No. 248, 1989

WAGES AND LABOUR SCARCITY
— The microfoundations of the determination of factor shares

by

Nils Henrik Schager

Nils Henrik Schager

WAGES AND LABOUR SCARCITY – THE MICROFOUNDATIONS OF THE DETERMINATION OF FACTOR SHARES

Abstract

It was early noted that the Phillips Curve explanation of wage dynamics lacks a solid microeconomic basis. As the explanatory unemployment variable in the Phillips relation is intuitively to be regarded as an indicator of labour scarcity, several authors have argued that the determination of factor shares (the wage/profit ratio) ought to be the logical intermediate step between unemployment and wage inflation. Contributions by Kuh, Solow and Stiglitz in the late sixties follow this line of thought.

As soon as wage–setting by firms instead of the impersonal forces of excess demand has been considered, problems have arisen, however. Given the ubiquity of diminishing returns to labour in the short run in static economic models, it is hard to derive the desired result that the wage share increases when unemployment falls. Monopolistic price–setting among firms can only yield such a result, if arbitrary rules for mark-up pricing are enforced. Even more disappointing, the temporarily monopsonistic firm of labour market search models has been shown to increase its (real) wage, when it faces larger flows of unemployed job applicants.

The aim of this paper is to show that search theory easily yields the result that factor shares turn in favour of wages, when unemployment falls. The only important prerequisite for this result is that money wages are treated as downward rigid. The analysis remains valid when firms experience constant returns to labour in the short run (up to a capacity limit or to a demand constraint). In fact, my analysis predicts that with such a simple, but presumably relevant, production technology, factor shares are determined by the stock of unfilled vacancies, which in turn is partly determined by the unemployment stock. As intuition says the share of wages in value added increases with the stock of vacancies and decreases with unemployment.

Nils Henrik Schager

Wages and Labour Scarcity – the Microfoundations of the Determination of Factor Shares

Introduction

It is convenient to start off the discussion of this paper by quoting the introductory lines of A. W. Phillips famous article in Economica: "When the demand for a commodity or service is high relative to the supply of it we expect the price to rise, the rate of rise being greater the greater the excess demand. It seems plausible that this principle should operate as one of the factors determining the rate of change of money wage rates, which are the price of labour services." (Phillips (1958)).

As it is stated Phillips argument appears plausible indeed. There are, however, unanswered questions as to the mechanism by which excess demand for labour gives rise to wage increases. First one has to recognize that the ultimate effect of the wage increase should be to eliminate the excess demand. The scarcity of labour is eventually to be reflected not in high wage increases but in a high wage (relative to the renumeration of other factors of production). We may of course believe that for example high wages lead to high demand for commodities, so that wage increases produce excess demand for products. Hence attempts to increase production make excess demand for labour reappear, so wages and prices increase in a perpetual sequence, if uninterrupted by a permissive monetary system. Nevertheless, the underlying theoretical model must conceptually be able to say something sensible about the wage level in short run equilibrium when demand in the product market is stabilised.

Second the notion of excess demand as the difference between the static demand and supply curves, which was an accepted device at the time Phillips wrote his article, is of course an artefact. It is observed by no one but the fictitious Walrasian auctioneer, who by assumption adjusts prices as to eliminate any discrepancy between demand and supply. As soon as price and wage

decisions are based on the optimal behaviour of individual agents, the simple excess demand (or supply) concept evaporates. It has turned out to be extremely difficult to reestablish a simple counterpart to it, let alone a clear-cut relation between any modified excess demand concept and the price dynamics.

The present paper is an attempt to clarify these two obscure points with special emphasis on wage determination and labour market phenomena; prices and the product market are treated more simplistically.

Unemployment, real wages and factor shares in static equilibrium

In this section I will adress the problem of how the wage level is determined in short run equilibrium. Such an analysis should logically precede any modelling of wage dynamics.

If unemployment reflects, in an inverse way, scarcity of labour, one would like to have it that wages are high at low rates of unemployment. Of course, scarcity must be reflected in relative, not absolute prices, so it is some kind of real wage that is relevant. As scarcity of labour can only be a meaningful concept in relation to labour demand, i.e. to the needs of employers, the relevant wage concept should be a product real wage, i.e. the money wage should in the first place be deflated by the relevant product price.

Labour productivity must not be forgotten, however, not even in the short run. Although different ideas about the effect of technical progress on factor productivity can blur the picture, the essential point is that a labour productivity increase makes labour more valuable to the employers at given money wages and product prices. Formally the most straightforward way is to assume that labour productivity increases are brought about by labour—augmenting technical progress, so that employers only need to care about the 'efficiency units' of labour they buy. Hence a high wage ought to imply a high wage in relation to the product value added, not only to the product price. In other words, when labour is scarce, the share of wages in value added should be high (and the profit share low).

