This paper reviews the empirical literature on tests of different wage theories of relevance in particular to the Swedish labour market. The empirical results are confronted with the institutional changes in Sweden during the last twenty years. Not much empirical support can be found for the competitive model, the shirking model or the insider-outsider model. The fair wage version of efficiency wage setting receives support, however. Efficiency wage setting appears to have become more important also for Sweden as a consequence of decentralisation of wage bargaining giving scope for firms to differentiate wages. Due to the obvious institutional importance, bargaining models of wage formation continues to play an important role for Swedish wage setting. Bargaining models combined with fair wage setting appear to capture much of present day wage setting in Sweden.

Keywords: Efficiency wages, trade union models
JEL classification: J31, J51, J53

September 19, 2005

*FIEF, Wallingatan 38, SE-111 24 Stockholm, Sweden. E-mail: per.lundborg@fief.se
1. Introduction

Wage formation is of central importance to understanding economic development in almost all dimensions including inequality, unemployment, inflation, productivity, welfare etc. Since labour market institutions largely determine wage formation, the issue of how wages are set cannot be separated from a country’s economic history. Thus, the institutions determining wages have evolved over time as a result of the social, economic and political objectives. Not only do wage formation practices differ over time but also across countries and across sectors within countries and even across firms within the sectors. Needless to say, understanding wage setting practices is a highly complicated matter calling for a great deal of generalisations.

Some sixty years ago the view of the labour market was no different from the view of other markets in those days, i.e. the labour market was treated a competitive one where the wage and employment was the outcome of the interaction of supply and demand. This view of the labour market may be reasonable for some sectors of the modern economy, particularly in the US, but for most European countries only very few sectors can be adequately described as competitive. For good reasons this view has been abandoned as a main instrument. Nevertheless, in for instance modern international economics or growth theory, students continue to assume that the wage of the economy is set in the orthodox competitive way.

Labour economists early realised that this view was too simplistic to serve as a realistic description of wage setting in the real world. There were many obvious observations that simply did not conform to the competitive model. The fact that the labour market generally does not clear, leaving large segments of the labour force unemployed, was an obvious sign of the failure of the competitive model. A related observation was the asymmetry that wages were upward flexible but downward rigid. The existence of trade unions that bargained with employers on the wage could obviously not be disregarded. Others noted that there was a coexistence of vacancies and unemployment in the labour market, which was incompatible with a competitive model. Finally, unlike commodities, workers react to incentives, suggesting that the firm could increase profits by raising the wage so as to raise workers’ effort and thus productivity of the firm.

Over the years, a large number of competing models have been launched and while a few have disappeared from the “market”, many continue to be used as analytical tools for the labour market. Depending on the issue at
hand, economists analyse wages by means of trade union models, search models, efficiency models and others. But theory has a strong lead over empirics and we do not have a particularly clear view of which model is the most relevant one and for which sectors. Clearly, the heterogeneity of a modern labour market leaves scope for a large set of wage setting models to be of relevance.

We shall not go into the technical details of the wage setting models since there is plenty of literature that fills that purpose. We shall instead review the development of models of wage formation and focus on their empirical relevance and, in particular, their ability to explain unemployment and wage stickiness. The assumptions in the models will be confronted with Swedish labour market institutions and we shall present the available empirical evidence.

A major motivation is that there exist many different theoretical models to explain wage setting and unemployment in the labour economics literature. However, not all of these are of empirical relevance, and we shall try to single out some that are relevant to a country like Sweden. This review is confronted with the institutional and structural changes in the labour market. Our results point at bargaining models and fair wage models as being of relevance and that search theory is an indispensable tool for certain analysing specific issues.

2. A review of wage setting models

The problems of the competitive model

Despite the periods of high unemployment in the 1930s, economists described the labour market in terms of supply and demand curves. The fact that labour supply exceeded demand was a problem that could be solved by demand management according to the Keynesian theories of the days. This view dominated also into the 1960s and to justify that unemployment existed in the competitive markets, the proponents of the Walrasian view acknowledged the frictional elements of the labour market. For instance, in an often cited passage, Friedman (1968) in explaining the natural rate, remarked that “the `natural rate of unemployment` is the level that would be ground out by the Walrasian system of general equilibrium equations, provided there is imbedded in them the actual structural characteristics of the labour and commodity markets, including market imperfections, stochastic variability in demands and supplies, the cost of
gathering information about job vacancies and labor availabilities, the costs of mobility, and so on.”

However, such frictions could hardly explain the high unemployment rates of the great depression and the shortcomings of the competitive model in describing the labour market became obvious when unemployment in the OECD area drastically increased and remained at high levels in the 1970s and 1980s. Moreover, when economists looked in more detail into the existing models, they were unhappy with the microeconomic underpinnings of the models that should explain why wages are not set so as to clear the labour market.

