)}{\partial \eta} \right|_{s_N^*}}{\left. \frac{\partial a(s, e)}{\partial s} \right|_{s_N^*} - \theta \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_N^*}} \leq 0 \end{aligned}$$

In a similar way we define a critical marginal product $m_N^*(\theta, e, \eta)$:

$$a(s_1, e) = \theta a(s_1, \eta) + (1 - \theta)(m_N^* - a_F)$$

If such an $m_N^* \in [m_0, m_1]$ exists, then for all $m > m_N^*$ there are no signals $s \in [s_0, s_1]$ that make the outside option bind. The critical marginal product is an increasing function of the firm's bargaining power θ and its belief η , and a decreasing function of the amount of effort allocated in the firm e :

$$\begin{aligned} \frac{dm_N^*}{d\theta} &= \frac{m_N^* - a_F - a(s_1, \eta)}{1 - \theta} > 0, \\ \frac{dm_N^*}{de} &= \frac{\left. \frac{\partial a(s_N^*, e)}{\partial e} \right|_{s_N^*}}{1 - \theta} < 0, & \frac{dm_N^*}{d\eta} &= -\frac{\theta \left. \frac{\partial a(s_N^*, \eta)}{\partial \eta} \right|_{s_N^*}}{1 - \theta} \geq 0 \end{aligned}$$

The opportunity to leave the firm in case her outside option is binding changes the expected payoff of the worker. While for $(m, s) \in \{[m_N^*, m_1], [s_0, s_N^*]\}$ she will accept the wage w_N , she will choose her outside options for other combinations of m and s .

Applying equation (7.33), we now can write the expected additional payoff as

$$\begin{aligned} A_N &= - \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [1 - H(s_N^*|e)] G(m|e) dm \\ &= (1 - \theta) \int_{m_0}^{m_N^*} [1 - H(s_N^*|e)] G(m|e) dm. \end{aligned}$$

By equation (7.36), its derivative with respect to the amount of effort allocated in the firm is

$$\begin{aligned} \frac{\partial A}{\partial e} &= \int_{m_0}^{m^*} \int_{s^*}^{s^1} \frac{\partial a(s, e)}{\partial e} h(s|e) g(m|e) ds dm \\ &\quad - (1 - \theta) \int_{m_0}^{m^*} \frac{\partial H(s^*|e)}{\partial e} G(m|e) dm \\ &\quad + (1 - \theta) \int_{m_0}^{m^*} \frac{\partial f}{\partial m} [1 - H(s^*|e)] \frac{\partial G(m|e)}{\partial e} dm < 0. \end{aligned}$$

7.A.4 The expected additional payoff in the minimum wage scenario

Differentiating equation (7.13), that defines the critical signal s_M^* at which the minimum wage is exactly equal to the bargained wage, yields

$$\frac{ds_M^*}{dm} = - \frac{1 - \theta}{\theta \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_M^*}} \leq 0.$$

The expected additional payoff A_M from equation (7.14) is different from the previously analysed additional payoffs, because the minimum wage is binding for signals $s \in [s_0, s_M^*]$ and marginal products $m \in [m_0, m_M^*]$. Nevertheless, we can use the same integration rules as for the other A . Then,

$$\begin{aligned} A_M &= \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} [\bar{w} - w_N] h g ds dm \\ &= \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} [\bar{w} - \theta a(s, \eta) - (1 - \theta)(m - a_F)] h g ds dm \\ &= \int_{m_0}^{m_M^*} [\bar{w} - \theta a(s_M^*, \eta) - (1 - \theta)(m - a_F)] H(s_M^*|e) g dm \\ &\quad + \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} \theta \frac{\partial a(s, \eta)}{\partial s} H g ds dm \\ &= \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} \theta \frac{\partial a(s, \eta)}{\partial s} H g ds dm \\ &= G(m_M^*|e) \int_{s_0}^{s_M^*(m_M^*)} \theta \frac{\partial a(s, \eta)}{\partial s} H ds \\ &\quad - \theta \int_{m_0}^{m_M^*} \frac{ds_M^*}{dm} \left. \frac{\partial a(s, \eta)}{\partial s} \right|_{s_M^*} H(s_M^*|e) G dm \\ &= (1 - \theta) \int_{m_0}^{m_M^*} H(s_M^*|e) G(m|e) dm. \end{aligned}$$

Its derivative is

$$\begin{aligned}
 \frac{\partial A_M}{\partial e} &= \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} [\bar{w} - w_N] \left(\frac{\partial h}{\partial e} g + h \frac{\partial g}{\partial e} \right) ds dm \\
 &= \int_{m_0}^{m_M^*} [\bar{w} - \theta a(s_M^*, \eta) - (1 - \theta)(m - a_F)] \left(\frac{\partial H(s_M^*|e)}{\partial e} g + H(s_M^*|e) \frac{\partial g}{\partial e} \right) dm \\
 &\quad + \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} \theta \frac{\partial a(s, \eta)}{\partial s} \left(\frac{\partial H}{\partial e} g + H \frac{\partial g}{\partial e} \right) ds dm \\
 &= \int_{m_0}^{m_M^*} \int_{s_0}^{s_M^*} \theta \frac{\partial a(s, \eta)}{\partial s} \left(\frac{\partial H}{\partial e} g + H \frac{\partial g}{\partial e} \right) ds dm \\
 &= \theta \int_{s_0}^{s_M^*} \frac{\partial a(s, \eta)}{\partial s} \left(\frac{\partial H}{\partial e} G(m_M^*|e) + H \frac{\partial G(m_M^*|e)}{\partial e} \right) ds \\
 &\quad - \theta \int_{m_0}^{m_M^*} \frac{ds_M^*}{dm} \frac{\partial a(s_M^*, \eta)}{\partial s} \left(\frac{\partial H(s_M^*|e)}{\partial e} G + H(s_M^*|e) \frac{\partial G}{\partial e} \right) dm \\
 &= (1 - \theta) \int_{m_0}^{m_M^*} \left(\frac{\partial H(s_M^*|e)}{\partial e} G + H(s_M^*|e) \frac{\partial G}{\partial e} \right) dm. \tag{7.40}
 \end{aligned}$$

The sign of this derivative is in general ambiguous. However, by assumptions (7.1) and (7.2) we get the following signs for extreme effort allocations:

$$\begin{aligned}
 \lim_{e \rightarrow 0} \frac{\partial A_M}{\partial e} &= (1 - \theta) \int_{m_0}^{m_M^*} H(s_M^*|e) \frac{\partial G}{\partial e} dm < 0, \\
 \lim_{e \rightarrow 1} \frac{\partial A_M}{\partial e} &= (1 - \theta) \int_{m_0}^{m_M^*} \frac{\partial H(s_M^*|e)}{\partial e} G dm > 0
 \end{aligned}$$

CHAPTER 8

Conclusions

8.1 SUMMARY

The aim of this thesis was to provide a better understanding of the factors that lead to the occurrence and persistence of wage gaps. The discussion in part I (chapters 4 and 5) centred around two themes, efficiency wages and bargaining power. If firms pay efficiency wages, employees or local labour representatives that have some bargaining power vis-à-vis the firm, can try to influence the relation between wage and effort in different ways. First, they can accept the effort function that results from their incentive-compatibility constraint. Firm and employees will then negotiate about which point on the effort function to choose. As we have argued in chapter 5, negotiations about the effort level will result in a commitment problem, if effort cannot be monitored perfectly.¹ The same problem will arise if points beyond the effort functions are chosen by simultaneous bargaining about wage and effort levels. Therefore, the bargaining parties can choose a point on the effort function only by bargaining about the wage. We have discussed local wage bargaining in chapter 4 of this thesis. A second way to influence the wage-effort relation is to bargain about the parameters that determine the shape of the effort function. This type of bargaining was discussed in chapter 5.