The argument is consequently that a theory of the functional distribution of income should be discernible behind the dynamics of the Phillips curve. This idea is certainly not new. The contribution by Kuh (1967) follows essentially this line of thought. Kuh, as well as many others, has, however, encountered a major stumbling block in the seemingly unavoidable neoclassical assumption of diminishing returns to labour in the short run.

Low unemployment means that employment is high, which, in the presence of diminishing returns, implies a lower marginal product of labour. If the firms are on their static demand curves for labour, as optimal behaviour seems to require, the money wage must equal the marginal value product of labour. Hence low unemployment is accompanied with a low real product wage as well as with a low wage share, unless labour demand is highly wage elastic in the short run.

Kuh (1967) circumvents this problem by postulating at first an "almost" neoclassical demand function for labour, which nevertheless he later "re-interprets" as to yield a stable labour–output ratio, a rigid real wage and hence stable factor shares. Unemployment is allowed to cause deviations from this short run wage/price equilibrium, but only by affecting the relative bargaining strength between unions and employers, an idea which has gained widespread popularity in recent years. Kuh's claim that labour market conditions cause deviations from an otherwise stable equilibrium distribution of factor shares appears attractive, but his way of substantiating it obviously violates the neoclassical conditions of optimal behaviour to a disturbing degree.

Almost at the same time as Kuh wrote his article, Solow and Stiglitz (1968) addressed explicity the connection between effective demand and the distribution of factor income. Concentrating on those aspects of their article which are of main interest for this paper, I leave alone their discussion on the dependence of effective demand on factor shares as well as their modelling of the ultra-short-run dynamics of wages and prices (which is in the tradition of excess demand analysis and, in fact, goes back to Hansen (1951)).

Solow and Stiglitz conclude that when the firms are on their neoclassical demand curves for labour (i.e. firms are not demand constrained) the short run equilibrium relation between demand and factor shares is the conventional neoclassical one. Higher demand can only lead to higher production and employment, if the product real wage falls (or rises less than any improvement in labour productivity). The effect on the wage share depends as usual on the wage elasticity of labour demand.

When firms are demand constrained, the marginal product of labour is only a notional concept at production volumes above the constraint. It may be, Solow and Stiglitz argue, that there can exist a mark up of prices over marginal cost in such a situation. If this mark up decreases when demand expands, higher employment may be accompanied with a higher real product wage and a higher wage share. This is just a possibility, however, and it is unclear how a mark up is to be maintained in the first place, when there is excess supply in the product market. Reference to the existence of monopolistic competition offers no real solution to the problem, the authors remark.

Turning to more recent contributions such as Layard and Nickell (1986) (see also Nickell (1986)), we find that received theory has still the same tentative attitude towards the relation between unemployment on one hand and product real wages and wage shares on the other. True, the effect of unemployment on the money wage level through the bargaining behaviour of unions is more elaborated and the importance of price surprises is stressed in the fashion of modern macroeconomics. The behaviour of the real wage and of wage shares remains obscure, however, and its significance is played down.

According to Layard and Nickell (1986) it is the pricing rules followed by firms, which determine the relationship between the product real wage and unemployment, as already Solow and Stiglitz claimed. Relying on the neoclassical demand curve for labour, Layard and Nickell repeat the traditional results, when the pure competition pricing mechanism applies. They find it more plausible that monopolistic pricing prevails, however. More specifically they consider the case of "normal cost" pricing, where the mark up of prices over wages is fixed and hence diminishes over marginal costs, when production and employment increase as to exactly offset the effects of diminishing

returns to labour. Hence the product real wage remains fixed (by assumption) and the wage share increases when employment expands and unemployment falls.

The problem is again that normal cost pricing is optimal only under very special assumptions as regards the relation between the level and the price elasticity of product demand. Neither theory nor empirical evidence point to the existence of such a stable relation. It is understandable that economists tend to find this approach increasingly unrewarding.

In a following section of this paper I intend to show that the failure of the standard models to say something definite and convincing about real wages, wage shares and unemployment does not depend on some undiscovered trick as to details. Static neoclassical models are simply not well suited for dealing with the phenomenon of labour scarcity and its effect on real wages and wage shares. As a natural bridge to that discussion I now turn to the basic subject of wage determination, when the conditions of perfect competitive equilibrium do not hold.