The observation that wages are rigid while employment varies much over the business cycle constituted a more specific puzzle that spurred the search for alternative models of wage setting. The real wage is, however, only moderately pro-cyclical. These observations cast doubts on the competitive model of wage setting and are consistent with a Walrasian labour market only if two conditions are met. The first is that labour supply must be elastic, because if it is inelastic, large fluctuations in labour demand should lead to large changes in real wages but only to small changes in employment and hence unemployment. For men, there is a consensus that labour supply is very inelastic. Basically, all men that are able to work want to work at prevailing wages implying that the labour force participation rate approaches one. For women, the available empirical evidence suggests that labour supply is more elastic than for men but yet very low. The overall labour supply must therefore be considered to be generally inelastic. Clearly, the first condition is not met.

The second condition for the competitive model to be consistent with unemployment is that shifts in labour supply play an important role in determining employment fluctuations. This, however, seems highly unlikely.

There have, though, been attempts to save the competitive model. One of these was provided by Lucas and Rapping (1969). They made a distinction between the long-run supply of labour, which is inelastic and the short-run supply, which can be argued to be considerably more elastic. In their theory of inter-temporal substitution of labour they argued that faced with a temporary fall in the real wage, workers will choose to work less and quit jobs, thus substituting work today for work in the future when their labour was expected to be better rewarded. However, in assessing the literature on inter-temporal labour supply, Card (1994) casts strong doubts on the plausibility of the size of the response of labour supply to short-run variations in the wage that are needed to explain the data. Thus, it is unbelievable that the unemployment that we observe in recessions is the
result of a voluntary choice by workers. Moreover, Bewley (1999) concludes about the Lucas and Rapping (1969) model from his interview study on American wage setting, that “This theory does not accurately describe labor market behavior during a recession...and counselors of the unemployed knew of almost no one who had quit because of a pay cut”

Nevertheless, while the competitive model, including the version of Lucas and Rapping, quite easily can be rejected as a description of representative segments of the labour market, it still may be a relevant tool in analysing certain segments of a modern labour market, in particular those that are not under the influence of trade unions. In search of a model that has the potential to describe major parts of the labour market, we must consider the existence of the institution that has a dominating influence on wage setting, particularly in Europe, i.e. the trade union.

The rise of trade union models

The monopoly union model

The analysis of the influence of trade unions on wage setting dates back long before the rise of European unemployment in the seventies. Already in the 1940s, Dunlop (1944) discussed possible union objectives like maximisation of total union employment, of average wages, of the wage bill of the employed plus benefit payments of the unemployed. The objective he favoured and strongly argued for was maximisation of the wage bill of the membership. The utility function that unions were assumed to maximise was therefore written as $U = wn$ where $w$ is the wage and $n$ is membership.

Towards the end of the 1960s, the issue of the objectives of trade unions was again discussed. With similarities to the formulation of Dunlop’s, it was suggested that unions are rent maximisers. de Menil (1971) assumed that unions maximise a real wage surplus defined as the difference between the real wage bill in the union sector (as in Dunlop’s formulation) and the wage bill in a perfectly competitive sector. The objective function was then $U = (w - w^c)n$ where $w^c$ is the wage in the competitive sector. Like in Dunlop’s maximand, workers are assumed risk-neutral, i.e. the utility function is assumed linear in wages.

There are today two dominating objective functions in research on trade unions’ wage setting: The utilitarian objective function and the expected utility approach. In a very popular version of the trade union model, the so called monopoly union model, the union sets the wage level unilaterally to maximise utility or expected utility subject to the firm’s labour demand
curve. Undeniably, this is a simplistic view, denying the obvious fact that the firm is an active partner in the wage bargaining process. Nevertheless, this model cannot easily be dismissed in a review since i) it is a special case of the standard bargaining model and ii) it is a very popular version in the theoretical and in the empirical work.

The monopoly union is the version that has been applied in the relatively few existing studies that assess the objectives of trade unions over wages and employment. In this extreme case the union can determine a wage that extracts all surplus while the firm passively determines employment according to this imposed wage. In a pioneering work, Farber (1978a, b) applied the monopoly union model to the US coal industry using annual data for the period 1948-1973.

Farber’s estimate of the relative risk aversion of the US coal miners is significantly different from zero. This implies that the utility function is not linear in wages and hence that the rent maximisation hypothesis must be rejected. The estimated parameter of risk aversion is so high that Farber concludes that the United Mineworkers Union place great emphasis on the employment consequences of their wage policies.

Using a slightly less sophisticated model, Carruth and Oswald (1985) estimate wages and employment in the British coal industry 1950-1980. Again it is assumed that the union, in this case the National Union of Mineworkers, acts as a monopoly union. As was the case for Farber’s study on the US coal workers, Carruth and Oswald find that the coefficient of relative risk aversion is significantly different from zero. Again, the rent maximisation hypothesis is rejected and also British coal miners’ union consider the employment consequences of their wage policies.

The approach used by both Farber and Carruth and Oswald does not, however, imply that employment matters to the union: this is simply a maintained hypothesis. Dertouzos and Pencavel (1981) assume a utility function with wages and employment. Using data for the International Typographical Union in the US 1946-1965 and estimating a two-equation system, they argue carefully that the union’s wage and employment policies reflect the true preferences of the union.

