What does economic theory tell us about labour market tightness?

Andrew Brigden
and
Jonathan Thomas

Working paper no. 185

Bank of England, Threadneedle Street, London, EC2R 8AH. andrew.brigden@bankofengland.co.uk, jonathan.thomas@bankofengland.co.uk

The views expressed here are those of the authors and do not necessarily reflect those of the Bank of England, or the Monetary Policy Committee. The authors gratefully acknowledge comments received from Larry Ball, Ian Bond, Mike Joyce, Stephen Nickell, Chris Pissarides, Mark Schweitzer, two anonymous referees, and seminar participants at the Bank of England.

Copies of working papers may be obtained from Publications Group, Bank of England, Threadneedle Street, London, EC2R 8AH; telephone 020 7601 4030, fax 020 7601 3298, e-mail mapublications@bankofengland.co.uk

Working papers are also available at www.bankofengland.co.uk/wp/index.html

The Bank of England's working paper series is externally refereed.

ISSN 1368-5562

Contents

Abstract

Labour market tightness is a phrase often used by commentators and policy-makers, but it is rarely defined. In this paper, the phrase 'labour market tightness' is interpreted as describing the balance between the demand for, and the supply of, labour. A logical consequence of this approach is that tightness is not a helpful concept in those models of the labour market, such as the standard competitive and the basic matching model, where there are insufficient rigidities to create imbalances between labour demand and supply. It is proposed that changes in the labour share of income are a convenient yardstick for measuring changes in labour market tightness. In response to certain kinds of shock, changes in the labour share will give misleading signals, but this is likely to occur less frequently than with other oft-cited tightness indicators such as the unemployment rate or the employment rate. The paper concludes by considering the links between labour market tightness and inflation. A key lesson from this analysis is that any attempt to infer the relationships between labour market tightness, various market indicators of it, and inflation, requires both a clear definition of tightness and depends on the specific model of the labour market.

Key words: Labour market tightness, labour share, inflation.
JEL classification: E240, E310, J230.

Summary

The aim of this paper is to offer a coherent framework for examining the underlying drivers of labour market tightness, and the relationship between labour market tightness and inflation. Our motivation stems from the fact that although the phrase 'labour market tightness' is frequently used in the economics literature, it is rarely defined. Nonetheless, a variety of empirical evidence on labour market quantities and prices, such as unemployment and average earnings growth, is often cited as evidence of changes in the tightness of the labour market. Without a clear definition of the phrase it is difficult to evaluate the usefulness of any evidence offered; and a proper understanding of the relationship between tightness and inflation is also problematic.

In our view 'labour market tightness' can be defined in terms of its implications for the labour share of income. This follows from the notion that the labour market is tight (loose) when there is an imbalance between labour demand and labour supply, which will exert upward (downward) pressure on real unit labour costs, or equivalently on the labour share. Because the words tight and loose imply a degree of imbalance, we assert that the labour market can only be considered tight or loose out of steady state. This has two important implications. First, no shock can cause the labour market to become tight or loose unless it pushes the labour market away from its steady state. In practice, this is not too restrictive, since the kinds of rigidity that are present in most popular macro-models are sufficient to do this. Second, any shock that alters the steady-state value of the labour share cannot be said to have made the labour market permanently tighter or permanently looser. This is because movements in the steady-state do not involve any change in the balance between the demand for, and supply of, labour.

We use our definition of labour market tightness and its associated properties to examine the tightness implications of several popular labour market models. We start with the basic competitive model, and then work through models of efficiency wages, insider power, skill mismatch and matching frictions. A key message of this exercise is that the implications of much-cited indicators of changes in labour market tightness, such as unemployment, depend critically upon both the underlying economic shock and any market rigidities. For example, in the model of insider power a positive shock to nominal money balances leads to a tightening of the labour market that is accompanied by a decline in unemployment, which subsequently rises over time back to its unchanged steady-state value. On the other hand, an adverse labour supply shock in the perfectly competitive or efficiency-wage models leads to a tightening of the labour market that is accompanied by a rise in unemployment to a higher steady-state value.

We then turn to the relationship between tightness and inflation. By our definition, a tightening of the labour market will cause the labour share of income to rise. Since labour market tightness is a real phenomenon, it will have no implications for inflation unless the economy is subject to some form of nominal rigidity. Examples of such rigidities that could plausibly underpin a link between tightness and inflation include sticky price expectations, and restrictions on the frequency with which firms can alter prices. If such frictions are present, it is possible for out of steady-state movements in the labour share to influence inflation. A key lesson from this analysis is that any attempt to infer the relationships between labour market tightness, various market indicators of it, and inflation, requires both a clear definition of tightness and depends on the specific model of the labour market.

1. Introduction

Labour market tightness is a phrase often used by economic commentators and policy-makers alike, yet it is rarely defined. Nonetheless, a variety of empirical evidence on labour market quantities and prices, such as unemployment and average earnings growth, are often cited as evidence on changes in the tightness of the labour market. Without a clear definition of the phrase it is difficult to evaluate the usefulness of any evidence offered. A proper understanding of the relationship between tightness and inflation is also problematic.

With these issues in mind, this paper has two key aims. The first is to outline what popular labour market models have to say about the drivers of labour market tightness as we interpret it. To this end, we start with the basic market-clearing model, and then work through models of efficiency wages, insider power, skill mismatch, and matching frictions. The second aim is to investigate the linkages between tightness and inflation. Policy-makers and commentators tend to speak in terms that imply a link between a tightening of the labour market and inflation. However, labour market tightness, as we interpret it, is fundamentally a real concept. A tightening of the labour market may cause firms to pay more in real terms for a given quality of labour, but it does not necessarily follow that there are consequences for the general price level. Given plausible nominal rigidities in the economy, it is, however, possible for movements in labour market tightness to have inflationary consequences. Several examples of such nominal inertia, such as sticky price expectations, are discussed.

Before we go any further, it is worth setting out some important concepts. First, what does labour market tightness mean? Ideally, one needs a working definition that is sufficiently broad to embrace all of the models considered. To this end, we propose that labour market tightness be taken to describe the balance between the demand for, and the supply of, labour. If the demand for labour increases relative to supply, the labour market tightens, then we can expect some upward pressure on the real price of a given quantity of labour. So a tightening of the labour market is likely to produce an increase in real unit labour costs, or equivalently an increase in the labour share of income.(1)

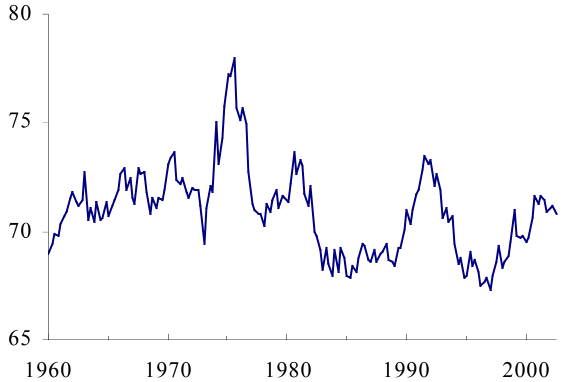
This suggests that movements in the labour share can be used to measure changes in market tightness. However as the paper proceeds, it will soon become apparent that, even in the simplest

(1) The labour share is the fraction of total output that goes to workers. In our context it should be thought of as excluding the self-employed as they are ignored in the labour market models we consider. Batini, Jackson and Nickell (2000) consider a measure that (i) includes the self-employed and public sector, (ii) includes the self-employed and excludes the public sector. Their exclusion of the public sector can be justified by the fact that they are primarily interested in modelling the pricing decisions of private sector firms. Both measures give similar results in regressions explaining the 1972-99 behaviour of inflation.

labour market models, focusing on the labour share can lead to stark, and arguably counter-intuitive, conclusions. To illustrate, when the production function is Cobb-Douglas, both product and labour markets are perfectly competitive, and firms have the right to manage, the labour share is constant and equal to the exponent on labour in the production function. This means that shocks to labour supply, caused perhaps by a change in the replacement ratio, which intuitively one might associate with a tightening or loosening of the labour market, have no effect on our proposed indicator. The flip side is that shocks to the exponent on labour in the production function, or (if we relax the assumption of perfectly competitive product markets) shocks to the degree of product market competition, cause a permanent shift in the labour share, and therefore market tightness.