Wage setting outside static equilibrium

In a much quoted contribution Arrow (1959) criticised price theory for lacking a convincing theory of how prices are set outside equilibrium. Agents cannot take prices as given in such a situation but have to realize that their own price decisions influence the amount they can buy or sell. In other words, Arrow sought for the micro-foundations of the price dynamics in situations of excess demand or supply in the market. Clearly, Arrow's criticism is of immediate relevance for the original interpretation of the Phillips curve as reflecting the wage dynamics of a labour market in disequilibrium.

When quantity and price decisions are taken simultaneously by agents, we are in the realm of monopolistic or monopsonistic behaviour. We have encountered models of monopolistic behaviour in the product market in the contributions referred to in the preceding section. In a way these models meet Arrow's criticism, but they are of the static kind which requires that

products are imperfect substitutes or, else, are based on an arbitrary partitioning of the market among firms. The first more thorough response to Arrow's challenge was the development of models of dynamic monopolistic competition in homogeneous markets with imperfect information, i.e. the models of search. The contributions in Phelps (1970a) constitute the standard references.

The Phillips curve issue was immediately addressed in several of the contributions in Phelps (1970a), notably Mortensen (1970). Mortensen developed an advanced model of the optimal dynamic wage policy of a firm, which has a temporary monopsonistic position in a search labour market (but is price−taker in the product market). As is common with such complex models, the properties of the optimal policy are contained in a functional equation, which makes them not too easily accessable. Nevertheless, Mortensen was able to establish that the firm's optimal wage is increasing in unemployment (i.e. in the intensity of the flow of unemployed job applicants to the firm). This result runs counter to the intuition of the Phillips curve, that firms should offer a lower wage, when unemployment is higher (as stated e.g. by Phelps himself in Phelps (1970b)). Consequently search theory does not seem to predict a high (real) wage when labour appears more scarce to firms (as reflected in lower unemployment).

It should be noted that Mortensen's result depends in no way on the assumption that the marginal value product of labour is decreasing in employment. In contrast to the static models of the preceding section, the effect of unemployment in search models is not just a reflection of the aggregate effect of employment. Even if the marginal product is constant, the optimal wage of the firm increases in unemployment.

This result of rigorous search theory, when applied to the wage behaviour of firms, has had important consequences. First, the search theory approach to Phillips curve dynamics has focused on labour supply phenomena. More specifically, the wage-unemployment dynamics have been regarded as reflecting labour supply responses to unexpected and temporarily misunderstood price or wage changes, the causal chain running from such

changes to unemployment (of course, assumptions on the formation of expectations are amendments to, not part of search theory per se). Second, while search theory has been amply applied to the analysis of job search, including the determination of job applicants' reservation wages, the search-cum-wage behaviour of firms has since Mortensen (1970) been addressed sparsely and with less degree of technical sophistication. As a consequence the search theory approach to labour market analysis has in recent contributions (Mortensen (1988), Pissarides (1985, 1987)) been restricted to deal with the "technology" of contacts between job searchers and job vacancies, leaving the determination of wages to other approaches such as bargaining models (see below).

The dominant professional answer to Arrow's question, as far as the labour market is concerned, has in recent years been that unions set wages, possibly in negotiations with employers. The idea that the bargaining strength of unions influences the wage level, at least in money terms, is of course an old one. We have encountered it above as an ad hoc amendment to Kuh's model. The novel feature consists of formalising union behaviour by deriving it from a union utility function, according to which the union balances off the benefits from employment against the benefits from (real) wages. The important relationships are easily demonstrated in the model specification, where the union determines the wage and the firm determines employment according to its traditional downward sloping demand curve for labour. A higher real wage means less jobs at the firm, but if the unemployment rate in the market is low, the disutility of a worker being unemployed is less (because of more job opportunities and shorter expected duration of the unemployment spell). Hence lower unemployment causes the union to settle for a higher (real) wage. This result is easily extended to the case, in which increases in labour productivity occur.

The union model of wage determination manages to break the link between low unemployment and low real wages, but only at the cost of being partial. It may give a satisfactory account of what happens at an individual firm, but if all employment is generated by profit-maximizing firms its predicted relation between unemployment and real wages cannot hold in the aggregate. Then unemployment is always the complement to such employment as is

consistent with the neoclassical demand curve for labour, and such employment must be decreasing in the real wage. On the other hand, the proposed mechanism may well serve as a building block in a model of wage and price inflation, if firms reduce continuously the real wage by their mark up of prices, as Layard and Nickell (1986) have it (inflation caused by 'a battle of mark ups').

As a recent alternative to wage setting through unions, some authors have proposed wage setting by firms according to the notion of "efficiency wages". The general idea behind this notion is that firms can influence the productivity of its employees by offering a wage above the market equilibrium level. In one version of those models firms set wages as to establish a premium above the value to the employee of being unemployed. Such a premium is necessary for making the employee abstain from shirking granted that he is fired when caught shirking. As the relative value of being unemployed decreases with the expected duration of unemployment, higher unemployment requires a lower wage premium. The total effect of unemployment on the real wage is thus the outcome of an increasing impact on the equilibrium wage and decreasing impact on the wage premium.