Their model nests the rent maximization and wage-bill hypotheses as special cases. First, both these hypotheses are rejected by their data. Secondly, local unions were obviously concerned with the surplus of wages above some basic level as well as the surplus of employment above some level. For wages the difference between the union outcome and the basic level is represented by $w - b$, and for employment (here identical to membership) by $n - \bar{n}$. There were, though, a considerable variation across the local unions’ weights of these excess wages and employment. In a
follow-up study, Pencavel (1984) split the data into large and small unions. The wage bill hypothesis was rejected in all specifications and the estimates of the elasticity of substitution between wages and employment suggested only limited substitution. For the large unions, the results suggested some support for the rent maximisation hypothesis while smaller unions were more concerned with employment than were the large ones.

Based on these studies, it may be concluded that wages and employment matter to the trade unions analysed in these studies, i.e. coal miners and printing industries. Preferences appear to differ quite considerably, though, between US and British unions.

In a sequel study, Pencavel (1985) attempts to apply the trade union model to Swedish data on mining and manufacturing sectors 1968-1982. While previous studies have focused on the microeconomic level, the purpose is to study wage and employment on the macroeconomic level as the period in focus was one of highly centralized wage setting. Unfortunately, the study is left somewhat unfinished in terms of the empirical application and the extent of wage drift etc. hides much of the true preferences and objectives of trade unions. As an overall characterisation of the empirical literature on trade union objectives, Pencavel concludes that “...our knowledge of the way wages and employment are determined in unionised labor markets is meagre. In view of this, it would seem ill-advised to place much reliance on these models for the purpose of macroeconomic policy evaluation and prescription.”

It should be remembered that the empirical studies are based on the monopoly union model, i.e. the union is assumed to set the wage and the firm accepts this wage and determines employment. Hence, any conflict in terms of union-firm preferences is hidden as there is no bargaining assumed. It implies, for instance, that wage moderation may be attributed a concern for employment (besides wages) while the observed wage moderation may have resulted from employers resistance to the unions’ wage demands.

This line of research, so strongly dependent on the monopoly union model, faded in the eighties and the obvious existence of wages set in union-firm bargaining is, of course, something that cannot be neglected.

**Union-firm bargaining models: The Right-to-manage model or efficient bargaining**

One of the most popular models is the right-to-manage (RTM) model, which assumes union-firm bargaining and retains the assumption that firms determine employment. The union’s and the firm’s gains from reaching a bargain are defined and wages are then determined from maximisation of
the product of the two agents’ gains. With no bargaining power of the firm, the model is simply the monopoly union model. The surplus of the bargaining process is shared according to the parties’ bargaining strength. The wage that comes out of this model is such that the proportional marginal benefits to both of the parties of a unit increase in wages exactly equals the bargaining strength weighted proportional marginal costs to each party.

Much discussion on the right-to-manage model has concerned the fact that its solution is Pareto-inefficient: the union, the firm or both could be made better off by bargaining over employment as well as wages as in the efficient bargaining (EB) model. This spurred a literature on how employment could enter the models, if at all. Bargaining over employment is, however, despite the prediction of efficiency gains, not observed other than in severe crises and large-scale lay-offs. In particular, employers would most surely oppose such a proposition since it implies more bargaining power to the union. It also turned out that the theoretical result is not very robust. For instance if uncertainty is introduced, the contract is no longer incentive compatible.

The fundamental difference between the two models can be seen by writing the predicted marginal product that the EB model generates as

\[ pq' = w - (u(w) - u(b)) / u'(w) \]

(1)

while the RTM model implies only \( pq' = w \). The RTM model implies a solution on the labour demand curve while the EB model, in general, does not. The empirical literature has not been able to discriminate efficiently between the RTM and the EB models. This is dissatisfactory since the predictions of these two models differ widely. The first study that tried to discriminate the two models was Brown and Ashenfelter (1986) who tested if the wage was a function only of the marginal product of labor (as in the right-to-manage model) or if it also was a function of the union’s preference for a solution with higher wages and higher employment as obtainable under efficient bargaining. Using data from the US printing industry, their results were not conclusive.

However, MaCurdy and Pencavel (1986) specified a more general parameterisation of the unions’ preference function and a two-stage procedure. They first estimate the marginal productivity of labour and then the relation between marginal utility of employment and marginal utility of wages where the latter is a function of employment, wages and the alternative wage. Their results favoured the EB model and reject the RTM model but they nevertheless note that this is no proof of the EB model: the
fact that employment does not lie on the labour demand curve does not necessarily imply that it is situated on the contract curve.

In an often cited study, Card (1990) argues that when the wage rate of the firm enters into an employment regression, the OLS estimate of the slope of the demand curve is biased. To overcome the spurious wage-employment correlation, one needs to look at unexpected real wage shocks. Card instruments these with unexpected price shocks. These shocks, represented by indexation, differed across the wage contracts that constituted Card’s panel data. Card found that firms’ labour demand adjusts to wages well in line with the RTM model. Card concludes that “…there is no evidence that employment is related to outside wages in a manner consistent with simple efficient bargaining models”.