To address these issues, we shall argue that shocks to labour supply can only affect labour market tightness if they move the economy out of steady state in such a way as to generate imbalances between the demand for and supply of labour. This will occur if there are certain rigidities in the economy, such as sluggish employment adjustment, or adaptive wage and price expectations. These sorts of rigidity are built into some of the labour market models we consider, but they are not present in all of them. For example, the textbook competitive and matching models are always in equilibrium. As a result the notion of a 'tight' or 'loose' labour market is not a helpful concept in these cases. Moreover, we assert, for reasons that we hope will become clear, that shocks to labour demand which cause a permanent shift in the labour share do not indicate a permanent shift in labour market tightness. As Chart 1 below shows, the UK labour share has tended to fluctuate around a well-defined mean, suggesting that permanent shocks to the UK labour share have been either very infrequent, offsetting, or small.

Chart 1
The labour share of income in the United Kingdom[(a)]



(a) The numerator is compensation of employees. The denominator is GDP at factor cost minus an estimate of self-employment income.

Our framework implies that it is more convenient to judge whether the labour market has become tighter or looser by looking at labour market prices. But changes in labour market tightness often have implications for labour market quantities. Indeed, it is sometimes argued that the effects of changes in labour market tightness will be seen in labour market quantities before they are seen in prices. This belief is implicit in the structure of the Bank of England's macroeconometric model (Bank of England (1999)), where a shock to aggregate demand first raises hours worked. Other things equal, this increase in hours worked raises the number of people in employment and reduces the number of people in unemployment. It is this reduction in unemployment that finally puts upward pressure on real unit wage costs, or equivalently the labour share of income, through pay growth.

So in the Bank of England's macroeconometric model a tighter labour market implies a lower unemployment rate. As we review the different theoretical models, we shall note what each has to say about the kind of changes in labour market quantities that one might see when the labour market becomes tighter. In some cases we find that unemployment is likely to be falling, but in other cases it is likely to be rising. Thus one of the key lessons from this paper is that focusing on quantity-based measures can sometimes give a misleading impression about the evolution of labour market tightness.

To sum up, we define labour market tightness as describing the balance between labour demand and supply. Changes in labour market tightness are likely to produce changes in real unit labour costs, or equivalently, changes in the labour share of income. In our view, the labour market can never be considered permanently tighter or permanently looser: demand/supply equilibrium must eventually be re-established. Consequently, economic shocks can only affect labour market tightness if, and for as long as, they move the labour market away from that steady state. A tightening of the labour market, in our terms, will cause the labour share of income to rise. But since we define labour market tightness as a real phenomenon, such movements will have implications for inflationary pressure only if the economy is subject to some form of nominal rigidity. Examples of such rigidities that could plausibly underpin a link between tightness and inflation include sticky price expectations, and restrictions on the frequency with which firms can alter prices. Naturally, none of this affects the conclusion that the rate of inflation is ultimately determined by the monetary stance.

The layout of the paper is as follows. Section 2 sets out a number of popular labour market models and outlines what each model has to say about the determinants of the labour market tightness. While there is some overlap in the underlying drivers, several of the models point to determinants that are not covered by others. In Section 3, we outline several mechanisms by which a tightening of the labour market will lead to an increase in inflationary pressure. A final section concludes.

2. Labour market models

2.1 The standard competitive model

Background

The standard competitive model assumes that firms and workers are price takers in the product and labour markets, and that wages and employment are set at the level where aggregate labour demand and aggregate labour supply are equalised. Thus unemployment is voluntary and is defined as the difference between some exogenous time endowment and actual hours worked.

We begin with a very simple structure. On the demand side, output is produced according to a Cobb-Douglas production function with labour and capital and as factor inputs. This generates a conventional downward sloping labour demand curve. We assume that firms are always on their labour demand curves which means that they set employment (firms have the 'right to manage'). The labour supply curve is derived from a utility function defined over consumption and leisure and is assumed to be upward sloping. Equilibrium wages and employment are affected by technological changes that are biased towards a particular factor input, and by changes in the level of unemployment benefits. Technological changes shift the labour demand curve, while benefit changes shift the labour supply curve.

Key equations

Labour demand is derived by assuming that each identical firm faces a Cobb-Douglas production function given by **(2.1.1)**.

$$Y = N^{\alpha} K^{1-\alpha} \qquad \textbf{(2.1.1)}$$

Y is output, N is employment and K is the (fixed) capital stock at firm. For simplicity, we assume that $0 < N < 1$ so that N is also the employment rate. Profit (π) is given by:

$$\pi = PY - WN - rK \qquad \textbf{(2.1.2)}$$

where P is the price level, W is the wage rate, and r the rental rate of capital. Differentiating **(2.1.2)** with respect to N and using **(2.1.1)** leads to the following labour demand equation,

$$\frac{W}{P} = \alpha \frac{Y}{N} \qquad \textbf{(2.1.3)}$$

The supply side of the labour market is based on an aggregate utility function of workers, which is defined over a composite consumption good, *C*, and leisure, 1-*N*. Utility is described by the following CES utility function:

$$U = [\beta C^{(\gamma-1)/\gamma} + (1-\beta)(1-N)^{(\gamma-1)/\gamma}]^{\gamma/(\gamma-1)} \qquad \textbf{(2.1.4)}$$

0<β<1 captures the weight which workers place on consumption and leisure, while γ is the elasticity of labour supply. Let $b < W$ be the real level of unemployment benefits. Then the workers' budget constraint is the sum of wage and benefits weighted by the share of time allocated to employment and unemployment respectively. Following Pissarides (1998), this budget constraint can be written as:

$$C = NW + (1-N)b \qquad \textbf{(2.1.5)}$$

Substituting **(2.1.5)** into **(2.1.4)**, it follows that workers choose *N* to maximise the following utility function:

$$U = [\beta(NW + (1-N)b)^{(\gamma-1)/\gamma} + (1-\beta)(1-N)^{(\gamma-1)/\gamma}]^{(\gamma-1)/\gamma} \qquad \textbf{(2.1.6)}$$

Maximising utility with respect to *N* leads to the following labour supply schedule:

$$N/(1-N) = [(\beta(W-b)/(1-\beta))^{\gamma} - b]/W \qquad \textbf{(2.1.7)}$$

Differentiating the labour supply schedule with respect to *W* shows that it is upward sloping if we assume that $\gamma > 1 - b/W$. Market equilibrium is given by the intersection of labour demand **(2.1.4)** and labour supply **(2.1.7)** which can be solved for wages and employment. The exogenous variables in the model are α, the exponent on labour in the production function, which shifts the labour demand function, and *b*, the real level of benefits, which shifts labour supply.