Mortensen (1988) analyses an efficiency wage model of the shirking version in an advanced labour market equilibrium setting, in which he has search and matching between the unemployed and job vacancies. He shows that in the presence of diminishing returns to labour in the aggregate (but not at the firm level) lower unemployment is connected with a higher (real) wage. Mortensen's result is, however, the outcome of the interaction of several assumptions as to the production technology, the matching technology in the labour market, the vacancy creation mehanism etc, which makes it difficult to discern in detail the connection between assumptions and results. More crucial for assessing the importance of the derived relation between (real) wages and unemployment is of course what faith one has in the relevance of the supposed disciplinary wage premium, that characterises the shirking model.

In the present context it is important to realise that although both the bargaining model and the efficiency wage model discuss the effects of

'unemployment' in general, they really focus on the effects of the duration of unemployment. A decline in the expected duration of unemployment increases the opportunity cost to an employee of his present job and he or his union forces the employer to give compensation through a higher wage. The emphasis on the duration of unemployment or its inverse, the probability of a worker's transition from unemployment to employment, is a search theory constituent of both models. As a matter of fact, it is surprisingly straightforward to demonstrate the effect of labour market duration variables on the wage share in a simple search—cum—bargaining setting, such as is used as the basic model in Pissarides (1987).

A search—cum—bargaining model of factor shares

In Pissarides (1987) the matching between an unemployed worker and a vacant job is isolated from any other such matching. Neither the value added produced nor the wage paid is influenced by the existence of other workers being employed by the firm. The wage is determined by a Nash bargain between the worker and the firm, so that the surplus of the match is divided by the worker and the firm in proportions β and $(1-\beta)$.

Pissarides's further analysis is strongly influenced by his assumption that vacancies are created up to a point, where the marginal value of a vacancy is equal to the direct costs of announcing the vacancy. I agree with Blanchard and Diamond (1989), that this assumption is not appropriate for an analysis of the short run, in which the number of vacancies are more likely to be given to the firm as the number of potentially productive jobs. Hence the value of a vacancy is equal to the value of a filled job multiplied by the discounted expected duration of the vacancy, minus the discounted value of any direct vacancy cost. Applying this relation and assuming that direct vacancy costs are negligeable, but otherwise retaining the assumptions as well as the notation in Pissarides (1987), we get

The surplus of a match to the firm:

$$J - V = \frac{y - w}{p(\theta)/\theta + s + r}$$

The surplus of a match to the worker:

$$W - U = \frac{w - b}{p(\theta) + s + r}$$

J	=	the value to the firm of a filled job
V	=	the value to the firm of a vacant job
W	=	the value to the worker of being employed
U	=	the value to the worker of being unemployed
y	=	value added produced at the job
w	=	wage paid at the job
b	=	worker's evaluation of leisure
$p(\theta)$	=	the transition probability from unemployment to employment
$p(\theta)/\theta$	=	the transition probability from vacant to filled job
s	=	the transition probability from filled to vacant job and from employment to unemployment
r	=	the discount rate

Hence the Nash bargain solution for the wage w immediately yields the following equation for the functional distribution of income

$$\frac{w - b}{y - w} = \frac{\beta}{1 - \beta} \cdot \frac{p(\theta) + s + r}{p(\theta)/\theta + s + r} \tag{1}$$

$p(\theta)$ is the inverse of the expected duration of unemployment and $p(\theta)/\theta$ is the inverse of the expected duration of a vacancy. Hence shorter unemployment durations and/or longer vacancy durations turn the functional distribution of income in favour of wages and vice versa. Given reasonable assumptions on the matching technology between the unemployed and the vacancies, the expected duration of unemployment is non–decreasing in the stock of unemployment and non–increasing in the stock of vacancies. Corresponding relations hold for the duration of vacancies. So lower

unemployment is associated with a higher share of wages in value added through its effects on the durations of vacancies and unemployment.

(1) highlights the excellent possibilities of search theory to represent labour scarcity – and, for that matter, job scarcity – and its consequences. It is able to do so, because it focuses on the properties of the dynamic processes in the labour market, not on the equilibrium, at which the labour market eventually settles, if left undisturbed for a long time. More specifically, the existence of unfilled vacancies signifies that the firm is not at its static demand curve for labour. The condition that the wage should equal the marginal value product of labour just characterises the locus of possible wage – employment equilibrium points. As I will show in the next section, the position of this locus is but one of the determinants of the wage–employment outcome, which depends also on how far away the firm is from the equilibrium locus and how long time it takes the firm to get there. The notion of labour scarcity is contained in these latter entities.