Teulings and Hartog (1998) argue that one possible reason for the differences in MaCurdy and Pencavel’s results and those of Card is that the former study a highly homogenous and strong union with strong control of the workplace while the latter studies broad and heterogeneous unions bargaining with large corporations. They also suggest that one reason for why MaCurdy and Pencavel find that employment and wages are determined simultaneously is that they deal with a declining industry, i.e. the printing industry. The unions’ influence on employment would then be stronger than if the industry is growing. The EB model could therefore apply to declining industries, yielding overemployment, and the RTM model to growing industries, yielding underemployment. It should be remembered that these studies are applied to decentralized labour markets rather than to more centralized European type of labour markets.

There is a long list of problems involved in the attempts to discriminate between the models. As noted by MaCurdy and Pencavel, the predictions are sensitive to the particular form of the unions’ objective function. In some utility functions the alternative wage cancels from the efficiency condition so that, in effect, the predicted marginal products of the two models become identical. Moreover, some functional forms of the utility curve and the labour demand curve imply that the alternative wage also can have an impact in the RTM model. Thus an effect of $b$ on the marginal product in (1) could be consistent with the RTM model implying that a variety of functional forms must be investigated. On top of these problems, it is clear that there are other models, like the efficiency wage model and even the perfectly competitive labour market, that generate a solution implying that the wage differs from the marginal product. Lockwood and Manning (1989) show that in a dynamic framework any variable affecting profits and unions’ utility will also have an impact on employment. One can also argue that under special conditions the contract curve and the
labour demand curve are identical. Carruth and Oswald (1987), Oswald (1993) and Booth (1995) give examples when the EB model generates the same outcome as the RTM model.

In view of these problems, it appears as if researchers more or less have given up the attempts to discriminate between the different forms of the trade union models. An argument that speaks against the EB model is that there is little evidence that firms and unions bargain over employment other than, possibly, in recessions with lay-offs as noted by Teulings and Hartog (1998). The Swedish “Medbestämmandelagen” (MBL) gives the unions the right to negotiate with the firm’s representatives about lay-offs. The influence on the actual number of lay-offs is, however, weak and is limited to influence on who are selected to leave and when(?).

To conclude, it seems quite unlikely that the efficient bargaining model should be of great empirical relevance and using this model as a basis for empirical research is not likely to be fruitful, in particular not on more aggregate levels. For a study on wage setting in a particular union one would need to have strong evidence that the union, during bargaining, had a direct influence on employment. The decision on factor intensities is one of efficiency in production and a crucial determinant of profitability and employers are not likely to hand over much influence on such matters to the trade union representatives.

Lessons from trade union models
The history of trade union models is somewhat peculiar. In the 1980:s the trade union models was a very hot research topic. Trade unions represented an obvious labour market institution that to researchers studying most European countries could not be dismissed. There were theoretical developments that resulted in a new view on wage formation.

However, during the 1990:s this research gradually was exhausted for new ideas and researchers appear to have learned these models (monopoly union model, the Right-to-Manage model, efficient bargaining model) in such detail that the flow of research papers faded. Prominent researchers in the field left this branch of labour economics or even labour economics for other types of research. One interpretation was that research had come to a point where the utility of research effort was lower than in other fields, i.e. there was more or less full knowledge about the models.

It is true that some basic models, like the RTM model, the EB model or monopoly union model, had been penetrated in great detail. The basic static models yielded a solution on the labour demand curve with the marginal revenue product of labour equal to the wage. Manning (1994) comments on
the results of the basic static models: In comparing RTM with the competitive model:

i) For a given wage, employment in the RTM firm will be the same as in a competitive firm.

ii) As unions raise wages above the competitive level, employment will be lower in an RTM firm than in an equivalent competitive firm.

In comparing the EB and the RTM models:

iii) For a given wage, employment is higher in the EB than RTM models as long as unions care about employment.

iv) If unions do not care about employment, the RTM and EB models are indistinguishable.

Point iii) above is true in a partial equilibrium model but not necessarily in a general equilibrium model. A final result that Manning comments is an empirical one:

v) A test of the EB against RTM models can be based on testing whether measures of the alternative wage are of use in predicting employment conditional on own wages

In a simple model, in which it is assumed that the ex post substitutability between labour and capital is less than the ex ante substitutability, none of these results are necessarily true! Thus, the results obtained from the basic trade union models are not robust with respect to fairly simple and quite reasonable extensions. An alternative explanation as to why the interest in the bargaining models faded could therefore be that the mission of finding truly robust models was given up.

**Efficiency wage models**

**The rationale for efficiency wage setting**

In discussing why wages could end up on a level higher than the one that is consistent with full employment, the light fell naturally on the behaviour of the trade unions. But unemployment was not limited to unionised sectors and in the late seventies and early eighties some researchers asked
themselves if also rational firms could be interested in anything else than a wage as low as possible. If that was the case, also unemployment in labour markets where firms have a dominating influence on wages could be explained.