In chapter 4, we focused on bargaining about the division of rents as a reason for wage gaps. After the contract wage has been determined by collective bargaining between union and employers' federation, employees (or a local union) can force another round of bargaining on individual firms. This kind of local wage bargaining is common in Scandinavian countries, where collective bargaining is highly centralized, but a similar bargaining situation can also arise in other countries, possibly in an implicit way. As chapter 3 has shown, the different studies on this kind of bargaining, Holden (1988, 1989, 1998), Muysken and van Veen (1996) and de Gijssel (1996), deal with the concept of bargaining power in a very stylized way. Moreover, the studies make different assumptions about the role of efficiency wages in the determination of the threat points of firm and workers. While Muysken and van Veen assume a single effort function to hold in situations of agreement as well as during bargaining, Holden and de Gijssel argue that workers are able to reduce their effort during bargaining, in order to make a disagreement costly for the firm. As turns out, the predictions

¹And that is just one of the reasons why firms pay efficiency wages.

about the occurrence of wage gaps are highly sensitive to the assumptions about the threat points.

Chapter 4 dealt with the issue of bargaining power in two ways. First, it modelled the process of bargaining explicitly, introducing a noncooperative bargaining game in which bargaining power is expressed by the frequency at which the bargaining parties can make wage proposals. Second, it considered the choice of the workers' threat point to be another expression of bargaining power. It was shown that different constellations of the contract wage, the profit-maximizing (efficiency) wage, and the worker's disagreement payoff lead to different predictions about the occurrence of wage gaps, and that the above-mentioned studies can all be drawn back to one of these constellations.

Chapter 5 addressed bargaining about the shape of the effort function. It started out from the observation that in Western European countries, unions and works councils have a say in the determination of working conditions and in the personnel policy of employers. They can, for example, use their say in these issues in order to influence the monitoring technology the firm applies, or in order to influence the possible sanctions a worker can expect when she is caught shirking. We argued that workers' representatives cannot protect workers from being laid off who fall short of any performance standard, but that it will back workers against the firm, as long as they put forth at least a certain *reference level of effort*. Because these "semi-shirkers" are protected from unemployment, the severest sanction the firm can impose is to cut their wage back to the contract-wage level.

The model we have analysed in chapter 5 provided some interesting insights in the behaviour of wage gaps and the effects of works councils (or unions in general) on productivity. First, by introducing a reference effort level that is tied to the contract wage, we provided an argument for why the firm's profit-maximizing (efficiency) wage may depend on the contract wage. We showed that a raise in the contract wage will lead to an even stronger increase in the effective wage, i.e., that the size of the wage gap depends positively on the contract wage. Second, we compared the impact of different types of unions on effort and employment. The result was that, if unions care about employment, they might induce an effective effort level that is higher than it would be without unions.

In part II (chapters 6 and 7) of this thesis we discussed possible motives of workers and firms to settle on wage-setting systems that imply the occurrence of wage gaps. Both chapters of this part stressed the importance of incentive issues for this decision. In chapter 6 we focused on the employment relation between a single employee and her employer. In this chapter, we argued that both parties are typically interested in a fixed, predetermined wage in order to avoid hold-up of investments. If a party can undertake specific investments that generate returns which directly accrue to the investing party, but which are of no value if the employment relation is abolished, the other party has an interest to change the distribution of surpluses after the investment is done, in order to acquire a share of the returns. If the investing party is not protected by a fixed wage from being ripped off in this way, the investment level will be inefficiently low. However, a fixed wage can also put the

investment and even the whole employment relation at risk, if it makes the outside option of one of the parties bind. Therefore, we argued, the parties must on the one hand determine a fixed (contract) wage, while on the other hand they must make a provision for renegotiations of their contract, if one of their outside options threatens to become binding. In chapter 6, a noncooperative renegotiation game was proposed, that leads to outcomes that carefully weigh the need to leave the contract wage unchanged against the need for wage flexibility: the wage is changed only if an outside option becomes binding, and it is only changed to the extent that is necessary to prevent the party with the binding outside option from abolishing the employment relation. In this view, wage gaps, deviations of the renegotiated wage from the predetermined contract wage, occur due to changes in the outside options, i.e., they are induced solely by factors from outside the employment relation. In spite of the underlying incentive mechanism, the chapter provided a similar view on wage gaps as the Phillips-curve approach.

In the model of chapter 6 wage gaps occurred as necessary adjustments of the wage to exogenous factors, while in principle both the worker and her employer had an interest in keeping the wage fixed, in order to protect the returns of their specific investments. Chapter 7, on the other hand, has emphasized the agency aspect of an employment relation. We assumed another type of investment, which is more typical of employment relations. The worker has to decide about the allocation of effort over two uses: she can put forth effort in the firm, or she can invest in her outside option. The chapter addressed the problem of finding an incentive scheme that leads to efficient investments in the firm, when firm and worker cannot contract on the effort allocation. We compared the performance of different wage-formation systems, especially a fixed-wage scheme similar to the one of chapter 6, and a scheme with local wage bargaining similar to the one of chapter 4. We concluded, that the latter is the most efficient system, because it combines a fixed base wage rate (the contract wage) with a performance dependent pay (the wage gap as share in the firm's revenues). It was argued that the contract wage is necessary to prevent the worker from investing in her outside option, but that it does not provide an incentive to invest in a greater work performance. This incentive can only be provided by the wage gap. The outcome thus supports the basic hypothesis of this thesis, that both bargaining and incentive pay are important determinants of the wage gap.

8.2 RECOMMENDATIONS FOR ECONOMIC POLICY

The recommendations for economic policy that can be drawn from this study can only be indirect. Perhaps the most important is that wage gaps are nothing policy makers should be worried about.

Wage gaps, or more precisely, wage drift used to be a public concern in the 1950s and 1960s, when the drift rates were much higher than nowadays. In those years, some Western European governments, including the Dutch,² interfered actively in the wage-formation process in order to keep wage increases under control—after the

²Recall the "guided wage policy" (chapter 2).

Second World War, most governments were more concerned about economic growth than about income policies. However, the governments' control mostly applied to collective agreements on wages only, and so the rapid increases of effective wages due to a rising demand for labour aroused their suspicion. Another concern, that was uttered in many contemporary studies on wage drift, was that unions and employer federations might lose their grip on the development of wages.³ This claim, however, can be rejected, because in all models discussed in this thesis, wage gaps are co-determined by the contract wage. Hence, it is likely, that unions and employers' federations will take them into account when negotiating collective agreements.

The discussion about wage drift has faded after the early 1970s, and with it, for about twenty years, the interest in the topic. This may be due to two reasons. First, the wage-drift rates were much lower, and even close to zero, after the mid-1970s.⁴ Second, the *Zeitgeist* had turned against government intervention, and nowadays, governments usually respect the autonomy of the parties involved in wage formation.

A closer look, however, reveals that the discussion has not faded but that only its emphasis has shifted. In view of high and persistent unemployment in European countries, the recent years have seen a vivid discussion about the performance of different wage-formation systems, initiated by the seminal contribution of Calmfors and Driffill (1988). They claimed that the performance of labour markets is better in countries where wage setting is either highly centralized or highly decentralized. Countries in between, like Germany and the Netherlands, where collective agreements are bargained at the level of industries, perform worse according to their study.⁵ Their reasoning is that if collective bargaining occurs on a centralized, say nationwide, level, the bargaining parties will have to take the employment effects of their decisions into account. If, in contrast, wage bargaining is highly decentralized and takes place at the level of individual firms, the bargaining parties are disciplined by the market for the firm's output. Only if bargaining occurs at the industry level, the bargaining parties might attempt to shift the negative employment effects of wage increases to other industries. But if the bargaining parties in all industries act this way, a higher unemployment rate will be the result.