The model, from which (1) is derived, gives just a rudimentary picture of the implications of search theory, however. It states quite convincingly that it is important for the firm to get a profitable job vacancy filled quickly as well as for worker to get a job quickly and that both the firm and the worker are prepared to pay for such a benefit. The Nash bargain approach, however, does not produce any reason why they should pay, as the wage is not allowed to influence either the unemployment or the vacancy duration. The wage decision is in fact not integrated into the dynamic processes of the labour market. As the model of Mortensen (1970) accomplishes such an integration, but does not produce results compatible with (1), the suspicion must remain that the qualitative relations given by (1) are misleading special results of over–simplified assumptions.

Fortunately, this suspicion can be dissolved. In the next section I will report results from an analysis of mine, which integrates the wage decision of the firm into the vacancy process along the lines in Mortensen (1970). If a plausible restriction is laid on the set of feasible dynamic wage policies, the firm will respond to changes in the duration of vacancies in a way that is qualitatively consistent with (1). Hence it seems that search theory is not

only capable of giving the notion of labour scarcity an operative meaning, but also able to show how optimal responses by firms to higher degrees of labour scarcity increases the real wage and the wage share, as economists' intuition has said but their theory has refused to confirm.

A search—cum—wage model of factor shares

In this secton I will present and comment on parts of the analysis and results given in Schager (1988b). Those readers who want to study the technicalities of my analysis are referred to that paper.

The model of this section combines several of those features which characterise the approach in Kuh (1967) with the methods of search theory along the lines in Mortensen (1970). The production technology of the firm is characterised by fixed coefficients, so production takes place at constant returns to labour up to a capacity limit. The model is short run, so labour is the only variable factor of production. The product market is represented by an exogeneous product price, which the firm takes as given. Hence only the labour market is explicitly considered. Most important is the assumption that the firm regards it as impossible to lower its wage in money terms.

The labour market is characterised by search, the aim of which is to match job applicants and job vacancies. Job applicants are assumed to be distributed over reservation wages in a non-degenerate way, as a response to a non—degenerate distribution of wage offers over job vacancies, which in turn reflects that firms have different values added associated with their vacancies as well as different vacancy/employment ratios. At present, however, the mutual consistency of the two wage distributions as to reflect strict optimality conditions is not ascertained. It is the theory of the firm part of the model which is rigorously worked out.

It is an important feature of search models of the firm that they study the dynamic path of adjustment of the firm's employment without assuming that the firm is always — or ever — at the locus where price equals marginal labour cost (or on its static demand curve for labour, if it exists). Such a locus defines instead the equilibrium states of the adjustment process. The fact that

the firm is not in equilibrium but announces vacancies in order to get there is important for the possibility to conceptualise (one part of) labour scarcity.

Standard search models of the Mortensen type, in which the reservation wage distribution is non-degenerate, equip the firm with a monopsony power in the labour market, which it uses to regulate the speed of its employment adjustment process. A higher wage, say, means more rapid net recruitment and a higher rate of elimination of unfilled vacancies. It is consequently intuitively appealing that a higher vacancy ratio (in relation to the stock of initial employment), signifying a higher degree of unsatisfied labour demand, causes the firm to increase its optimal wage. Since the model has price and productivity as given, the real product wage and the wage share of value added are raised, too.

Unfortunately, the standard search model produces other less appealing results. As wages are completely flexible, firms decrease their wages as vacancies are successfully filled and employment expands. This is a counterfactual intermediate result, but one could have hoped that it should be compensated by the effect of lower unemployment in the aggregate as employment expands. As I noted above, however, that is not so. A lower unemployment causes the firm to lower its optimal wage in standard search models. As the mechanism works through the contact intensity of job applicants to the firm, so that lower unemployment means prolonged durations of vacancies, this result has indeed been discouraging for anyone who believes in a positive relation between labour scarcity and (real) wages.

What was not discovered – and not easy to suspect, either – is that the effect of unemployment on the optimal wage in standard search models is crucially dependent on the assumption that money wages are completely flexible. If that assumption is replaced by the assumption that the money wage of the firm is downward rigid, the possibility occurs that the optimal wage is decreasing (or, rather, non-increasing) in the contact intensity and hence in unemployment. A sufficient condition for this result to hold is that the firm's vacancies are in expectation filled before the planning horizon is reached at the initially given wage level (see Schager (1988b)).