Implications for labour market tightness

Chart 2
Competitive model with instantaneous employment adjustment: adverse labour supply shock

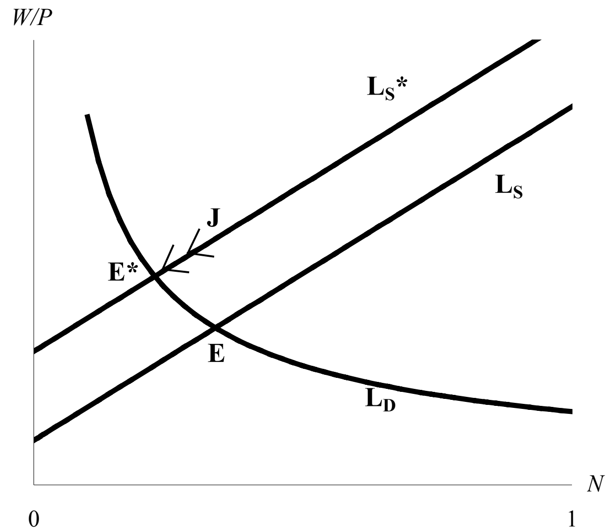


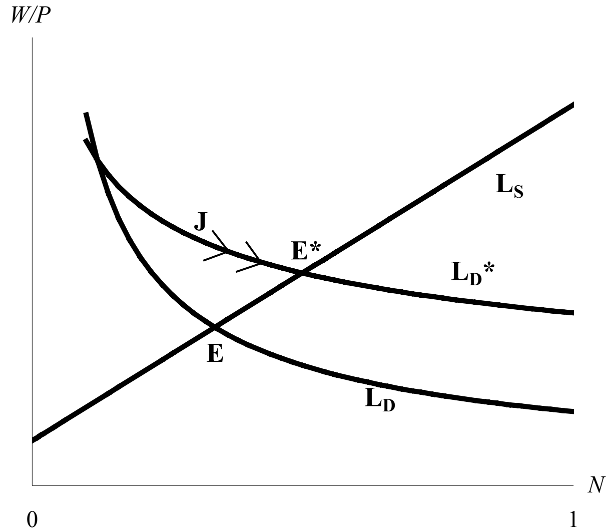
Chart 2 plots the labour demand and labour supply curves in real wage/employment space. Suppose that unemployment benefits, *b*, rise. The labour supply curve shifts from $\mathbf{L_S}$ to $\mathbf{L_S^*}$ and the equilibrium moves from **E** to **E***. The real wage is now higher, and employment is lower, but what of the labour share? Looking back at equation **(2.1.3)**, we can see that all along the labour demand curve, the labour share ($W.N / P.Y$) is constant and equal to α. And since firms have the right to manage, we know they are always on their labour demand curve. As the economy moves from **E** to **E***, the real wage increases but employment falls. As employment falls productivity rises. Indeed, employment falls and productivity rises just enough such that the labour share is unchanged.

Does this mean the labour share is not a good indicator of labour market tightness? Not necessarily. In the above example, there is no imbalance between the demand for and the supply of labour. At point **E***, and at point **E**, firms are operating on their labour demand curve and workers are operating on their labour supply curve. When unemployment benefits rise, firms rationally cut production and hence their need for labour declines. In this perfectly competitive model, the economy is always in steady state and, according to our definition, the labour market can be neither tight nor loose.

In practice, it seems unlikely that firms will immediately adjust employment following an economic shock, perhaps because of hiring and firing costs.[2] Consider instead how the labour share would respond if employment were fixed in the short term. Following the shift in the supply curve from $\mathbf{L_S}$ to $\mathbf{L_S}$*, the system must jump to a point like **J**. Now there is an imbalance between the demand for and supply of labour. In order to maintain staffing levels, firms must pay a higher real wage. But employment and productivity are unchanged. Hence the labour share rises. It is not optimal for firms to maintain such high levels of production, so employment starts to fall, and we move back down $\mathbf{L_S}$* to the final equilibrium at **E***. In an economy where the labour supply curve is hit by a series of infrequent shocks, and firms cannot adjust employment immediately, the labour share will vary over time but have a constant mean. This seems to be a reasonable description of the data in Chart 1.[3]

In this example, an assumption that employment is fixed in the period immediately following a shock is sufficient to push the economy out of steady state. It is this departure from steady state that allows scope for the labour market to be either 'tight' or 'loose' for short periods of time.

Chart 3
Competitive model with sluggish employment adjustment: positive labour demand shock



Equation **(2.1.3)** shows that the position of the labour demand curve $\mathbf{L_D}$ depends only on α, the exponent on labour in the Cobb-Douglas production function. Chart 3 shows how the labour demand

(2) Nickell (1986) and Bentolila and Bertola (1990) provide empirical evidence that such costs are important.
(3) If the production function is of the CES form then an adverse labour supply shock will lead to a rise in the labour share if the elasticity of substitution between capital and labour is less than one. The labour share will fall if this elasticity is greater than one. In the Cobb-Douglas case this elasticity is equal to one.

curve might shift following a rise in α.(4) Sluggish employment adjustment means that wages jump up in the short run to a point like **J**. Since **J** is on a higher labour demand curve, we know that the labour share has risen. This new higher labour share is maintained as real wages decline, employment rises and we reach a new steady state at **E***.

In the introduction we stated, without attempting to justify the statement, that we did not believe that a permanent shift in the labour share indicated a permanent shift in labour market tightness. We hope this example goes some way towards explaining our reasoning. At point **E*** there is no imbalance between the demand for, and the supply of, labour. Firms are operating on their labour demand curve and workers are operating on their labour supply curve. Workers enjoy an increased share of national output since, following the shock, they are producing an increased share of national output.

In models that have staggered wage and price setting and nominal rigidities, such as the one due to Layard, Nickell and Jackman (1991) that we consider in Section 2.3, then shocks other than to the exponent on labour in the production function can seemingly cause a labour demand curve that is drawn in actual real wage (W/P) employment space to move. That is because the labour demand curve is only stable when drawn in expected real wage (W^e/P) employment space. Consider a positive shock to nominal money balances that occurs after prices have been set. Demand and employment are likely to increase in the short term causing an increase in nominal wages that firms had not expected when they set prices. One could argue that this amounts to a temporary upward shift in the labour demand curve $\mathbf{L_D}$ as drawn in Chart 2. Yet we would argue this is not really a shock to labour demand in the true sense, rather that the rigidities in the model have forced firms to employ, temporarily, more labour than they would like to and hence operate off their labour demand curve.

---

(4) Notice that $\mathbf{L_D}$ and $\mathbf{L_D}$***** cross as N moves towards zero. At very low rates of employment, technology shocks that are biased towards labour and away from capital cause the real wage to fall. That is because such technology shocks are harmful to output when the capital stock is relatively large and hence reduce average labour productivity. This situation is unlikely to occur in practice.

Table A: The consequences of different shocks in the standard competitive model

Shock	Positive labour demand (α up)	Adverse labour supply (b up) (instantaneous adjustment)	Adverse labour supply (b up) (sluggish adjustment)
Unemployment	↓	↑	↗
Real wages	↑	↑	↑ ↘
Labour share	↑	→	↑ ↘

Table A lists the responses of three different variables (unemployment, the real wage rate, and the labour share) to shocks in the standard competitive model.

The first column outlines the implications of a labour-biased technology shift (ie a positive labour demand shock). Unemployment falls, real wages rise and the labour share rises. Yet we do not believe this signals a tighter labour market as there is no resulting imbalance between the demand for and supply of labour.

An increase in unemployment benefits (adverse labour supply shock) has no implication for the labour share and therefore market tightness, unless sufficient rigidities exist to shift the economy out of steady state. One such rigidity is a restriction that output cannot change in the immediate period. This would cause real wages to overshoot, and the labour share to jump up. In the long run, unemployment would rise, real wages would fall back slightly and the labour share would return to its (unchanged) steady-state value.