However, the results of this study partly qualify this claim. First, wage gaps are always determined at the level of individual firms. Because wage gaps are common in countries with collective bargaining at a nationwide *and* industry-wide level (those countries "in between"), one could argue that there are actually no countries with a strictly centralized wage-setting system.⁶ Second, we have shown in chapters

³See also the discussion of the Phillips-curve theory in chapter 3.

⁴Recall that the early literature on the topic was always about wage *drift*, not wage *gaps*. However, a drift rate close to zero does *not* imply that relative wage gaps are decreasing.

⁵This discussion is not purely academic. In Germany, for example, there is a public debate whether firms should be allowed to deviate from collective agreements and pay less than the contract wage when this saves jobs.

⁶Notice that Calmfors and Driffill (1988) mention the Scandinavian countries as examples for highly centralized wage setting. Yet, we have seen in the chapters 3 and 4, that just in these countries explicit *local* wage bargaining about wage gaps is institutionalized. Hence, in these countries wage setting is actually far from being centralized. In view of the results of the mentioned chapters it could even be argued that those countries should perform much worse: unions could

5, 6 and 7, that wage-setting systems that allow for the occurrence of wage gaps, determined at the level of individual firms, can enhance the efficiency of employment relations. Especially if there is asymmetric information about the workers' effort, as was assumed in the models of chapters 5 and 7, these wage-setting systems can even lead to lower wage cost per effective unit of labour.

8.3 SUGGESTIONS FOR FURTHER RESEARCH

One of the conclusions of the literature survey in chapter 3 was that the outcome of models of local wage bargaining (or wage renegotiations) between a firm and its employees is highly sensitive to the assumed bargaining structure. As has become clear in chapter 4, which deals explicitly with this problem, and also in chapter 7, where different wage-setting schemes are evaluated, especially the assumptions about parties' disagreement payoffs are decisive. Although these chapters analyse the consequences of different assumptions, the determinants of the disagreement payoffs are still not fully understood. Probably, the solution will be found at other stages of the wage-formation process. The discussion in chapter 5 already hinted at this direction by arguing that collective agreements do not only determine contract wages but that they also (co-)determine some standards for workers' effort.

Another severe problem is the lack of data on wage gaps. Highly aggregated data on wage *drift* is available for Germany and the Netherlands from the statistical offices, but micro-data that would allow testing the models developed in this thesis is extremely scarce. The German company panels mentioned in chapter 2 might help filling the gap. Yet for the Netherlands, such studies are still missing to the knowledge of the author.

(ab)use the contract wage set on a centralized level to rip off firms in a second round of bargaining. Notice however, that we have shown in chapter 7 that local wage bargaining can enhance the efficiency of employment relations.

Bibliography

- ADDISON, J. T., K. KRAFT, AND J. WAGNER (1993): "German Works Councils and Firm Performance," in *Employee Representation: Alternatives and Future Directions*, ed. by B. E. Kaufman, and M. M. Kleiner, pp. 305-377. Madison.
- AGHION, P., O. COMPTE, AND P. JEHIEL (1998): "Bargaining while Searching for Outside Options: A Potential Role for Unions," *European Economic Review*, 42, 909-918.
- AKERLOF, G. A. (1982): "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics*, 97, 543-569.
- (1984): "Gift Exchange and Efficiency Wage Theory: Four Views," *American Economic Review*, 74(2), 79-83.
- ALBEDA, W., W. J. DERCKSEN, AND F. H. TROS (1998): *Arbeidsverhoudingen in Nederland*. Samson Bedrijfsinformatie, Alphen aan de Rijn, 6 edn.
- BAKER, G., M. GIBBS, AND B. HOLMSTROM (1994): "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109(4), 921-955.
- BELLMANN, L., AND S. KOHAUT (1995): "Effektiv- und Tariflöhne in der Bundesrepublik Deutschland: Eine empirische Analyse auf Basis des IAB-Betriebspanels," in *Determinanten der Lohnbildung: Theoretische und empirische Untersuchungen*, ed. by K. Gerlach, and R. Schettkat, vol. 3 of *Sozialwissenschaftliche Arbeitsmarktforschung*, pp. 72-93. Edition Sigma, Berlin.
- BINMORE, K. (1987): "Bargaining Models," in *Natural Gas Markets and Contracts*, ed. by R. Golombek, M. Hoel, and J. Vislie, vol. 161 of *Contributions to Economic Analysis*, pp. 239-252. North-Holland, Amsterdam.
- BINMORE, K., A. RUBINSTEIN, AND A. WOLINSKY (1986): "The Nash Bargaining Solution in Economic Modelling," *Rand Journal of Economics*, 17(2), 176-188.
- BINMORE, K., A. SHAKED, AND J. SUTTON (1989): "An Outside Option Experiment," *Quarterly Journal of Economics*, 104, 753-770.
- BLANCHARD, O. J., AND L. H. SUMMERS (1986): "Hysteresis and the European Unemployment Problem," in *NBER Macroeconomics Annual 1986*, ed. by S. Fischer, pp. 15-78. MIT Press, Cambridge, Mass.

- BLUM, J. (1983): *Von der Tarif- zur Effektivlohnstruktur: Eine theoretische und empirische Analyse zum Einfluß der Gewerkschaften auf die vertikale und horizontale Lohnstruktur*, vol. 404 of *Europäische Hochschulschriften, Reihe V*. Verlag Peter Lang, Frankfurt a.M., Bern.
- BRONSTEIN, I., K. SEMENDJAJEW, G. MUSIOL, AND H. MÜHLIG (1997): *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt a.M., 3 edn.
- BULKLEY, G., AND G. D. MILES (1996): "Trade Unions, Efficiency Wages, and Shirking," *Oxford Economic Papers*, 48(1), 75–88.
- (1998): "Response to Goerke," *Oxford Economic Papers*, 50, 305.
- CALMFORS, L. (1990): "Wage Formation and Macroeconomic Policy in the Nordic Countries: A Summary," in *Wage Formation and Macroeconomic Policy in the Nordic Countries*, ed. by L. Calmfors, pp. 11–60. SNS Förlag and Oxford University Press, Stockholm and Oxford.
- CALMFORS, L., AND J. DRIFFILL (1988): "Bargaining Structure, Corporatism and Macroeconomic Performance," *Economic Policy*, 6, 13–62.
- CENTRAAL BUREAU VOOR DE STATISTIEK (1975): *Indexcijfers van regelingslonen, methodebeschrijving en reeksen 1926–1974*. Staatsuitgeverij, The Hague.
- (2000): "StatLine," <http://www.cbs.nl/nl/cijfers/statline/index.htm>.
- CHIANG, A. C. (1984): *Fundamental Methods of Mathematical Economics*. McGraw-Hill, Auckland etc., 3 edn.
- VON DER DECKEN, W. (1964): "Die übertarifliche Bezahlung der Arbeiter in der gewerblichen Wirtschaft und im Dienstleistungsbereich im Oktober 1962," *Wirtschaft und Statistik*, 1964, 707–712.
- DEUTSCHE BUNDESBANK (2000): "Bestimmungsgründe und gesamtwirtschaftliche Bedeutung von Produzenten- und Konsumentenlohn," *Monatsbericht*, 56(7), 15–27.
- (various volumes): "Monatsbericht," Frankfurt a. M.
- DEUTSCHMANN, C., AND R. SCHMIEDE (1983): *Lohnentwicklung in der Bundesrepublik 1960–1978: Wirtschaftliche und soziale Bestimmungsgründe*. Campus Verlag, Frankfurt am Main, New York.
- DOSTAL, W., B. HADLER, J. KÜHL, M. LAHNER, E. ULRICH, E. WENDLER, AND M. WOLFSTEINER (1994): "Konzeption des IAB-Betriebspanels und erste Teilergebnisse," in *Firmenpanelstudien in Deutschland*, ed. by U. Hochmuth, and J. Wagner, pp. 237–254. Francke Verlag, Tübingen and Basel.