As far as empirical evidence tells vacancy durations are much shorter than any reasonable assessment of the planning horizon of a business firm. The empirical relevance of downward money wage rigidity is beyond dispute, regardless of economists' frustration at their failure to give a really convincing theoretical explanation of its existence. Hence I find the result reported in the preceding paragraph highly relevant.

Let me now present an equation, which tells how the optimal wage is determined under the assumptions given in the beginning of this section. I want to add that the assumption of constant returns to labour is not necessary for the analysis to carry through. It is adhered to partly because it seems to reflect fairly well the production technology firms usually experience in the short run, partly because it brings out the essential points of my analysis more clearly, and, finally, partly because it shows that diminishing returns are not needed in order to get a determinate result. Furthermore I present the version in which a Holt technology of vacancy-unemployment matching applies ("increasing returns"). Finally, I use the condition that the expected horizon is much longer than the expected duration of vacancies as to approximate the equation. These two last assumptions are made in the interest of simplicity and Schager (1988b) gives the full details. Hence the optimal wage policy of the firm is a constant wage level w* such that

$$w^* = \frac{y \cdot h \cdot V_r}{L_r + h \cdot V_r}$$

if an interior solution obtains, where

w^* =   optimal wage.

y =   value added per worker in current prices.

L_r =   employment plus new vacancies (labour demand) as a ratio of employment.

V_r = unfilled vacancies as a ratio of employment = new vacancies multiplied by the expected duration of vacancies (as a ratio of employment).

h = the discount intensity (the inverse of the expected horizon) multiplied by the wage elasticity of the hiring intensity, the hiring intensity being the inverse of the expected duration of vacancies.

Both V_r and b depend on w^* through the wage dependence of the hiring intensity. If we allow the firm to be the representative firm of the aggregate, however, and assume that shifts in the reservation wage distribution follow closely those in the wage offer distribution (see the discussion in Schager (1988a)), the equation above with V_r and h interpreted as parameters characterises the optimal wage of the representative firm in short run equilibrium. This equilibrium is characterised by stable (or at least only randomly shifting) wage offer and reservation wage distributions. Firms are not on their static demand curve for labour, however, but have job vacancies.

Some rearranging of the equation yields immediately the functional distribution of income as profits per worker are equal to $y - w^*$

$$\frac{w^*}{y - w^*} = \frac{h \cdot V_r}{L_r} \qquad (2)$$

So the model yields the extremely simple result that in short run equilibrium the wage/profit ratio is approximately linearly increasing in the ratio of unfilled vacancies to total labour demand. Unfilled vacancies are the product of announced vacancies (signifying how much firms want to expand) and the duration of vacancies (signifying how difficult it is to expand), so they reflect the degree of labour scarcity to the firms. As the duration of vacancies is monotonically decreasing in unemployment the wage share is decreasing in unemployment

Of course, equation (2) is too simple to be expected hold literally. Matching in the labour market may follow a more complicated pattern than indicated by the Holt 'technology'. Labour supply is presumably not as passive a

response to labour demand, as the aggregate interpretation of the model has it. More specifically, the position of the reservation wage distribution relative to the wage offer distribution should be affected by changes in the expected duration of unemployment, reflecting the symmetric effect of job availability (or scarcity) on the decisions of job applicants. This would enhance the quantitative effect of the stock of unfilled vacancies compared to (2), particularly so if a Holt matching technology applies, where the inverse of the duration of unemployment increases linearly in the vacancy stock. Finally, if there exist diminishing returns to labour in the short run, account must by taken of the fact that the number of new vacancies announced by the firm is dependent on its wage decision. In that case not only the duration but also the inflow of vacancies is a decreasing function of the wage.

Now it is clear that the qualitative properties of (2) and eq. (1) of the preceding section are the same, especially if we take into account that the stock of unfilled vacancies has a stronger impact than (2) suggests through its effect on the reservation wage decisions of job applicants. Applying the conditions of a Holt technology in matching to (1) turns it into

$$\frac{w - b}{y - w} = \frac{\beta}{1 - \beta} \cdot \frac{V_a(k \cdot V_a + s + r)}{V_a + s + r} \tag{1'}$$

where the additional notation is

V_a = unfilled vacancies in absolute number

v_a = new vacancies in absolute number

k = the 'efficiency' factor in unemployment–vacancy matching

Note that L_r in (2) contains new vacancies (as a ratio to employment). The change from ratios with index r in (2) to absolute numbers with index a in (1) is due to the assumption behind (1) that wages agreed on in matchings are independent of each other.