2.2 Efficiency wage models

Background

The key assumption here is that worker productivity is increasing in the wage paid by the firm. This may reflect the idea that more productive workers set higher reservation wages, so that higher pay attracts better job applicants and raises workforce quality. Alternatively, a higher wage may discourage worker shirking because it raises the cost of being unemployed. This link between worker productivity and pay provides a rationale for the existence of involuntary unemployment because firms have an incentive to pay wages above the level that clears the market.

We consider the shirking variant of efficiency wages developed by Shapiro and Stiglitz (1984). In this case, firms cannot perfectly observe worker behaviour. So to discourage shirking they attempt to pay a wage premium which exceeds any premium paid by their competitors. Such behaviour by all firms pushes average wages above the competitive equilibrium which generates involuntary unemployment. This unemployment acts to further reduce shirking because it raises the cost of job loss if the worker is discovered shirking and fired.

Compared to the competitive model, the efficiency wage model delivers a richer set of demand and supply-side variables that can influence the labour market equilibrium and thus the labour share. For example, monopoly power in the goods market can feed through to labour demand because the marginal product of labour is equal to the product of the mark-up and the real wage. On the supply side, changes in the ability of firms to monitor workers or the rate of labour turnover can also affect market tightness.

We employ a standard price-setting and wage-setting framework in real wage / employment space (Layard *et al* (1991), Blanchard (1997)). Specifically, identical imperfectly competitive firms produce a single good under a Cobb-Douglas production technology, and set employment to maximise profits. This leads to a downward sloping price-setting curve. Once again it is assumed that employment adjustment is sluggish. The wage-setting schedule, which Shapiro and Stiglitz (1984) call the 'no-shirking condition' is upward sloping, reflecting the idea that higher real wages are necessary to induce positive effort as the economy tends towards full employment.

Key equations

Each firm faces a Cobb-Douglas production function given by **(2.2.1)** and a constant-elasticity of substitution demand function given by **(2.2.2)**.

$$Y_i = N_i^{\alpha} K_i^{1-\alpha} \qquad \textbf{(2.2.1)}$$

$$Y_i = \left(\frac{P_i}{P}\right)^{-\eta} Y_d \qquad \textbf{(2.2.2)}$$

Y_i is the output of firm i, N_i is the number of people employed in firm i, K_i is the (fixed) capital stock at firm i, P_i is the price charged by firm i, P is an aggregate price index, η is the elasticity of demand, and Y_d is an aggregate demand index. Profit (π_i) is given by:

$$\pi_i = P_i Y_i - WN_i - rK_i \qquad \textbf{(2.2.3)}$$

where W is the wage rate (common to all firms) and r the rental rate of capital. Using **(2.2.1)** to substitute for N_i then **(2.2.2)** to substitute for Y_i and rearranging we obtain:

$$\pi_i = P_i^{1-\eta} P^{\eta} Y_d - W\left(P_i^{-\eta} P^{\eta} Y_d\right)^{\frac{1}{\alpha}} K_i^{\frac{\alpha-1}{\alpha}} - rK_i \qquad \textbf{(2.2.4)}$$

Differentiating with respect to firm i's price, setting this differential to zero, and using the assumption that all firms are identical, $P_i = P \ \forall \ i$ gives:

$$P_i = \left(\frac{1}{\alpha\kappa}\right) W \left(\frac{Y_d}{K_i}\right)^{\frac{1-\alpha}{\alpha}}$$
$$\text{where } \kappa = 1 - \frac{1}{\eta} \qquad \textbf{(2.2.5)}$$

Using this in the demand function **(2.2.2)**, we can see that actual output for all firms (Y_i) is equal to the demand index, Y_d. Making this substitution back into **(2.2.5)**, and using the production function to remove K_i, we obtain:

$$P = \left(\frac{1}{\kappa}\right)\left(\frac{W}{\alpha(Y/N)}\right) \qquad \textbf{(2.2.6)}$$

(2.2.6) says that product prices are equal to a constant mark-up over marginal cost (where marginal cost is the wage rate times the inverse of the marginal product of labour, $W/(\alpha Y/N)$. The mark-up is

$1/\kappa = \eta/(\eta - 1)$. Since we are conducting our analysis in real wage / employment space it is useful to rearrange the price-setting curve accordingly. Thus:[5]

$$\frac{W}{P} = \alpha\kappa \frac{Y}{N} \qquad \textbf{(2.2.7)}$$

The main innovation in the shirking model concerns the determinants of the supply side of the labour market which is given by the 'no-shirking condition'. The argument is neatly summarised by Pissarides (1998). Consider a stationary environment where identical risk neutral workers can be employed or unemployed. Let the expected returns in each state be E and U respectively. Workers have a utility function $U(w_i, e) = w_i - e$, where w is the real wage and e is the level of effort devoted to the job. The effort level is 0 if the worker shirks or $e > 0$ if he or she does not. All shirkers face the probability, q of being caught and dismissed into unemployment. Utility is zero if the worker is unemployed. Further suppose that there is an exogenous rate of job separations, s, and that workers discount the future at rate r.

Let the expected returns of a non-shirker and shirker be E_i^{ns} and E_i^{s} respectively. For a non-shirker, these returns are the discounted value of utility, $w_i - e$, less the discounted expected utility loss from exogenously losing the current job and moving into unemployment $s(E_i^{ns} - U)$. Thus,

$$rE_i^{ns} = w_i - e - s(E_i^{ns} - U) \qquad \textbf{(2.2.8)}$$

In the case of a shirker the utility of the job is w_i but the probability of entering unemployment is $s + q$. Thus we can write,

$$rE_i^{s} = w_i - (s + q)(E_i^{s} - U) \qquad \textbf{(2.2.9)}$$

Firms offer a wage that makes the individual indifferent between shirking and non-shirking. Thus in equilibrium $E_i^{ns} = E_i^{s} = E$. From **(2.2.8)** and **(2.2.9)** it follows that $E = U + e/q$. Substituting this

[5] Notice that this rearrangement makes the price-setting schedule look like a standard labour demand schedule. But the two are fundamentally different. The labour demand relation is derived under the idea that firms take wages and prices as given because they operate in competitive labour markets. However, the price-setting curve allows firms to set prices. Therefore the level of competition in the goods market will affect the price-setting curve but not the labour demand curve. Similarly the standard labour supply curve gives the wage at which a given number of workers are willing to work, while the wage-setting curve is the outcome of worker-firm bargaining or the unilateral decisions of firms. Therefore factors such as the decentralisation of pay bargaining will affect the wage-setting curve but not the labour supply curve.