The idea that was exploited was quite simple: If the firm raises the wage, the individual will work harder. A very simple idea for why an employer could be interested in raising the wage above the competitive level is if the competitive wage does not allow for an adequate food intake. A higher wage would then make the worker better nourished and more productive. This, of course, is not an argument of relevance to developed countries but only a concrete example of the benefits of higher wages. For developed countries, the arguments for a higher than competitive wage are different and several versions exist. First, a higher wage may make the worker grateful since the raise is considered as fair and this fairness is returned to the employer in the form of a higher effort. (Akerlof (1982).) In another version, the firm has limited monitoring abilities so that they have to provide their workers with an incentive in the form of a higher wage to exert effort. (Shapiro and Stiglitz (1984)) A third version suggests that by paying higher wages a firm may attract a pool of higher-quality applicants for new jobs and thus obtain a better workforce. A fourth version claims that firms pay higher wages so as to reduce costly labour turnover.

**Empirical studies on general efficiency wage effects**

In view of the large number of credible reasons for firms to raise the wage above the market clearing level, we should expect to see this supported by empirical results. The empirical studies can be divided into those that test for any efficiency wage effect, without committing to any particular version of efficiency wage models, and those that test for some specific version. In general, the presence of persistent inter-industry and inter-firm wage differentials that do not appear to be due to compensating differences or to differences in human capital represents some support for efficiency wages. The presence of these differentials has been found by Dickens and Katz (1987), Krueger and Summers (1988) and Groschen (1991). However, as noted by Levine (1992), these studies do not directly test if efficiency wages are the reason for the dispersion of wages.

To perform a more direct test, Levine (1992) tests the fundamental prediction that marginal wage increases raise productivity sufficiently to pay for themselves. Levine finds a positive relation between changes in relative wages and changes in total factor productivity. He also found that the output elasticity of wages was of the magnitude predicted by efficiency
wage theory. Moreover, unionisation seemed to yield a weaker relation between changes in wages and changes in productivity.

In general, however, the results are mixed. Leonard (1987) finds some modest support from California that higher wages reduce turnover. Campbell (1993) finds that firms paying higher wages allow for more selective hirings. However, in both latter studies, the estimated coefficients are lower than those suggested by theory. Ackum (1994) finds some limited support for the US, UK and Sweden. However, she finds no evidence that relative wages are related to the proportion of time spent working while the threat of unemployment does have a significant effect. This latter argument does not, however, support the efficiency wage story since the effort-wage relation is the crucial one.

Edin and Chen (2002) test for general efficiency wage effects in Sweden. Their study represents a quite different empirical approach as they test for efficiency wage theories by means of data on different methods of pay, i.e. piece-rates and time wages. Piece rate workers are compensated for their performance, i.e. for effort and competence. This implies that the shirking and adverse selection models do not apply very well to piece rates. Also the fairness model and labour-turnover models are less applicable to piecework. Hence, Edin and Chen test if there is less efficiency wage setting in piece-work than in time work by testing if industry wage differentials are less prominent for piece-rate compensation. They find mixed support and conclude: “For individuals who are paid under both piece rates and time wages (i.e. mixed rate workers), industry effects exhibit less variability and explain a smaller proportion of overall wage variability for piece-rates pay than for time wages. These findings are consistent with our predictions. In contrast, the industry effects are of roughly equal importance in piece rates and in time wages for individuals who are paid under either method of pay but not both (i.e. pure-rate workers.)”

One implication of efficiency wages is that there should be a positive effect on effort of unemployment. Agell and Lundborg (1995, 2003) find that this is the case in Sweden. This is consistent with the idea of Akerlof (1982), that workers are grateful to have employment in periods of high unemployment and that they “pay back” by high effort. It is also consistent with the Shapiro and Stiglitz (1984) model since the economic penalty of shirking rises. It is also be a consequence of firms laying off the least productive employees during a recession, an idea discussed at length in Bewley (1998). When Agell and Lundborg (2003) investigate this proposition in more detail, however, they do not find support.
Empirical studies on specific efficiency wage models

Shapiro-Stiglitz’s shirking model
An early version of the efficiency wage theory is Shapiro and Stiglitz (1984). The idea here is that firms have limited possibilities of monitoring work effort of their employees, which gives firms incentives to set a higher wage than the competitive one so as to stimulate workers’ effort. The equilibrium wage is set by the firm so that the expected value of discounted lifetime utility from the present moment forward of being employed and exerting effort exactly equals the expected value of discounted lifetime utility of being employed and shirking.

At this wage equilibrium unemployment is derived from rational behaviour and is thus a potential candidate to explain real world unemployment. There are, however, several problems with the model. First, since the model is highly stylised, it is extremely hard to tell what unemployment level it predicts. Nor can it be used to determine how unemployment should vary over the business cycle.

A number of studies have failed to establish a trade-off between wages and supervision and a version of the shirking model presented in Bulow and Summers (1986) cannot explain the differentials of inter-industry wages across countries. Neal (1993) on US data finds no relation between inter-industry differences in the frequency of worker monitoring and inter-industry wage differentials.