- DRIEHUIS, W. (1975): "Inflation, Wage Bargaining, Wage Policy and Production Structure: Theory and Empirical Results for the Netherlands," *De Economist*, 123(4), 638-679.
- FAITH, R. L., AND J. D. REID (1987): "An Agency Theory of Unionism," *Journal of Economic Behavior and Organisation*, 8, 39-60.
- FRANZ, W. (1994): *Arbeitsmarktökonomik*. Springer Verlag, Berlin, etc., 2 edn.
- FUDENBERG, D., AND J. TIROLE (1990): "Moral Hazard and Renegotiation in Agency Contracts," *Econometrica*, 58(6), 1279-1319.
- (1991): *Game Theory*. The MIT Press, Cambridge, Massachusetts.
- GAHLEN, B., AND H. J. RAMSER (1987): "Effizienzlohn, Lohndrift und Beschäftigung," in *Arbeitsmärkte und Beschäftigung: Fakten, Analysen, Perspektiven*, ed. by G. Bombach, B. Gahlen, and A. E. Ott, pp. 129-160. Mohr und Siebeck, Tübingen.
- GERFIN, H. (1969): "Ausmaß und Wirkung der Lohndrift," in *Lohnpolitik und Einkommensverteilung*, ed. by H. Arndt, vol. 51 (Neue Folge) of *Schriften des Vereins für Socialpolitik*, pp. 472-522. Duncker & Humblot, Berlin.
- GIEBELS, R. (1998): "Adieu, Poldermodel," NRC Handelsblad, 4 June 1998.
- DE GIJSEL, P. (1996): "Effort, Bargaining Power, Wage Gap, and Employment in Local Wage Bargaining," Universität Osnabrück.
- DE GIJSEL, P., AND R. OLTHOFF (1998): "Local Wage Bargaining, Efficiency Wages, and the Wage Gap. A Game Theoretic Approach," in *Aspects of the Distribution of Income*, ed. by F. Haslinger, and O. Stöninger-Venkatarama, pp. 441-465. Metropolis-Verlag, Marburg.
- GOERKE, L. (1998): "Comment on 'Trade Unions, Efficiency Wages, and Shirking' by Bulkley and Myles," *Oxford Economic Papers*, 50, 305.
- (2000): "On the Structure of Unemployment Benefits in Shirking Models," *Labour Economics*, 7, 283-295.
- GOULD, F. W. (1967): "The Dynamics of Wage Drift," *The Economic Record*, 43, 371-394.
- GROSSMAN, S., AND O. HART (1983): "An Analysis of the Principal Agent Problem," *Econometrica*, 51, 7-45.
- HALL, R. E., AND E. P. LAZEAR (1984): "The Excess Sensivity of Layoffs and Quits to Demand," *Journal of Labour Economics*, 2(2), 233-257.

- HANSEN, B., AND G. REHN (1956): "On Wage Drift: A Problem of Money-Wage Dynamics," in *25 Economic Essays in Honour of Erik Lindahl*, pp. 87-138. Ekonomisk Tidskrift, Stockholm.
- HART, O. D., AND B. HOLMSTRÖM (1987): "The Theory of Contracts," in *Advances in Economic Theory*, ed. by T. F. Bewley. Cambridge University Press, Cambridge.
- HASLINGER, F., AND O. STÖNNER-VENKATARAMA (eds.) (1998): *Aspects of the Distribution of Income* Marburg. Metropolis-Verlag.
- HAUCAP, J., U. PAULY, AND C. WEY (2000): "The Incentives of Employers' Associations to Raise Rival's Costs in the Presence of Collective Bargaining," in *The Unemployment Debate: Current Issues*, ed. by P. de Gijssel, R. Olthoff, and T. Zwick, pp. 185-226. Metropolis-Verlag, Marburg.
- HOEL, M. (1987): "Bargaining Games with a Random Sequence of Who Makes the Offers," *Economics Letters*, 24, 5-9.
- (1989): "Efficiency Wages and Local Versus Central Wage Bargaining," *Economics Letters*, 30, 175-179.
- HOLDEN, S. (1988): "Local and Central Wage Bargaining," *Scandinavian Journal of Economics*, 90, 93-99.
- (1989): "Wage Drift and Bargaining: Evidence from Norway," *Economica*, 56, 419-432.
- (1998): "Wage Drift and the Relevance of Centralised Wage Setting," *Scandinavian Journal of Economics*, 100(4), 711-731.
- HOLMLUND, B. (1986): "Centralized Wage Setting, Wage Drift and Stabilization Policies under Trade Unionism," *Oxford Economic Papers*, 38, 243-258.
- ISACHSEN, A. J. (1977): "A Note on Wage Drift, the Case of Sweden," *Scandinavian Journal of Economics*, 79, 366-374.
- KAHNEMANN, D., J. L. KNETSCH, AND R. THALER (1986): "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 76(4), 728-741.
- KATZ, L. (1986): "Efficiency Wage Theories: A Partial Evaluation," in *NBER Macroeconomic Annual 1986*, ed. by S. Fischer, pp. 235-276. The MIT Press, Cambridge, Massachusetts and London, England.
- KLEIERL, F. (1993): "Zum Verhältnis von Effizienzlöhnen und Tariflöhnen - Theoretische Erklärungsansätze und empirische Evidenz," Master's thesis, Universität Regensburg, Regensburg.

- KLEINHÜCKELSKOTEN, H. D., AND D. SPAETLING (1980): *Aspekte der Interdependenz zwischen Tarif- und Effektivlohnentwicklung*, vol. 2943 of *Forschungsberichte des Landes Nordrhein-Westfalen*. Westdeutscher Verlag, Opladen.
- LAFFONT, J.-J., E. MASKIN, AND J.-C. ROCHET (1987): "Optimal Non-linear Pricing with Two-dimensional Characteristics," in *Information, Incentives, and Economic Mechanisms*, ed. by T. Groves, R. Radner, and S. Reiter. University of Minnesota Press.
- LAZEAR, E. P. (1979): "Why is there Mandatory Retirement?," *Journal of Political Economy*, 87, 1261-1284.
- (1981): "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *American Economic Review*, 71, 606-620.
- (1995): *Personnel Economics*. MIT Press, Cambridge, Massachusetts and London, England.
- LERNER, S. W., AND J. MARQUAND (1962): "Workshop Bargaining, Wage Drift and Productivity in the British Engineering Industry," *The Manchester School of Economic and Social Studies*, 30, 15-60.
- LEVER, M. H. C. (1993): "Union Wage Formation and (Un)employment," Ph.D. thesis, Universitaire Pers Maastricht, Maastricht.
- LINDBECK, A., AND D. J. SNOWER (1988): "Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach," *American Economic Review*, 78(1), 167-188.
- MACLEOD, W. B., AND J. M. MALCOMSON (1993): "Investments, Holdup, and the Form of Market Contracts," *American Economic Review*, 83, 811-837.
- MALCOMSON, J. M. (1997): "Contracts, Hold-Up, and Labor Markets," *Journal of Economic Literature*, 35, 1916-1957.
- MARQUAND, J. (1967): "Wage Drift: Origins, Measurement and Behaviour," Woolwich Economic Papers 14, Woolwich London.
- MARSHALL, A. (1920): *Principles of Economics*. Macmillan, London, 8 edn.
- MCAFEE, P., AND J. MCMILLAN (1988): "Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory*, 46, 335-354.
- MCCALL, J. J. (1970): "Economics of Information and Job Search," *Quarterly Journal of Economics*, 84, 113-126.
- MCLAUGHLIN, K. J. (1994): "Rigid Wages?," *Journal of Monetary Economics*, 34(3), 383-414.