The duration of a job match s, which appears in (1'), is not explicitly visible in (2). That difference is more formal than real, though, because in deriving (2) I have used the technique of Markov decision processes, which is a branch

of stochastic controls (see Schager (1988b) and Schager (1987)). While thus the decision problem of the firm is made truly stochastic, tractability imposes the cost, however, of restricting the analysis to a pure recruitment model without quits. As I have demonstrated in Schager (1987), p. 157–162, there is a structural correspondence between a stochastic recruitment control and a deterministic recruitment–cum–quit control, the latter being the standard approach in search theory models. Consequently, the duration of a vacancy in the model from which (2) is derived, can be interpreted as the expected time until a net recruitment occurs. Of course, this can be explicitly worked out, but I favour the view that research efforts in search theory should rather be directed torwards the harder task of developing the stochastic control methods properly.

The main task of this paper is hereby completed. We have demonstrated that search theory is able to give the notion of labour scarcity a sensible interpretation and that the wage response of firms to such scarcity is what intuition says: it is optimal for the firm to increase its wage, at given productivity and product prices, when the degree of labour scarcity increases. Hence a more pronounced labour scarcity implies a higher real wage and a higher wage share.

Conluding remarks on the empirical evidence and the inflation mechanism

The framework of this paper has been analytical. We have analysed the behaviour of the wage as a response to labour scarcity in a model, where the product market has been reduced to an exogenous price. As far as empirical studies of the behaviour of real wages and of wage shares are concerned, one must of course consider whether other assumptions as to the price formation mechanism are more appropriate in the area under investigation. Different models of monopolistic price setting are obvious alternatives. A search model of dynamic monopolistic competition in the product market accords well with the labour market analysis of this paper. I would prefer it to models of static monopolistic competition, but an extension along such lines would clearly entail more of analytical difficulties. Anyhow, the issue of price setting is outside this paper.

Another caveat with respect to the application of an equation such as (2) to empirical observations stems from the fact that short run equilibrium does not prevail, as long as the wage distributions adjust to the positions determined by unfilled vacancies, prices and productivity. As these entities are changing frequently, not to say continuously, a strict short run equilibrium may seldom be observed. As to the adjustment mechanism, by which the non–degenerate reservation wage and wage offer distributions change, we have still to rely on ad hoc assumptions, although the notions of excess demand and supply can now be interpreted in terms of the deviation of the wage level from short run optimality as defined by (2) or some amended version of it.

Despite these qualifications it is worth emphasising some interesting implicatons of the search–cum–wage model. The equation (2) is easily transformed to yield a price/wage ratio, which looks like a mark up equation, although prices are exogenenously given. The "mark up" varies inversely as the unfilled vacancy ratio. My model predicts that product real wages or seeming mark ups vary over the business cycle as vacancies do, although neither non–constant returns to labour nor monopolistic price–setting is present.

It is also evident that the product real wage (adjusted for labour productivity increases) and the wage share are high for at least two reasons, that are usually regarded as characterising different phases of the business cycle. The first reason is that firms want to expand and announce more new vacancies than before, although the speed of recruitment is unchanged. The other one is that the speed of recruitment goes down, i.e. vacancy durations get longer, as unemployment falls, although expansion plans are unaltered.

As to the development of employment the first phase ought to be characterised by increasing employment. In the second phase a part of the employment expansion plans have been fullfilled, so employment should be higher and unemployment lower. The depressing wage effect of a lower ratio between new vacancies and employment is counteracted, though, by the prolonged vacancy durations. Consequently, (2) predicts that product real

wages and wage shares should vary procyclically in the sense that high wages and wage shares are associated with increasing and/or high employment.

There exists a long tradition of empirical studies on the behaviour of the 'real wage' over the business cycle. It has to a large extent been associated with the theoretical controversy, referred to above, why not employment and real wages exhibit such a clear inverse relation as the neoclassical demand curve predicts. I have no ambition to survey that literature, but offer Barsky and Solon (1989) as a recent reference. They conclude that there is clear evidence against the countercyclicality of real wages, but that the evidence is more mixed as to whether real wages are non–cyclical or pro–cyclical. Their own 'analysis of longitudinal (U.S.) microdata... suggests that real wages are indeed quite procyclical', however (Barsky and Solon (1989), p. 28). I can only add that the theoretical analysis of this paper offers a good basis for explaining such findings, at the same time as it points to the necessity of defining clearly what is meant by concepts such as the 'real wage' and 'procyclicality'.

Let us conclude by reflecting for a moment on the Phillips curve issue, which was the starting point of this paper. Any correspondence between a theory of real product wages and wage shares and a theory of wage and price inflation is of course crucially dependent on the way in which product prices are formed. Labour scarcity will be associated with price inflation, only if some mechanism allows firms to increase their prices and thus to reduce again, at least temporarily or partly, the real wage. The mechanism may be thought of as increased demand in the product market, caused automatically by the higher wage share but with lagged impact. It may be thought of as monopolistic cost–based mark–up pricing practices by firms, which are inconsistent with the 'mark up' that is implicit in the (high) degree of labour scarcity (an inflation–causing 'battle of mark ups', but of another kind than envisaged by Layard and Nickell, cf. above).