into **(2.2.8)**, solving for w_i and imposing symmetry so that $w_i = w$ allows us to derive the aggregate real wage in terms of U. Therefore,

$$w = rU + (r + s + q)(e / q) \qquad \textbf{(2.2.10)}$$

The return from unemployment can be defined as the discounted value of unemployment benefits, b, plus the probability of moving from unemployment to employment (sN/U) multiplied by the discounted expected gain from this transition.[6] Thus,

$$rU = b + (sN / u)(E - U) \qquad \textbf{(2.2.11)}$$

Substituting **(2.2.11)** into **(2.2.10)** and using the fact that $E - U = e/q$ leads to the following wage-setting curve which is increasing and convex in the employment rate, N:[7]

$$w = b + (r + s + q)(e / q) + N / (1 - N)(se / q) \qquad \textbf{(2.2.12)}$$

Implications for labour market tightness

The equilibrium is given by the intersection of equations **(2.2.7)** and **(2.2.12)**. As in the competitive model technological shifts towards labour cause a permanent rise in the labour share. A rise in product market competition has a similar impact because the resulting rise in output demand feeds through to the demand for labour. However, in both cases labour market tightness is unchanged. Thus the outcome follows Chart 3 in Section 2.1 where the labour market jumps to point **J** in the short run and the labour share rises. In the long run the labour market moves to point **E*** where the new higher labour share is maintained. On the supply side, recall that the competitive model indicated that only movements in b affect the labour share. The shirking model delivers a richer set of tightness factors. Specifically, changes in the probability of job separation, s, and the probability of being caught shirking, q, are also important. For example, suppose that s rises. This increases the incentive to shirk because the worker faces a higher chance of entering unemployment. Thus the wage-setting curve or no-shirking condition shifts to the left. Like a rise in b, the sluggish adjustment of employment means that market tightness will temporarily rise. In this case the outcome follows Chart 2 in Section 2.1, where the labour market moves to **J** in the short run, and the labour share rises. In the long run the labour market moves to **E*** as the labour share declines to its original value. In contrast, an increase in q lowers the expected returns to shirking so that the

(6) The probability of moving from unemployment to employment is sN/U because the hiring rate equals the separation rate (sN) in stationary equilibrium and all hirings are from the unemployed (U).
(7) Shapiro and Stiglitz (1984) call this the 'no-shirking condition' as it shows the wage that firms must pay to induce workers to supply a non-zero level of effort, conditional upon the employment rate.

wage-setting curve shifts to the right. Consequently, labour market tightness declines in the short run before returning to its long-run value.

Table B: The consequences of different shocks in the efficiency wage model

Shock	**Positive labour demand (α up)**	**Adverse labour supply (b up, s up, q down) (instantaneous adjustment)**	**Adverse labour supply (b up, s up, q down) (sluggish adjustment)**
Unemployment	↓	↑	↗
Wages	↑	↑	↑↘
Labour share	↑	→	↑↘

Table B lists the responses of three different variables (unemployment, the real wage rate, and the labour share) to shocks in the efficiency wage model. While the efficiency wage framework allows a wider variety of shocks than the standard competitive model, the underlying message remains the same. Positive labour demand shocks cause a permanent reduction in unemployment, a permanent increase in the real wage and a permanent increase in the labour share. There are no implications for labour market tightness. An adverse labour supply shock will cause a temporary tightening of the labour market (accompanied by a rising unemployment and higher real wages), but only if output cannot adjust immediately.

2.3 The model of Layard, Nickell and Jackman (1991)

Background

In chapter 1 of their book, Layard, Nickell and Jackman (1991) sketch out a small-scale macro-model based on a labour market characterised by insider power. Their model differs from those we have considered so far in two important respects. First, the authors impose a nominal rigidity, in the form of sticky price expectations, which acts as an alternative to the sticky employment assumption that we have imposed thus far. Second, they include money in their model. This allows us to trace out the effects of shocks that originate outside the labour market, and which move neither the labour supply nor the labour demand curve, yet affect market tightness. It also allows us to consider the linkages between labour market tightness and inflation.

Models of insider power are characterised by the notion that firms rank existing staff (the insiders) above unemployed workers (the outsiders). This ranking gives the insiders a source of monopoly power when setting the wage, and provides micro-foundations for a wage-setting curve that is upward sloping in real wage / employment rate space. Unemployment persists in these models because, at low rates of unemployment, the competing claims of firms and workers over what is produced are incompatible. As the economy nears full employment, insiders know that, should they leave the firm, they can find another job very quickly. In order for them to stay with the firm they must therefore be offered a high real wage (high w-p). At the same time, as the economy nears full employment, labour productivity falls. By implication, the marginal cost of production rises which leads firms to push for a higher mark-up of product prices over the wage rate (high p-w and hence low w-p).

The functioning of the labour market in many large scale macroeconomic models of the UK economy, such as those employed by HM Treasury, NIESR, and the Bank of England, is essentially the same as that envisaged by Layard, Nickell and Jackman (1991) model.

Key equations

Period by period, the sequence of events is as follows:

1. There are many imperfectly competitive firms indexed by i. Each firm i sets a price p_i for its product, maximising expected profits conditional on the expected level of demand and expected money wages.

2. A demand shock is realised.

From here on the model is deterministic, and the following three events can be regarded as occurring simultaneously:

3. Each firm i supplies whatever is demanded (y_i^d) after the demand shock at the price (p_i) which they set at the beginning of the period. This assumption is crucial because it means that firms can be operating inefficiently, and produce more than they would choose to if prices could be immediately adjusted. It prevents firms from cutting output immediately following an adverse labour supply shock, and hence plays a similar role to the arbitrary restriction on employment adjustment that we made in earlier sections.

4. Capital is fixed. A Cobb-Douglas production function is sufficient to determine uniquely the

amount of labour (n_i) that each firm needs to hire in order to produce y_i^d.

5. The wage rate is determined as the outcome of a Nash bargaining process between firms and trade unions.

It can be shown that the labour demand and labour supply curves take the following general form.[8]

$$p - w^e = \beta_0 - \beta_1 u^e \qquad \textbf{(2.3.1)}$$

$$w - p^e = \gamma_0 - \gamma_1 u \qquad \textbf{(2.3.2)}$$

(2.3.1) is the labour demand curve (or price-setting rule). It is based on the idea that prices (p) are set as mark-up over the expected marginal cost.[9] **(2.3.2)** is the labour supply curve (or wage-setting rule). The Nash bargaining process produces an expected real wage that varies inversely with the unemployment rate. As unemployment falls, insiders know that, should they be laid off, they could quickly find work elsewhere. This raises the value of their fall-back point. The four parameters β_0, β_1, γ_0 and γ_1 should not be regarded as deep structural parameters. Rather they are functions of other parameters, such as the level of unemployment benefits, the degree of product market competition, and a measure of trade union power — all of which may change over time (see appendix).

The aggregate demand side of the model is represented by a reduced-form equation linking the unemployment rate to real money balances ($m - p$):

$$u = -\frac{1}{\lambda}(m - p) \qquad \textbf{(2.3.3)}$$

In steady state, all expectations are fulfilled. Using $p = p^e$, $w = w^e$ and $u = u^e$ in equations **(2.3.1)** and **(2.3.2)** we obtain the following expression for the unemployment rate in steady state (u*):

$$u^* = \frac{\beta_0 + \gamma_0}{\beta_1 + \gamma_1} \qquad \textbf{(2.3.4)}$$

---

(8) The labour demand curve or price-setting rule is a transformation of the standard labour demand curve derived in Section 2.2. Micro-foundations for the wage-setting rule are set out in the appendix.

(9) In general terms, the expected marginal cost is equal to the product of the expected wage and the inverse of the expected marginal product of labour. Under a Cobb-Douglas production function, the expected marginal product of labour is increasing in the expected unemployment rate (u^e), which means its inverse is decreasing.

To examine behaviour out of steady state, we need to specify some process for w^e, p^e and u^e. Layard, Nickell and Jackman (1991) assume that (i) errors in the forecast for w^e and p^e are the same and (ii) expectations for inflation and the unemployment rate are formed adaptively. Using $w-w^e = p-p^e$, $\Delta p^e = \Delta p_{-1}$ and $u = u^e$ in **(2.3.1)** and **(2.3.2)** we obtain:

$$\Delta p = -\theta_1 (u - u^*) + \Delta p_{-1}$$
$$\text{where } \theta_1 = \frac{\beta_1 + \gamma_1}{2} \qquad \textbf{(2.3.5)}$$

Equation **(2.3.5)** says that, whenever u lies below u^*, inflation will be rising and whenever u lies above u^*, inflation will be falling. Inflation can only be stable when u equals u^*, hence u^* is referred to as the non-accelerating inflation rate of unemployment (or NAIRU).