In repeated surveys on Swedish employers, Agell and Lundborg (1995, 2003) investigate among other things the relevance of the shirking model for the Swedish labour market. While firms acknowledge that some employees shirk on the job, this does not, however, seem to constitute a major problem. One reason for this may, or course, be that firms go to some length to monitor employees and firms admitted that they monitor workers’ effort. Moreover, the equilibrium in the Shapiro and Stiglitz model is a no-shirking one so monitoring should not necessarily be perceived as a major problem. These findings cannot be used to dispute the model.

However, doubts can be raised based on the penalties imposed on shirkers. The Shapiro-Stiglitz model assumes that firms fire a worker who is detected as a shirker. However, seventy percent of Swedish firms claim that they would never fire a worker and another twenty percent would hardly ever dismiss a worker. Much against the assumptions of the Shapiro-Stiglitz model, harsh economic penalties do not seem to be crucial
in eliminating shirking and we conclude that this model is of little relevance to the Swedish labour market.

Would the shirking model be more relevant to a labour market like the US, where the firm is perceived to have more influence on wage setting? In one of his very detailed interview studies, Bewley (1999) reports that US managers, while interested in the mechanisms of the model, claimed that the model does not describe their own behaviour very well. Rather, the model describes what they called “bad management”. Bewley found some minor evidence of the theory in markets with low-paid temporary workers where wages are downward flexible. However, the shirking model is intended to explain high wages and wage rigidity. Thus, also Bewley, for the US labour market, in general rejects the shirking model. Eighty-seven percent of the firms approached by Bewley claimed that the model “Does not apply” and only four percent claimed that it “Applies”. Also Campbell and Kamlani (1997) find no real evidence in favour of the shirking hypothesis.

Our conclusion is that the Shapiro-Stiglitz model of efficiency wage setting is of limited relevance not only to the Swedish labour market but also to the American one. The fact that interest has been maintained for this model is probably because the model is highly attractive from a theoretical point of view but not at all for the empirical relevance.

**Fair wage models**

Maybe the most convincing argument for why firms would pay efficiency wages is that a high wage builds loyalty and stimulates hard work. A low wage may cause anger and generate shirking or even sabotage. There are studies showing that workers who believe they are underpaid may consciously perform their work in a way that reduces the firm’s profits.

In this model, output is a function of the number of workers employed multiplied by the effort that the worker offers. The firm maximizes profits by determining a wage in efficiency units and the final outcome is one where the wage is set so that the elasticity of effort with respect to the wage is 1. The firm thus wants to hire effective labour as cheaply as possible.

A basic motivation of the fair wage–effort hypothesis derives from human behaviour and simply states that people try to get even when they do not get what they deserve. In a wage contract this would mean that the perceived value of the labour input should equal the perceived value of the remuneration. Psychologists have considered it obvious that agents who feel under-rewarded will supply less effort and several studies also yielded supportive results. Psychologists’ experiments have mainly been directed
towards examining whether overpayment leads to effort increases. Such experiments have, however, yielded ambiguous results which is consistent with the hypothesis that overpayment does not increase effort, i.e. that $e=1$ for $w>w^*$ where $e$ is effort supplied, $w$ the actual wage and $w^*$ the fair wage. Akerlof bases his model on the convincing empirical sociological and psychological evidence of a relation between fair wages and effort, particularly those studies showing that underpayment reduces effort. Since the actual wage may be lower than the fair wage but still be higher than the competitive wage, unemployment will occur in equilibrium.

An implication of the model is that the real wage does not respond to demand shifts. The efficiency wage is determined exclusively by the properties of the effort function, and firms have no reason to adjust their wages. Shifts in demand of the business cycle would thus lead to large movements in employment and to only small changes in real wages. These implications of the models seem easy to reconcile with real world observations of the business cycle. However, the predictions should apply also to the long run and if economic growth shifts demand outwards the real wage remains unchanged and the result is that unemployment should show a downward trend. No such long run trend is observable, however.

Over the years, researchers have retained their interest in Akerlof’s version of the efficiency wage theory. There is a great deal of appealing economic intuition in the basic assumptions, in particular that wages have incentive effects. Moreover, the model seems to stand up reasonably well to empirical tests in the form of econometric applications as well as interview and experiments and below we shall present these at some length.

Since efficiency wage theory assumes that firms set wages, the model may seem to carry more promise for US labour markets than for European ones, including the Swedish. We consider first the available evidence for US economies.

Wadhwani and Wall (1991) is the first empirical study to test for the fair wage approach. Some evidence are found on UK data that productivity at the firm level increases when relative wages rise or the level of unemployment rises. While this provides some support for the efficiency wage models their estimated effort-wage elasticity (.6) is significantly less than the value implied by the Solow equilibrium (1.0). This could suggest that, while efficiency wage considerations are important, some other institution, like wage bargaining, could affect the results.

The possibility that the results are consistent with other theories cannot be ruled out. A possibility is that the link between wages and productivity is explained by unobserved human capital. It could potentially also be explained by the existence of rent-sharing implying that higher productivity
leads to higher wages. But the authors rule out this possibility since the effects of wage and unemployment do not vary systematically with unionism. Finally, they take the result that productivity responds to changes in relative wages as evidence of the sociological versions of the efficiency wage models though such a result could be consistent with other models as well.