The important thing, however, is not that price inflation cannot be stopped or at least substantially reduced at high degrees of labour scarcity. Presumably it can, at least by some draconic policy measures. What is important is that the effect of such price stabilization will be a low profit

share. The short run effect on labour demand may not be overwhelmingly large, if diminishing returns to labour in production are not prevalent. Nevertheless, investment for expansion must suffer and consequently labour demand will eventually fall. With lesser degree of labour scarcity, the profit share has a potential to increase. I say potential, because if money wages are not flexible downwards — and it would be unwise to think of falling money wages being but an extraordinary case — a low degree of labor scarcity is a necessary but not sufficient condition for a high profit share. What is also needed is a sufficiently high value added. If that prerequisite is to be accomplished by productivity increases only, the recovery of the profit share may well be a protracted process of high unemployment and low growth.

So even if a high degree of labour scarcity does not by itself produce price inflation, it must be tempting for a government, committed to full employment and economic growth, to try to undo the effects of labour scarcity on factor shares by allowing product prices to rise through expansionary policy measures. Such a policy may indeed be sucessful in counteracting the lingering effects of a former labour scarcity on the present profit share. So the impact of labour scarcity on price increases via its effect on the wage and profit share can be at work through many channels. In a long run context they may serve as equilibrating mechanisms, which ensure that factor shares balance at a near—constant ratio, as long time series seem to demonstrate ('Bowley's law').

References

Arrow, K.J. (1959), "Toward a Theory of Price Adjustment", in M. Abramovitz et al., *The Allocation of Economic Resources* (Stanford: Stanford University Press).

Barsky, R. and Solon, G. (1989), "Real Wages over the Business Cycle", NBER Working Paper No. 2888.

Blanchard, O.J. and Diamond, P. (1989), "The Beveridge Curve", *Brookings Papers on Economic Activity*, 1/1989, 1–76.

Hansen, B. (1951), *A Study in the Theory of Inflation* (London: Allen and Unwin).

Kuh, E. (1967), "A Productivity Theory of Wage Levels – an Alternative to the Phillips Curve", *Review of Economic Studies*, 34, 333–360.

Layard, R. and Nickell, S. (1986), "Unemployment in Britain", *Economica*, 53 (suppl.), 121–169.

Mortensen, D.T. (1970), "A Theory of Wage and Employment Dynamics", in E.S. Phelps (ed.), *Microeconomic Foundations of Employment and Inflation Theory* (London: Macmillan).

Mortensen, D.T. (1988), "The Persistence and Indeterminacy of Unemployment in Search Equilibrium", paper presented at the *Scandinavian Journal of Economics 90th Anniversary Symposium "Whither Macroeconomics?"*

Nickell, S. (1986), "Unemployment and the Real Wage", in C.–H. Siven (ed.), *Unemployment in Europe. Analysis and Policy Issues* (Stockholm: Timbro).

Phelps, E.S. (ed.) (1970a), *Microeconomic Foundations of Employment and Inflation Theory* (London: Macmillan).

Phelps, E.S. (1970b), "Money Wage Dynamics and Labor Market Equilibrium", in E.S. Phelps (ed.), *Microeconomic Foundations of Employment and Inflation Theory* (London: Macmillan).

Phillips, A.W. (1958), "The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957", *Economica*, 25, 283–299.

Pissarides, C.A. (1985), "Short-run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages", *American Economic Review*, 75, 676–690.

Pissarides, C.A. (1987), "Search, Wage Bargains and Cycles", *Review of Economic Studies*, **54**, 473–483.

Schager, N.H. (1987), *Unemployment, Vacancy Durations and Wage Increases: Applications of Markov Processes to Labour Market Dynamics*, Research Report No. 29, Industrial Institute for Economic and Social Research, Stockholm

Schager, N.H. (1988a), "Causes of Wage Increases in Swedish Manufacturing", IUI Working Paper No. 195, Industrial Institute for Economic and Social Research, Stockholm.

Schager, N.H. (1988b), "Search Theory, Downward Money Wage Rigidity and the Microfoundations of the Phillips Curve", IUI Working Paper No. 200, Industrial Institute for Economic and Social Research, Stockholm.

Solow, R.M. and Stiglitz, J.E. (1968) "Output, Employment, and Wages in the Short Run", *Quarterly Journal of Economics*, **82**, 537–560.