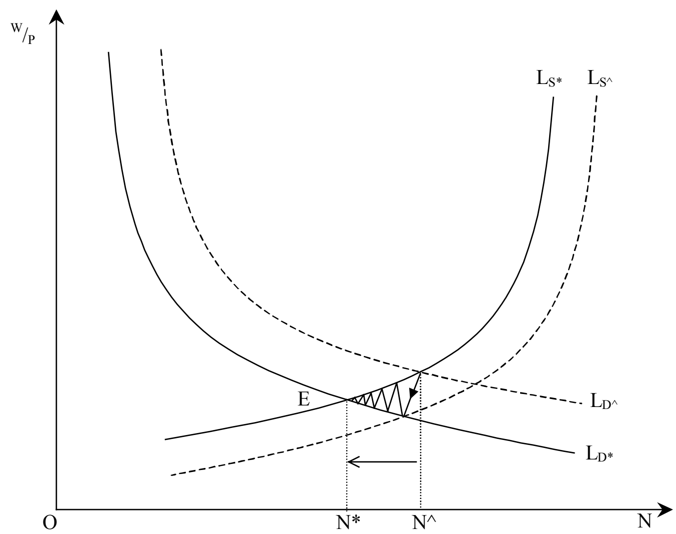
Using the model outlined above, we can analyse the consequences of many different kinds of shock (recall that β_0, β_1, γ_0 and γ_1 are all functions of other parameters). For our purposes, the shocks can be split into two camps: shocks to aggregate demand (which move u), and shocks either to labour demand or to labour supply (which move u^*). Starting from a position of long-run equilibrium (where $u = u^*$), let us consider first what happens following a shock to aggregate demand. This takes the form of an unanticipated increase in nominal money balances. Prices are fixed in the immediate period, so from **(2.3.3)** unemployment must jump down below u^*. From **(2.3.5)** this will be associated with a subsequent increase in inflation. What actually happens is as follows: to cope with the increase in demand, firms need to employ more labour than they had expected when product prices were set. This leads to a reduction in the cost to an insider of being fired (since immediate re-employment is more likely), and hence to an increase in the wage rate determined by the Nash bargain. Temporarily, insiders enjoy a higher real wage. But firms have been caught out on two fronts: not only have nominal wages gone up against a fixed product price, but increased employment means that, under Cobb-Douglas production, labour productivity has fallen. Consequently, firms aiming for a constant mark-up over marginal cost will raise their prices, and by more than the previous increase in nominal wages. This is the beginning of the classic wage-price spiral. Period by period, wages increase more rapidly than firms had been expecting, and product prices increase more rapidly than workers had been expecting. This will take place until price rises have eroded the initial shock to nominal money balances and the level of demand implied by **(2.3.3)** is again consistent with employment at u^*.

Shocks to labour demand or labour supply cause a jump in u^* rather than a jump in u. Again, starting from a position of equilibrium, consider the case where unemployment benefits rise. Now at any given unemployment rate, insiders are less concerned by the prospect of becoming unemployed. That means the real wage rate determined by the Nash bargain is higher at any given unemployment rate. From the labour demand curve, we know that if real wages are to be higher, productivity must rise, hence employment must fall and unemployment must rise. As in the case of demand shocks, the assumption of adaptive expectations means the economy does not jump immediately to the new equilibrium. Following the increase in benefits, workers push for a higher money wage. Firms raise their price in response, real money balances fall and through **(2.3.3)** unemployment begins to rise.

Implications for labour market tightness

In the Layard, Nickell and Jackman (1991) model, *u-u** provides a convenient quantities-based measure of labour market tightness.

Chart 4
LNJ (1991) model: positive nominal demand shock



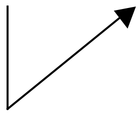
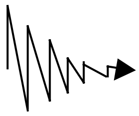
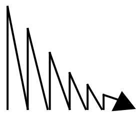
But it turns out that there is no simple mapping between *u-u** and the labour share. We use Chart 4 which plots labour demand (or price-setting), and labour supply (or wage-setting) curves in actual employment / actual real wage space, to illustrate this point. In steady state, when wage and price expectations are fulfilled, the economy will be at point **E**, at the intersection of $\mathbf{L_D}$***** and $\mathbf{L_S}$*****. Now consider what happens following an unanticipated increase in nominal money balances. Prices are fixed in the immediate period, so more goods will be demanded than firms had intended to supply. That means ***N*** must rise above ***N**** to a point like ***N^***. In order to attract extra staff, firms end up paying a higher real wage than they had anticipated. When drawn in actual employment / actual real wage space, the effective labour demand curve jumps to $\mathbf{L_D}$**^**. Real wages are higher, and productivity is lower, so the labour share has risen. Next period, the firm will raise its price to offset not only the rise in money wages, but also the decline in productivity. Firms are back on their long-run labour demand curve, $\mathbf{L_D}$*****. Now it is the insiders who are fooled. When drawn in actual employment / real wage space, the effective labour supply curve jumps to $\mathbf{L_S}$**^** and the labour share is back at its steady-state value. This process continues until price rises have fully eroded the real effects of the increase in nominal money balances. The labour share flips from being above its steady-state value, to being at its steady-state value, as the economy flips from being off the labour demand curve and on the labour supply curve, to being on the labour demand curve and off the labour supply curve.

The model of Layard, Nickell and Jackman is of interest in part because it allows for variations in the labour share under Cobb-Douglas technology, but also because it suggests a link between tightness and inflation (see **(2.3.5)**). But recall that **(2.3.5)** is only valid under adaptive expectations. When a shock to nominal money balances pushes u below u^*, workers bid up money wages on a view of product prices that turns out to be wrong. Then firms raise product prices on a view of money wages that turns out to be wrong. As this wage-price spiral escalates, both parties are fooled period by period. Both parties fail to perceive that the shock was a nominal one and can have only nominal consequences in the long run. If instead expectations were formed rationally, the increase in m would still cause a jump down in u below u^* (because prices are fixed in the immediate period). But p would then rise by the same amount as the rise in m, real balances and hence demand would be unchanged and u would return to u^*. In this example, tightness is only instantaneous and is associated with jumps in the price level, rather than sustained increases in the rate of inflation.

The assumption of adaptive expectations is one kind of nominal rigidity that is sufficient to provide a link between real developments in the labour market and inflation. For those who are uncomfortable with adaptive expectations, then other forms of nominal rigidity (such as overlapping contracts) could be used in conjunction with rational expectations to derive similar results.

Table C lists the responses of unemployment, the real wage rate and the labour share to a positive nominal demand shock in the LNJ (1991) model. This is one case where a reduction in unemployment is a reliable indicator of a tighter labour market. A positive nominal demand shock causes unemployment to jump down and then drift back up to its starting value. Under a strict interpretation of the LNJ (1991) model, both the real wage rate and the labour share would oscillate. This oscillation would go away if we assumed that different firms adjusted their wages and prices at different times. For example, one group of firms might set their prices in periods 1 and 3 and their wages in periods 2 and 4 while a second group might set their prices in periods 2 and 4 and their wages in periods 1 and 3.

Table C: The consequences of a demand shock in the LNJ (1991) model

Shock	**Positive nominal demand shock (*m* up)**
Unemployment	
Wages	
Labour share	

2.4 Models of skill mismatch

Background

The models considered so far have all focused on aggregate shifts in labour demand or supply. In contrast, models of skill mismatch emphasise changes in the composition of aggregate demand and supply. In particular, these models focus explicitly on the consequences of a change in balance between the demand and supply of skilled labour.