- MEYER, W. (1994): "Übertarifliche Bezahlung als Anreizinstrument," in *Firmen-panelstudien in Deutschland*, ed. by U. Hochmuth, and J. Wagner, pp. 219–235. Francke Verlag, Tübingen and Basel.
- (1995): "Analyse der Bestimmungsfaktoren der „übertariflichen Entlohnung“ auf der Basis von Firmendaten," in *Determinanten der Lohnbildung: Theoretische und empirische Untersuchungen*, ed. by K. Gerlach, and R. Schettkat, vol. 3 of *Sozialwissenschaftliche Arbeitsmarktforschung*, pp. 50–71. Edition Sigma, Berlin.
- MILKOVICH, G. T., AND J. M. NEWMAN (1996): *Compensation*. Irwin, Chicago etc., 5 edn.
- MIRRELES, J. (1975): "A Pure Theory of Underdeveloped Economies," in *Agriculture in Development Theory*, ed. by L. A. Reynolds, pp. 84–106. Yale University Press, New Haven.
- MOENE, K. O. (1988): "Unions' Threats and Wage Determination," *Economic Journal*, 98, 471–483.
- MÖLLER, J. (1990): "Das Hysteresis Phänomen in Arbeitslosigkeit-Inflations-Modellen: Theorie, empirische Relevanz und wirtschaftspolitische Konsequenzen," habilitation thesis, Universität Konstanz.
- MORTENSEN, D. T. (1970): "Job Search, the Duration of Unemployment, and the Phillips Curve," *American Economic Review*, 60, 847–862.
- MUYSKEN, J., AND T. VAN VEEN (1996): "Efficiency Wages and Local Wage Bargaining," *Scandinavian Journal of Economics*, 98, 119–127.
- NASH, JR., J. F. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.
- OSWALD, A. J. (1985): "The Economic Theory of Trade Unions: An Introductory Survey," *Scandinavian Journal of Economics*, 87, 160–193.
- PEHKONEN, J., AND J. VISKARI (1994): "Wage Drift: Phillips Curve vs Bargaining Models," *Labour*, 8(3), 395–421.
- PHELPS BROWN, E. H. (1962): "Wage Drift," *Economica*, 29, 339–356.
- RASMUSEN, E. (1994): *Games and Information: An Introduction to Game Theory*. Blackwell, Cambridge, Massachusetts, 2 edn.
- VAN REIJN, H. (1982): "De ontwikkeling van de verdiende lonen en de regelingslonen in de jaren 1972–'80," in *Sociaal-economische maandstatistiek*, no. 1982:1, pp. 53–70. Centraal Bureau voor de Statistiek.
- (1984): "Verdiende lonen versus regelingslonen; de ontwikkeling in de jaren 1972–'83," in *Sociaal-economische maandstatistiek*, no. 1984:10, pp. 18–28. Centraal Bureau voor de Statistiek.

- ROBAK, B. (1978): *Industriezweigstruktur und übertarifliche Entlohnung: Zur Differenz von Tarif- und Effektivlöhnen in der westdeutschen Industrie*, vol. 271 of *Volkswirtschaftliche Schriften*. Berlin.
- ROCHET, J.-C. (1985): "The Taxation Principle and Multi-time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, 14, 113-128.
- ROCHETEAU, G. (2001): "Equilibrium Unemployment and Wage Formation with Matching Frictions and Worker Moral Hazard," *Labour Economics*, 8, 75-102.
- RUBINSTEIN, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.
- SALOP, S. C. (1979): "A Model of the Natural Rate of Unemployment," *American Economic Review*, 69(1), 117-125.
- SCHLICHT, E. (1978): "Labour Turnover, Wage Structure, and Natural Unemployment," *Journal of Institutional and Theoretical Economics (Zeitschrift für die gesamte Staatswissenschaft)*, 134, 337-346.
- (1992): "Wage Generosity," *Journal of Institutional and Theoretical Economics (Zeitschrift für die gesamte Staatswissenschaft)*, 148, 437-451.
- SCHMIDT, E. M. (1991): "Produzieren in Niedersachsen und Baden-Württemberg, Konzeptionen und Ergebnisse aus der Pilotstudie zum Firmenpanel," *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, pp. 726-734.
- (1992): "Produzieren in Niedersachsen," in *Produzieren in Niedersachsen: Empirische Untersuchungen mit Betriebsdaten*, ed. by E. Ertel, and J. Wagner, pp. 107-113. Niedersächsisches Institut für Wirtschaftsforschung, Hannover.
- SCHNABEL, C. (1995): "Übertarifliche Entlohnung: Einige Erkenntnisse auf Basis betrieblicher Effektivverdienststatistiken," in *Determinanten der Lohnbildung: Theoretische und empirische Untersuchungen*, ed. by K. Gerlach, and R. Schettkat, vol. 3 of *Sozialwissenschaftliche Arbeitsmarktforschung*, pp. 28-49. Edition Sigma, Berlin.
- (1997): *Tariflohnpolitik und Effektivlohnfindung: Eine empirische und wirtschaftspolitische Analyse für die alten Bundesländer*, vol. 3 of *Kollektive Entscheidungen, Wirtschaftspolitik und öffentliche Finanzen*. Peter Lang, Frankfurt am Main.
- SCHULTE NORDHOLT, E. (2001): "Incidentele loonontwikkeling van werknemers, 1995-1999 (Development of non-contractual wages of employees, 1995-1999)," in *Sociaal-economische maandstatistiek*, no. 2001:6, pp. 23-29. Centraal Bureau voor de Statistiek.
- SHAKED, A. (1994): "Opting Out: Bazaars Versus 'High Tech' Markets," *Investigaciones Económicas*, 18(3), 421-432.

- SHAKED, A., AND J. SUTTON (1984): "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model," *Econometrica*, 52(6), 1351-1364.
- SHAPIRO, C., AND J. E. STIGLITZ (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433-444.
- SMITH, A. (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Campbell, R. H. and Skinner, A. S. (1976). Clarendon Press (reprint 1979), Oxford.
- SOBEL, J., AND I. TAKAHASHI (1983): "A Multistage Model of Bargaining," *Review of Economic Studies*, 50, 411-426.
- SÖDERSTRÖM, H. T., AND E. UDDÉN-JONDAL (1982): "Does Egalitarian Wage Policy Cause Wage Drift? An Empirical Study of Sweden 1960-1979," Seminar Paper 203, Institute for International Economic Studies.
- SOLOW, R. M. (1985): "Insiders and Outsiders in Wage Determination," *Scandinavian Journal of Economics*, 87, 411-428.
- STATISTISCHES BUNDESAMT (1962): "Fachserie M, Reihe 17/I," Wiesbaden.
- STIGLITZ, J. E. (1974): "Wage Determination and Unemployment in L.D.C.'s: The Labor Turnover Model," *Quarterly Journal of Economics*, 88, 194-227.
- (1976): "The Efficiency Wage Hypothesis, Surplus Labour and the Distribution of Income in L.D.C.s," *Oxford Economic Papers*, 28(2), 185-207.
- (1987): "The Causes and Consequences of the Dependence of Quality on Price," *Journal of Economic Literature*, 25, 1-48.
- TESCHNER, E. (1977): "Lohnpolitik im Betrieb: Eine empirische Untersuchung in der Metall-, Chemie-, Textil- und Tabakindustrie," Studienreihe des Instituts für Sozialforschung, Institut für Sozialforschung, Frankfurt a. M., New York.
- TEULINGS, C. N. (1995): "Corporatist Wage Setting in Models of Surplus Sharing," Discussion Paper TI 3-95-197, Tinbergen Institute.
- TEULINGS, C. N., AND J. HARTOG (1998): *Corporatism or Competition: Labour Contracts, Institutions and Wage Structures in International Comparison*. Cambridge University Press, Cambridge (UK).
- VARIAN, H. R. (1992): *Microeconomic Analysis*. W.W. Norton & Company, New York, London, 3 edn.
- VAN VEEN, A. P. (1997): "Studies in Wage Bargaining: The Influence of Taxes and Social Security Contributions on Wages," Ph.D. thesis, Universitaire Pers Maastricht, Maastricht.