One could argue that surveys represent a more promising approach to investigate the existence of fair wage setting. Blinder and Choi (1990) offer survey evidence on the fair wage hypothesis for the US. When managers were asked how a reputation for having an unfair wage policy would affect work effort, 95 percent of the firms said that effort would fall. Blinder and Choi conclude that the notion of fairness is important to the labour market and that there is some agreement among workers and firms about what is fair and what is not fair. They specifically point at the ideas of Akerlof (1982) and Okun (1981) as highly relevant.

In another interview study, Campbell and Kamlnai (1997) find that firms fear that a wage cut would reduce workers’ effort. In line with Bewley (1999) they report that firms find that wages have a strong effect on effort as workers’ attitudes towards their employers are affected by the wage. It is, however, rather a wage cut than a low wage level per se that firms connect with low morale.

Agell and Lundborg (2003) find that the percentage of managers that rank high wages as the most important factor in motivating their employees, is considerably lower in Sweden than in the US. The report that only 6 % of Swedish managers rank high wages at the top while Campbell and Kamlnai find that 24 % of US managers top rank high wages. Agell and Lundborg conclude that this difference may be due to the considerably lower level of individualistic wage bargaining in Sweden. However, much has happened in this respect since the 1990s, an issue to which we return at the end of this paper.

**Insider-outsider models**

Assar Lindbeck and Dennis Snower are the fathers of the insider-outsider theory and, like most other wage setting models of the eighties, this model was constructed to yield a microeconomic foundation of unemployment. As such it focused on explaining the absence of wage underbidding even when there existed workers willing to work for lower wages.

In reviewing the empirical insider-outsider literature, Lindbeck and Snower (2001) find pieces of evidence in favour of the theory. However, as they note, much of the evidence (if not all) is also consistent with other
labour market theories. Studies with empirical evidence that have been successful in discriminating between insider-outsider models and other labour market theories simply does not seem to exist. Lindbeck and Snower (2001) indicate how such empirical evidence could be obtained.

There are, however, survey studies that have asked wage setters in firms to what extent their wage setting is influenced by insider-outsider mechanisms. Not much support has been found for the U.S. Bewley (1999), p. 403, e.g. bluntly states: “Insider-outsider theories do not correspond to what I observed.” Campbell and Kamlani (1997) who found that “...insider-outsider theory received weak support...” interpreted their results as if the theory does not apply to low unionisation economies. Agell and Lundborg (1995) asked wage setters in Swedish firms about the labour turnover version of the insider-outsider model they found no enthusiasm for the crucial mechanism of the model. Firms with a large share of white-collar workers had a slightly more positive attitude than those with a small share. This is as expected to the extent that it is more costly to hire and train white-collar workers.

**Other wage setting models: search theory and tournament theory**

We shall not go into details about the empirical support for search theory. Like the bargaining models, search models can hardly be discarded simply because they capture aspects of the labour market that are obvious facts of life. There is search and there is matching in the labour market. Thus, it suffices here to recognise the relevance of these models and to note that they continue to be useful analytical tools for the labour market. However, these models are not universally applicable to all types of issues, but rather highly useful for analysing dynamic issues where search and matching takes time.

Tournament theory applies to a limited part of the labour market, i.e. salaries of managers and as such we only refer to one relevant study for Sweden. Heyman (2005) test for several predictions of tournament theory using managers as well as an extended data set to white collar workers. Some results are supportive of tournament theory. For instance, a higher wage spread is found in firms operating in volatile product markets, i.e. where output uncertainty is high.
3 Institutional changes in Sweden and wage setting: A discussion

In this final section we shall discuss the empirical support for the different theories against the major institutional changes that have taken place in Sweden during the last thirty years. There is a general European trend that the degree of coordination in wage bargaining decreases. Visser (2000) has calculated levels of wage co-ordination in fifteen European countries. Measured on a scale from 0, which represents no bargaining above firm level, to 1, which represents complete coordination, the degree of wage coordination in Sweden has decreased from .745 during the period 1973-1977 to .490 during 1983-87 and to .389 in 1993-97. The decline in Sweden is one of the strongest in Europe. During the period 1973 to 1977 only Austria had a higher level of coordination than Sweden, while in the period 1993 to 1997 there were six countries with a higher degree of centralisation than Sweden: Austria, Ireland, Belgium, Finland, Netherlands and Norway.

Centralised wage setting implies strong trade unions, which could encourage union membership. Since there has been a trend away from centralised wage setting, for which employers have been an active force, one could suspect that this would lead to a weakening of unionisation in Sweden. One could also expect overall unionisation to fall also because of the secular decline in the number of blue-collar workers. However, despite these changes total union membership increased until the mid-nineties to around 3.9 million members (1995). As seen in Figure 1 there has been a decline in overall membership after the peak in the mid nineties and total union membership has fallen to about 3.7 million (2002). Also as a share of the labour force, overall unionisation has dropped slightly in recent years as seen in Figure 2. The fall in union density is modest in Sweden in comparison with other countries.