The standard mismatch model is based upon a labour market that is comprised of skilled and unskilled employees. A shift in labour demand towards skilled workers which is not met by a commensurate rise in supply, will tend to raise their wages and reduce their unemployment rate, while lowering the wages and raising the unemployment rate of the unskilled. Given a wage-setting function for each skill group that is decreasing and convex in unemployment, this will lead to a rise in aggregate unemployment. The magnitude of this extra unemployment depends on the flexibility of real wages, while its persistence will depend on the speed of adjustment of the relative supply of skilled workers.[10]

Our analysis is based on Manacorda and Petrongolo (1999) who develop an index of skill mismatch which captures the changes in the balance between the demand and supply of labour inputs differentiated by skill. The basic idea is that a net shift in demand for skilled workers will increase

[10] The mismatch idea also extends to shifts in the composition of demand across specific regions, industries, or occupations (Lilien (1982), Blanchard and Katz (1997)).

their share of the aggregate wage bill following upward pressure on skilled employment and wages. On the other hand, a net shift in the supply of skilled workers increases their share of the labour force. The index compares these relative demand and supply changes.

Key equations

The model has two central elements. The first is a production function with heterogeneous labour input which is used to derive the mismatch index. The second element is a wage-setting function for each labour group which is convex in their group-specific unemployment rates. These wage functions are used to show how a rise in mismatch affects aggregate unemployment.

To fix ideas, suppose that the aggregate production function has three inputs: capital (K) skilled labour (N_s) and unskilled labour (N_u) where capital is fixed. The production function takes the Cobb-Douglas form:

$$Y = AK^{\alpha_k} N_s^{\alpha_s} N_u^{\alpha_u} \qquad \textbf{(2.4.1)}$$

α_s, α_u and α_k are the relative demand indicators for each input where $\alpha_s + \alpha_u + \alpha_k = 1$. Following Nickell and Bell (1995), we define a relative demand shift towards skilled labour as $d\alpha_s = -d\alpha_u > 0$. A captures general technological progress. If W_s and W_u are the real wage rates of each skill group, then profit maximisation implies:

$$\frac{\alpha_s}{\alpha_u} = \frac{W_s N_s}{W_u N_u} \qquad \textbf{(2.4.2)}$$

Let L_s and L_u be the labour force of each skill group. Thus $E_i = N_i/L_i$ is the group employment rate. Dividing both sides by L_s/L_u gives:

$$\frac{\alpha_s}{\alpha_u} \bigg/ \frac{L_s}{L_u} = \frac{W_s (N_s / L_s)}{W_u (N_u / L_u)} \qquad \textbf{(2.4.3)}$$

Now let $l_i = L_i/L$ be the labour force share of group i, where L is the total labour force. Taking logarithms of both sides and totally differentiating throughout leads to:

$$d \ln\left(\frac{\alpha_s}{\alpha_u}\right) - d \ln\left(\frac{l_s}{l_u}\right) = d \ln\left(\frac{W_s}{W_u}\right) + d \ln\left(\frac{E_s}{E_u}\right) \qquad \textbf{(2.4.4)}$$

The left-hand side of **(2.4.4)** can be interpreted as the difference in the growth rates of the demand and supply of skilled labour relative to its unskilled counterpart. This makes it a natural index of movements in skill mismatch, which we denote by $d\ln SM$. A change in the relative demand for skilled labour is captured by $d\ln(\alpha_s/\alpha_u)$, which is the growth of the ratio of skilled and unskilled wage bill shares. l_s/l_u is the ratio of the skilled and unskilled labour force shares. So $d\ln(l_s/l_u)$ traces movements in the relative supply of skilled labour relative to unskilled labour. The right-hand side of **(2.4.4)** expresses the index in terms of relative wages (W_s/W_u) and employment rates (E_s/E_u). An increase in market imbalances translates into either a rise in the relative wages of skilled workers, a rise in their relative employment rates (which is interpreted as a rise in the unskilled unemployment rate relative to the skilled unemployment rate), or a combination of both.[11]

The model is completed by assuming that the wage function for skill group *i*, is of the form: $\ln(W_i)=z_i-\gamma\ln(u_i)$.[12] Thus wages are a decreasing convex function of unemployment. This convexity means that wages are more responsive to changes in unemployment when unemployment is low than when it is high. With these wage functions it can be shown that an increase in mismatch leads to a fall in the skilled unemployment rate, and a rise in the unskilled unemployment rate. The convexity of the wage-setting functions mean that the aggregate unemployment rate increases when the skilled have higher employment and/or wage rates. In practice, the skilled have higher rates of both employment and wages. Similarly, skilled wages rise, unskilled wages fall, and the aggregate wage rate rises.

A skill-biased demand shock has no permanent effect on the labour share, $s_l. =\alpha_s + \alpha_u$. This is because $d\alpha_s = -d\alpha_u$. Chart 5 shows the results of such a shock. The demand curve for skilled labour, L_{Ds} shifts rightwards to $L_{Ds'}$. This is offset by a leftward shift in the demand curve for unskilled labour from L_{Du} to $L_{Du'}$ so that total labour demand is unchanged. With no restriction on employment adjustment, the equilibrium in the skilled market moves from E^s to $E^{s'}$, while the unskilled market moves from E^u to $E^{u'}$. Skilled wages and employment rise, while unskilled wages and employment decline. The convex wage setting schedule (*WS*), means that aggregate wages and unemployment rise but market tightness is unchanged. This neutral effect of skill-biased shocks will also occur if employment adjustment is sluggish and the real wage is fully flexible. In this case, skilled wages

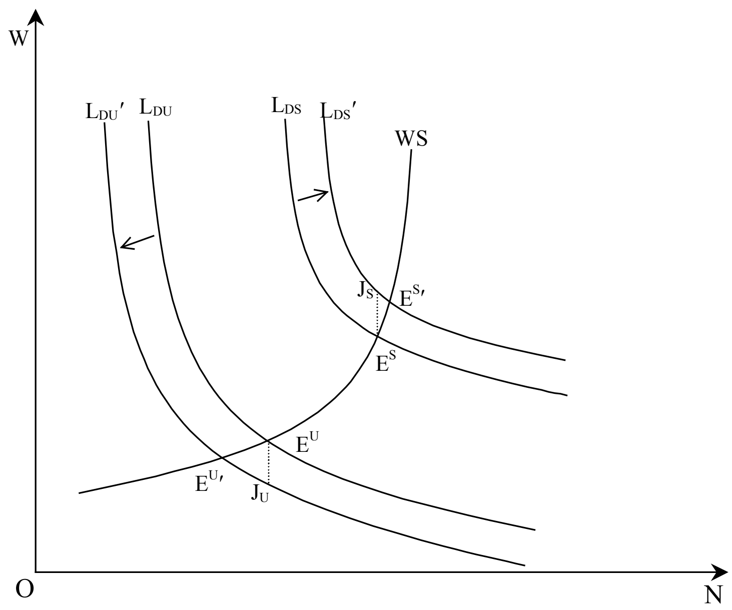
---

(11) Equation **(2.4.4)** assumes that the relative demand and relative supply of workforce skills are independent. This helps to simplify the analysis, but is unlikely to hold in practice. Consider a shift in demand towards skilled workers. Any resulting increase in the returns to skills should serve to increase the share of workers who will invest in acquiring the relevant attributes. This will eventually boost the relative supply of skills. Skill demand may also respond to an increase in supply. For example, a jump in the relative supply of skilled workers may induce firms to create skilled jobs.