- WALTHER, H. (1989): "Kündigungsschutz und Effizienzloohnhypothese," in *Effizienzlohntheorie, Individualeinkommen und Arbeitsplatzwechsel*, ed. by K. Gerlach, and O. Hübler, Sozialwissenschaftliche Arbeitsmarktforschung, pp. 97-120. Campus Verlag, Frankfurt a.M., New York.
- WEITZMAN, M. L. (1987): "Steady State Unemployment under Profit Sharing," *Economic Journal*, 97(1), 86-105.
- VAN DER WERF, R. E. J. (1987): "Vergelijking van de ontwikkeling van de verdiende lonen en regelingslonen, 1977-1985 - Cijfers en kanttekeningen," in *Sociaal-economische maandstatistiek*, no. 1987:1, pp. 13-18. Centraal Bureau voor de Statistiek.
- YELLEN, J. L. (1984): "Efficiency-Wage Models of Unemployment," *American Economic Review*, 74, 200-205.
- VAN DER ZWAN, J. (1995): "De ontwikkeling van het verdiend loon in componenten, 1985-1993 (The development of earnings in components, 1985-1993)," in *Sociaal-economische maandstatistiek*, no. 1995:7, pp. 26-29. Centraal Bureau voor de Statistiek.

Samenvatting (Dutch summary)

In landen met een stelsel van collectieve onderhandelingen tussen vakbonden en werkgeversorganisaties op landelijk niveau of het niveau van bedrijfstakken, betalen werkgevers vaak een toeslag op de CAO-lonen. Deze toeslagen, de zogenoemde incidentele lonen, kunnen een wezenlijk deel uitmaken van het inkomen van werknemers. Ondanks hun naam vertonen zij een opmerkelijke persistentie over de tijd.

Dit proefschrift is een theoretisch onderzoek naar het ontstaan en de persistentie van incidentele lonen. Het probleem wordt vanuit verschillende richtingen benaderd. Het leidmotief is echter doorgaans de rol die efficiëntielonen en onderhandelingsmacht van werknemers en werkgevers hierin spelen.

Onderhandelingsmacht beïnvloedt het loonvormingsproces op alle niveaus. Tijdens CAO-onderhandelingen is dit duidelijk zichtbaar. Onderhandelingsmacht is echter ook belangrijk op het bedrijfsniveau, namelijk als renten gecreëerd worden, bijvoorbeeld door transactiekosten of *moral hazard*.

Efficiëntielonen zijn voor dit onderzoek met name belangrijk vanwege hun motivatieaspect. Volgens de efficiëntieloontheorie bestaat er een positieve samenhang tussen de prestatie en het loon van een werknemer. Een onderneming kan van die samenhang gebruik maken en haar arbeidskosten minimaliseren door een hoger loon dan het marktloon te betalen. Werknemers, aan de andere kant, worden hierdoor deels afgeschermd van ontwikkelingen op de arbeidsmarkt, waardoor zij een sterkere onderhandelingspositie tegenover de onderneming kunnen opbouwen.

Hoofdstuk 1 bevat een beknopte inleiding in het onderwerp en een motivatie van het onderzoek. Verder wordt de opbouw van het boek uitgelegd.

Hoofdstuk 2 behandelt de empirische achtergrond voor het theoretisch onderzoek in de volgende hoofdstukken. Het beschrijft de arbeidsverhoudingen in Nederland en Duitsland en de ontwikkeling van de loondrift, de verandering van de incidentele lonen over de tijd, in de twee landen. Verder worden verscheidene empirische studies over de oorzaken van incidentele lonen en loondrift in Nederland en Duitsland besproken. Deze studies steunen de assumptie dat efficiëntielonen en onderhandelingsmacht een grote rol spelen bij het totstandkomen van incidentele lonen.

Hoofdstuk 3 geeft een overzicht van de bestaande theorieën over incidentele lonen en loondrift. De verscheidene modellen kunnen worden samengevat in drie groepen: de zogenoemde Phillipscurve-modellen, efficiëntieloon-modellen en modellen met locale onderhandelingen op bedrijfsniveau. Volgens de Phillipscurve-modellen worden de effectieve lonen bepaald op een competitieve arbeidsmarkt. In een situatie van een almaar stijgende vraag naar arbeid kunnen de CAO-lonen de effectieve lonen niet snel genoeg volgen, zodat er een verschil tussen beide ontstaat, het incidentele loon.

De efficiëntieloon-modellen verklaren incidentele lonen uit de eerder beschreven positieve samenhang tussen prestatie en loon. In deze modellen wordt die samenhang mede bepaald door het CAO-loon, waardoor bij een stijging van het CAO-loon ook het effectieve loon wordt opgedreven. De modellen met lokale onderhandelingen ten slotte geven als oorzaak van incidentele lonen expliciete onderhandelingen tussen management en werknemers op bedrijfsniveau, die op CAO-onderhandelingen op een centraler niveau volgen, aan.

De volgende hoofdstukken zijn ingedeeld in twee delen. In deel I (hoofdstukken 4 en 5) staan efficiëntielonen en onderhandelingen tussen werkgevers en werknemers centraal. Hierbij wordt ingegaan op een aantal problemen die bij de discussie van bestaande modellen van incidentele lonen, in het bijzonder de modellen van lokale onderhandelingen, naar voren zijn gekomen.

Hoofdstuk 4 behandelt de invloed van dreigpunten op het onderhandelingsresultaat in modellen van lokale loononderhandelingen met efficiëntielonen. In een algemene opzet wordt het onderhandelingsproces tussen werkgever en werknemers op bedrijfsniveau onderzocht. Het resultaat is dat de eerder besproken modellen kunnen worden samengevat onder twee parameterconstellaties van CAO-loon, efficiëntieloon en onderhandelingskosten, die - afhankelijk van de verdeling van onderhandelingsmacht - tot een incidenteel loon kunnen leiden. Tevens wordt een derde mogelijke constellatie geïdentificeerd, die onder alle omstandigheden tot een incidenteel loon leidt.

In hoofdstuk 5 wordt ingegaan op de invloed van vertegenwoordigers van werknemers op de gestalte van de effort-functie, d. w. z. de functionele samenhang tussen het loon en de prestatie van werknemers. Het hoofdstuk gaat uit van de observatie dat vakbonden en ondernemingsraden veelvuldig invloed uitoefenen op bijvoorbeeld de arbeidsomstandigheden en ontslagprocedures. In het model koppelen vakbonden het CAO-loon aan een bepaald prestatieniveau. Hierdoor maken zij het winstmaximaliserende loon, het zogenoemde efficiëntieloon, afhankelijk van het CAO-loon. Het resultaat is dat werkgevers hierdoor in de meeste gevallen gedwongen worden een toeslag op het CAO-loon te betalen. Bovendien wordt aangetoond dat het incidentele loon positief van het CAO-loon afhangt. Het laatste deel van dit hoofdstuk gaat in op de consequenties voor het gedrag van vakbonden, afhankelijk van de mate waarin zij rekening houden met de werkgelegenheid van hun leden.