Figure 1 also shows clear trends in the structure of unionisation as measured by the relative shares of the different confederations of unions. Since the mid eighties, LO has gradually lost its share of overall membership. During the 1970:s and 1980:s, TCO gained the most in membership while during the 1990:s membership in SACO increased at high rates. Between 1990 and 2002, SACO increased the share of total union members in Sweden from 8.6 to 14.3 percent while LO:s share fell from 58.1 to 51.4 percent.
These changes in the structure of unionisation mean that it becomes more and more difficult for a confederation with egalitarian ambitions to affect the overall earnings structure of the economy. As will be discussed below, wage differences have increased in Sweden ever since the peak of egalitarianism in the early 1980s. *Ceteris paribus*, to keep up the egalitarian welfare state, this development should imply that the tax and transfer system must be used to a higher degree to compensate for the increase in wage inequality.
To sum up, while the coverage of collective agreements and unionisation remain high in Sweden, we have nevertheless experienced a clear trend away from coordination in wage bargaining, at least from the mid seventies to the mid nineties. This indicates that though individual wage bargaining has increased, at the same time, there remain supporting collective agreements for most employees. The wage increases of the individual are set locally, while the central union determines a norm for the minimum wage increase.

Thus, Swedish wage setting has changed quite dramatically, yielding a considerably more diversified picture of wage formation. In the Sixties and Seventies, little scope existed for individual wage adjustments other than in the form of wage drift. It appears that different versions of a standard wage bargaining model, like the RTM model, would be a fairly adequate description of how bargained wages were established in those days.

The fact that the level of wage drift was very high, particularly during the 1970s and 1980s, implies that the bargaining models hardly can be used to determine final wages at the aggregate level. These wage increases, on top of the agreed central wages, were much the result of the fact that full employment was a leading objective of economic policy of the day. This stimulated wage increases also at the firm level, which added to the centrally bargained wages. An implication is that the standard wage bargaining models could be used mainly to determine a minimum wage increase from one contract period to another but not the final outcome of the wage level.

With the continued high union density and high coverage of collective agreements in Sweden, wage bargaining models continue to be indispensable tools. One could also argue that the high unemployment rates since the early 1990:s lowered the level of wage drift which in turn would enhance the importance of the bargaining outcome. On the other hand, there are strong forces against this development, particularly the fact that wage setting has been decentralised and individualised. For an analysis of the effects of individualistic wage setting in Sweden since the mid 1990, see Lundborg (2005).

Wage setting at the level of the individual worker and the firm, rather than at the centralised level, calls for a revision of the view of the appropriate model. Not only does it imply that models of wage bargaining should be revised so as to reduce or eliminate the considerations of aggregate unemployment in an appropriate model of wage setting. More importantly, however, is that, today, firms at the local level can affect the wage of the individual employee to a much larger extent than what used to be the case. The individualisation has in many unions been manifested in a
large increase in wage differentials. This suggests that efficiency wage models could be of much more relevance today compared to ten years ago. Our review above reveals empirical support among firms’ wage setters for the fair wage version of efficiency wage setting. While Agell and Lundborg (2003) found that in 1998 high wages was a major determinant of employees’ motivation in smaller share of firms than in the US, they concluded that this difference was due to less individualistic wage setting in Sweden. However, since this has increased in Sweden, and wage differences has increased as well, fair wage setting should be expected to be highly important also in Sweden. That fair wage setting has been an important issue during the process to individual wage setting is supported in a recent study by Lundborg (2005).

The decentralisation and individualisation of Swedish wage setting has undoubtedly raised the scope for efficiency wage setting in Sweden, particularly among white collar workers. Similar changes have not taken place among blue collar workers for which firms can be assumed not to have the same scope for efficiency wage setting.

Since unionisation remains high, a combination of the fair wage model and the bargaining model appears to hold promise as a useful approach to analysing current wage setting in Sweden and to explain why wages in many cases end up above market clearing levels. Indeed, much of the changes towards decentralised wage setting have been motivated by employers’ need for wage differences that should encourage the individual employees to invest in skills.
References

Calmfors, L. and K. Richardson (2003);”Marknadskrafierna och lönebildningen i landsting och regioner”, Landstingsförbundet.


Friedman, M. (1968);”The Role of Monetary Policy,” American Economic Review, 58, 1-17

Garino, G. and C. Martin (1999);”Efficiency Wages and Union-Firm Bargaining.” (mimeo), Brunei university, July


Working Paper Series/Arbetsrapport

FIEF Working Paper Series was initiated in 1985. A complete list is available from FIEF upon request. Information about the series is also available at our website: http://www.fief.se/Publications/WP.html.

2001


2002


187. Lundborg, Per and Rechea, Calin, “Will Transition Countries Benefit or Lose from the Brain Drain?”, 18 pp.

2003


2004


2005


203. **Skogman Thoursie, Peter**, "Happy Birthday! You are Insured. Differences in Work Ethics Between Female and Male Workers”, 8 pp.