Deel II behandelt het probleem van incidentele lonen vanuit een ander perspectief. Nu staat de vraag centraal of incidentele lonen een efficiënt antwoord kunnen zijn op onvolkomenheden in de arbeidsrelatie tussen een werkgever en één werknemer. Onder efficiëntie wordt hierbij verstaan dat de arbeidsrelatie in stand wordt gehouden zolang zij een positieve netto-uitbetaling boven de *outside options* genereert, en dat de twee partijen voor een efficiënt niveau van investeringen in de arbeidsrelatie kiezen.

In hoofdstuk 6 wordt uitgegaan van een specifiek type investeringen waarvan de opbrengsten naar de investeerder terugvloeien. Het wordt beargumenteerd dat in dit geval beide partijen baat hebben bij een vooraf afgesproken star loon opdat geen partij zich de opbrengsten van de ander kan toeëigenen hetgeen tot suboptimale

investeringen zou leiden. Een te star loon kan echter ook tot de ontbinding van de arbeidsrelatie leiden, met name als een partij vaststelt dat haar nut of winst met een andere partner hoger zou zijn. In dit geval zouden beide partijen juist baat hebben bij een flexibel loon. In het hoofdstuk wordt een onderhandelingsmodel voorgesteld dat alleen tot loonsveranderingen leidt als de arbeidsrelatie door de *outside options* in gevaar komt. Het incidentele loon is gelijk aan de afwijking van het onderhandelde loon en het vooraf afgesproken loon. In het hoofdstuk wordt ook aangetoond dat dergelijke veranderingen van het loon alleen mogelijk zijn in een situatie van complete informatie. In het geval van asymmetrische informatie over de *outside option* van de andere partij kan het loon echter niet voldoende worden aangepast om een inefficiënte ontbinding van de relatie te vermijden.

Hoofdstuk 7 neemt de structuur van het voorafgaande hoofdstuk over maar veronderstelt een ander type investeringen. Nu moet de werknemer kiezen tussen investeringen in zijn prestatie in het bedrijf en investeringen in zijn *outside option*; vergeleken met hoofdstuk 6 is de *outside option* van de werknemer nu geëndogeniseerd. Bovendien kan de werkgever de investeringen niet direct waarnemen. In dit hoofdstuk wordt nagegaan in hoeverre verschillende loonvormingsprocessen, die vaker in de arbeidseconomische literatuur aangehaald worden, onder deze omstandigheden tot efficiënte resultaten leiden. Het model toont aan dat door de endogenisering van de beslissing van de werknemer in geen van de onderzochte processen inefficiënte ontbindingen van de arbeidsrelatie dreigen. Efficiënte investeringen worden echter alleen bereikt in loonvormingsprocessen die in een vooraf vastgesteld loon (bijvoorbeeld het CAO-loon) en daarop opbouwende onderhandelingen over een loontoeslag voorzien. Het eerder in dit proefschrift besproken model van locale onderhandelingen leidt dus tot efficiënte resultaten.

In hoofdstuk 8 worden de resultaten van dit proefschrift samengevat. Tevens worden enkele suggesties voor economisch beleid en verder onderzoek gedaan.

Zusammenfassung (German summary)

In Ländern, in denen Tarifverhandlungen entweder auf nationaler Ebene oder auf Branchenebene geführt werden, bezahlen Arbeitgeber häufig einen Zuschlag auf den Tariflohn. Diese Zuschläge, die sogenannte Lohnspanne, können einen wesentlichen Teil des Arbeitnehmereinkommens ausmachen. Sie sind darüber hinaus über die Zeit gesehen auffallend beständig.

Diese Arbeit ist eine theoretische Studie über die Ursachen der Entstehung und der Persistenz von Lohnspannen. Das Problem wird aus verschiedenen Blickwinkeln betrachtet. Dabei stehen jedoch immer Effizienzlöhne und die Verhandlungsmacht von Arbeitnehmern und Arbeitgebern als wichtige Faktoren im Mittelpunkt.

Verhandlungsmacht beeinflusst den Lohnsetzungsprozeß auf allen Ebenen. In Tarifverhandlungen wird dies immer wieder deutlich. Verhandlungsmacht ist aber auch auf Betriebsebene wichtig, wenn nämlich, zum Beispiel durch Transaktionskosten oder moralische Risiken, Renten kriert werden, die es zu verteilen gilt.

Effizienzlöhnen wird in dieser Arbeit wegen des Motivationsaspekts eine besondere Bedeutung beigemessen. Laut Effizienzlohn-Theorie besteht zwischen der Leistung und dem Lohn eines Arbeitnehmers ein positiver Zusammenhang. Ein Unternehmen kann diesen Zusammenhang nutzen und seine Arbeitskosten minimieren, indem es einen höheren Lohn als den Marktlohn bezahlt. Arbeitnehmer wiederum werden hierdurch von den Entwicklungen auf dem Arbeitsmarkt teilweise abgeschirmt. Das versetzt sie in die Lage, eine Verhandlungsposition gegenüber dem Unternehmen aufzubauen.

Kapitel 1 beinhaltet eine kurze Einleitung in das Thema und eine Motivierung der Arbeit. Außerdem wird der Aufbau des Buches erläutert.

Kapitel 2 behandelt den institutionellen und empirischen Hintergrund der nachfolgenden theoretischen Modelle. Es beschreibt die wichtigen Arbeitsmarktinstitutionen in den Niederlanden und Deutschland, sowie die Entwicklung der Lohndrift, der zeitlichen Veränderung der Lohnspanne, in diesen beiden Ländern. Weiterhin werden verschiedene empirische Studien über die Ursachen von Lohnspannen und Lohndrift in den Niederlanden und in Deutschland besprochen. Diese Studien unterstützen die These dieser Arbeit, daß Effizienzlöhne und Verhandlungsmacht eine große Rolle beim Zustandekommen von Lohnspannen spielen.

Kapitel 3 gibt einen Überblick über die theoretische Literatur über Lohnspannen und Lohndrift. Die verschiedenen Modelle können zu drei Gruppen zusammengefaßt werden: den sogenannten Phillipskurvenmodellen, Effizienzlohnmodellen und Modellen mit lokalen Lohnverhandlungen auf Betriebsebene. Ersteren zufolge werden die effektiven Löhne auf einem kompetitiven Arbeitsmarkt bestimmt. Bei einer stetig

steigenden Arbeitsnachfrage können die Tariflöhne mit den effektiven Löhnen nicht mehr Schritt halten, so daß zwischen beiden eine Differenz entsteht, die Lohnspanne. Effizienzlohnmodelle erklären Lohnspannen aus dem oben beschriebenen positiven Zusammenhang zwischen Arbeitsleistung und Lohn. Dieser Zusammenhang wird in diesen Modellen unter anderem durch den Tariflohn bestimmt; bei einer Erhöhung des Tariflohns wird auch der effektive Lohn in die Höhe getrieben. Die Modelle lokaler Lohnverhandlungen schließlich sehen die Ursache von Lohnspannen in expliziten Verhandlungen zwischen Management und Arbeitnehmern auf Betriebsebene, die den Tarifverhandlungen auf zentraler Ebene nachgelagert sind.

Der Rest der Arbeit gliedert sich in zwei Teile. In Teil I (Kapitel 4 und 5) stehen Effizienzlöhne und die Verhandlungen zwischen Arbeitgebern und Arbeitnehmern im Mittelpunkt. Beide Kapitel gehen näher auf einige Probleme ein, die bereits bei der Diskussion bestehender Modelle über die Entstehung von Lohnspannen, insbesondere der Modelle lokaler Lohnverhandlungen, identifiziert wurden.

Kapitel 4 behandelt den Einfluß von Drohpunkten auf das Verhandlungsergebnis in Modellen lokaler Lohnverhandlungen, wenn Unternehmen Effizienzlöhne bezahlen. In einem allgemein gefaßten Modell wird der Verhandlungsprozeß zwischen Arbeitgebern und Arbeitnehmern auf Betriebsebene analysiert. Das Ergebnis ist, daß die im vorangehenden Kapitel besprochenen Modelle zu zwei Gruppen zusammengefaßt werden können, die jeweils mit einer Parameterkonstellationen von Tariflohn, Effizienzlohn und Verhandlungskosten korrespondieren. In beiden Konstellationen kann, abhängig von der Verteilung der Verhandlungsmacht, eine Lohnspanne entstehen. Zudem wird eine dritte mögliche Konstellation identifiziert, die unter allen Umständen zu einer Lohnspanne führt.

Kapitel 5 untersucht die Konsequenzen des Einflusses, den Arbeitnehmervertreter auf die Form der Effort-Funktion, den funktionalen Zusammenhang zwischen Arbeitsleistung und Lohn, ausüben. Ausgangspunkt ist die Beobachtung, daß Gewerkschaften und Betriebsräte vielfältig Einfluß auf Arbeitsumstände und das Entlassungsverfahren ausüben. In dem vorgestellten Modell koppeln Gewerkschaften den Tariflohn an ein bestimmtes Leistungsniveau. Dadurch machen sie den gewinnmaximierenden Lohn, den sogenannten Effizienzlohn, abhängig vom Tariflohn. Das Ergebnis ist, daß Arbeitgeber in den meisten Fällen gezwungen sind, einen Zuschlag auf den Tariflohn zu zahlen. Außerdem wird gezeigt, daß die entstehende Lohnspanne positiv von der Höhe des Tariflohns abhängt. Der letzte Teil dieses Kapitels beschäftigt sich mit den Konsequenzen dieses Modells für das Verhalten von Gewerkschaften.

Teil II (Kapitel 6 und 7) dieser Arbeit behandelt das Problem der Lohnspannen aus einem anderen Blickwinkel. Nun steht die Frage im Mittelpunkt, ob Lohnspannen eine effiziente Lösung für die Unvollkommenheiten des Beschäftigungsverhältnisses zwischen einem Arbeitgeber und einem Arbeitnehmer sein können. Unter Effizienz wird dabei verstanden, daß das Beschäftigungsverhältnis aufrechterhalten wird, solange es eine positive Netto-Auszahlung generiert, die die *Outside options* übersteigt, und daß beide Parteien sich für optimale Investitionen in das Beschäftigungsverhältnis entscheiden.

Kapitel 6 geht von einer für Beschäftigungsverhältnisse nicht unbedingt typischen Art von Investitionen aus, deren Erträge an den Investor direkt zurückfließen. In diesem Fall haben beide Parteien Vorteil an einem im vorhinein abgesprochenen festen Lohn, damit keine der Parteien sich die Investitionserträge der anderen aneignen kann, was zu suboptimalen Investitionen führen würde. Ein zu starrer Lohn kann allerdings auch zur Auflösung des Beschäftigungsverhältnisses führen, besonders dann, wenn eine der Parteien feststellt, daß ihr Nutzen oder Gewinn mit einem anderen Partner höher wäre. In solch einem Fall hätten beide Parteien gerade einen Vorteil an einem flexiblen Lohn. In dem Kapitel wird ein Verhandlungsmodell vorgestellt, das nur dann zu einer Änderung des Lohnes führt, wenn das Beschäftigungsverhältnis durch gestiegene *Outside options* in Gefahr kommt. Die Lohnspanne ist identisch mit der Abweichung des nachverhandelten Lohns vom vorher abgesprochenen Lohn. Darüber hinaus wird gezeigt, daß solche Lohnanpassungen nur möglich sind, wenn beide Parteien über vollständige Informationen verfügen. Bei asymmetrischen Informationen über die *Outside option* der jeweils anderen Partei kann der Lohn nicht ausreichend angepaßt werden, um eine ineffiziente Auflösung des Beschäftigungsverhältnisses zu vermeiden.

Kapitel 7 übernimmt die Struktur des vorangehenden Kapitels, geht aber von einem anderen Typ Investitionen aus. Jetzt muß sich der Arbeitnehmer zwischen Investitionen in seine Arbeitsleistung im Betrieb und Investitionen in seine *Outside option* entscheiden; im Vergleich zu Kapitel 6 ist die *Outside option* des Arbeitnehmers nun also endogenisiert. Außerdem wird angenommen, daß der Arbeitgeber die Investitionen nicht direkt beobachten kann. In diesem Kapitel wird der Frage nachgegangen, inwieweit unterschiedliche Lohnsetzungsprozesse, die häufiger in der arbeitsmarktökonomischen Literatur beschrieben werden, unter den genannten Umständen zu effizienten Ergebnissen führen. Es wird gezeigt, daß durch die Endogenisierung der Entscheidung des Arbeitnehmers in keinem der untersuchten Prozesse eine ineffiziente Auflösung des Beschäftigungsverhältnisses droht. Effiziente Investitionen werden allerdings nur in Lohnsetzungsprozessen erreicht, in denen ein leistungsabhängiger Aufschlag auf einen im vorhinein festgesetzten Lohn (zum Beispiel den Tariflohn) gezahlt wird. Dieser Lohnsetzungsprozeß entspricht weitgehend den an anderer Stelle in dieser Arbeit besprochenen Modellen lokaler Lohnverhandlungen. Diese führen also zu Ergebnissen, die nach den beschriebenen Kriterien effizient sind.

In Kapitel 8 schließlich werden die Ergebnisse dieser Arbeit zusammengefaßt. Zudem werden einige wirtschaftspolitische Empfehlungen sowie Anregungen für weiterführende Forschung gegeben.

CURRICULUM VITAE

Ralph Olthoff werd op 15 augustus 1969 geboren in Peine (Duitsland). Van 1981 tot 1988 volgde hij het voorbereidend wetenschappelijk onderwijs aan het Ratsgymnasium Peine. Na het behalen van het *Abitur* begon hij in 1988 met de studie economische wetenschappen aan de Universität Hannover (Duitsland). Bij het schrijven van zijn doctoraalscriptie „Alterssicherung und wirtschaftliche Entwicklung” werd hij begeleid door Prof. Dr. Dr. Franz Haslinger. De studie werd in april 1995 afgesloten met de titel Diplom-Ökonom.

Van 1995 tot 2002 was hij als assistent in opleiding en later als docent werkzaam aan de Universiteit Maastricht, waar hij onder leiding van Peter de Gijzel onderzoek verrichtte op het gebied van micro-economie en arbeidsmarkteconomie. De meeste resultaten van dit onderzoek zijn in dit proefschrift weergegeven.

Sinds september 2002 werkt hij als docent op de Universiteit Utrecht.

COLOPHON

This book was written and composed by the author, using Donald Knuth's typesetting programme \TeX with Leslie Lamport's macro package \LaTeX , numerous additional packages and own modifications.

The typeface used is Computer Modern Roman 10 on 12pt and 9 on 11pt, designed by Donald Knuth.

The offset-printing and binding work was done by Datawyse in Maastricht.

The cover shows a version of Otto von Guericke's Magdeburg sphere experiment as shown by Ferrari in *Philosophia Peripatetica*, 1745. In 1654, Guericke first demonstrated the enormous force of air pressure. He placed two copper bowls together to form a hollow sphere about 35.5 cm in diameter. After the air had been removed from the sphere, horses were unable to pull the bowls apart, although only the air around them held them together.

The Magdeburg sphere is an allegory for employment relations. Workers and employer are pulling in different directions. Yet, both parties prefer their relation over choosing their outside options. So it is the pressure from the outside world which holds them together. (And no, this should *not* suggest that employment relations only comprise a vacuum.)