(12) z_i includes the standard factors which move the wage-setting curve, including the level of benefits, worker bargaining power, and the long-term/short-term composition of the unemployed pool. The microfoundations of this double-logarithmic wage-setting function include the models of efficiency wages and insider power discussed in Sections 2.2 and 2.3 respectively. It can also be shown that the double-logarithmic specification is a log-linear approximation to the first-order conditions for wages in the wage bargaining model developed by Manning (1993).

will overshoot, while unskilled wages will undershoot, their long-run values, as the skilled and unskilled markets jump to J_s and J_u respectively.[13]

Chart 5
Skill mismatch model: skill-biased labour demand shock



The model also allows for aggregate supply shocks. These occur when an increase in standard wage-push factors such as benefits, or worker bargaining power leads to a rise in average wages above productivity growth (Jackman *et al* (1999)). In this case the wage-setting schedule in Chart 5 shifts upwards. With sluggish employment adjustment, this leads to a short-run rise in the labour share and increase in tightness. As before, wages and unemployment will eventually adjust to offset each other and the labour share will return to its original steady-state value (see Chart 2).

Implications for labour market tightness

The implications of the skill mismatch model for labour market tightness are summarised in Table D. A reallocation of labour demand towards skilled workers, which is not matched by a commensurate shift in the composition of labour supply, has no impact on tightness. Adverse aggregate supply shocks, which may reflect a rise in worker bargaining power or unemployment benefits, increase tightness in the short run. This framework also highlights the fact that movements in tightness might not be reflected in aggregate data. For instance, total hours worked is often used as an indicator of

(13) One way by which demand shocks can affect the labour share is if firms cannot cut unskilled wages, possibly due to minimum wage legislation. In this case, the labour share will increase.

aggregate demand.(14) However, adjustments in the structure of labour demand could increase the mismatch index, and yet have no effect on total hours.

Table D: The consequences of different shocks in the skill mismatch model

Shock	**Skill-biased labour demand (α_s up, α_u down)**	**Adverse labour supply (z up)**
Unemployment	↗	↗
Wages	↘	↘
Labour share	→	↘

2.5 Models of matching frictions

Background

Models of matching frictions focus on the process by which firms with job vacancies meet (or 'match') with individuals who are looking for work. There is persistent unemployment in matching models owing to the combination of two factors: (i) there is a steady flow into unemployment (adverse shocks will eventually cause a filled job to become unproductive and close down); and (ii) newly unemployed workers cannot immediately match with the stock of vacancies.

The vacancy-to-unemployment ratio (or *v/u*) is regarded as an important indicator of tightness in most matching models. This is because it measures the ease with which unemployed people find firms with vacancies, relative to the ease with which firms with vacancies find unemployed people. Many would consider this to be a natural use of the word 'tightness'. Indeed, in his comprehensive survey of matching models, Pissarides (2000) argues that 'by the structure of the model, [*v/u*] is an

(14) Of course observed hours are the outcome of the interaction of demand and supply in the market for working time. Nonetheless, working hours have been used as a proxy for labour demand or supply in the past. The November 1999 *Inflation Report* provides an example of the former. It states: 'There have been signs of a renewed pick-up in labour demand. According to the Labour Force Survey (LFS), total hours rose by 0.4% in the three months to August, having been broadly flat since last summer.' In contrast the February 2000 *Inflation Report* states: 'Total hours worked, often regarded as the broadest measure of labour supplied …'

appropriate measure of the tightness of the labour market'.(15) However, one may prefer not to use the expression in this context, as labour demand still equals labour supply. In the basic matching model that we describe here, the labour share is seen to depend on *v/u* as well as a number of exogenous parameters. However, without the imposition of additional rigidities, *v/u* is a jump variable. *v/u*, and by extension the labour share, are never away from their steady-state values. Since the standard matching framework does not admit any labour demand/supply imbalances, it follows that the labour market can never, in our terms, be 'tight' or 'loose'.

Key equations

At the heart of matching models is the matching function. This specifies the number of job matches (M) that occur per unit of time. In the benchmark model, M is an increasing function of the number of unemployed workers (U), and the number of job vacancies (V). Formally:

$$M = m(U, V) \qquad \textbf{(2.5.1)}$$

where the function m is homogenous of degree one such that a doubling of U and V will double the number of matches. We can use **(2.5.1)** to derive expressions for the probability that a vacancy is filled, and for the probability that an unemployed worker finds a job.(16)

$$\begin{aligned} P(\text{Vacancy is filled}) &= \frac{M}{V} \\ &= m\left(\frac{U}{V}, 1\right) \\ &= q(\theta) \\ \text{where } \theta &= \frac{V}{U} \end{aligned} \qquad \textbf{(2.5.2)}$$

$$\begin{aligned} P(\text{Unemployed worker finds a job}) &= \frac{M}{U} \\ &= \frac{V}{U} m\left(\frac{U}{V}, 1\right) \\ &= \theta q(\theta) \end{aligned} \qquad \textbf{(2.5.3)}$$

(15) Pissarides (2000, page 7).

(16) Other kinds of matching function have been suggested. Coles and Smith (1996) argue that firms and workers meet much more rapidly than suggested by **(2.5.1)**. In their model, if a large stock of unemployed people exist alongside a large stock of vacancies it is because the unemployed people and the vacancies are not compatible. They suggest the existing stock of unemployed will only meet with the flow of new vacancies and, equally, the existing stock of vacancies will only match with the flow of new unemployed.

Each filled job produces output with a value p>0 per period. A firm with a vacant job pays pc >0 per period to advertise the unfilled vacancy.[17] Jobs are destroyed at some exogenous rate λ. Once the match has taken place, a Nash bargaining process determines the wage rate. By implication, the wage rate (w) solves:

$$\max \Omega = (W - U)^{\beta} (J - V)^{1-\beta} \qquad \textbf{(2.5.4)}$$

where W is the present discounted value to an individual of being employed, U is the present discounted value to an individual of being unemployed, J is the present discounted value to a firm of a filled vacancy, V is the present discounted value to a firm of an unfilled vacancy and β is an index of the worker's bargaining strength. Expressions for W, U, V and J can be derived as follows:

$$rW = w + \lambda(U - W) \qquad \textbf{(2.5.5)}$$
$$rU = z + \theta q(\theta)(W - U) \qquad \textbf{(2.5.6)}$$
$$rJ = p - w + \lambda(V - J) \qquad \textbf{(2.5.7)}$$
$$rV = -pc + q(\theta)(J - V) \qquad \textbf{(2.5.8)}$$

(2.5.5) says the rental value of a job (defined as the capital value, W, times the interest rate, r) is equal to the wage per period plus the capital loss that would occur if the job were destroyed, times the probability that it will be destroyed. **(2.5.6)** says the rental value of being unemployed is equal to the flow utility of being unemployed (z) plus the capital gain that would occur if a job is found, times the probability that a job will be found. **(2.5.7)** says the rental value of a filled job is equal to the firm's per period profit (output net of the wage) plus the capital loss that would occur if the job were destroyed, times the probability that it will be destroyed. **(2.5.8)** says the rental value of a vacancy is equal to the negative of the per-period cost of advertising it, plus the capital gain that would occur if the vacancy were filled, times the probability that it will be filled.

A steady-state condition for unemployment, often referred to as the Beveridge Curve, is obtained by setting unemployment inflows $\lambda(1 - u)$ equal to unemployment outflows $\theta q(\theta)u$ and rearranging:

$$u = \lambda/(\lambda + \theta q(\theta)) \qquad \textbf{(2.5.9)}$$

[17] It does seem reasonable that recruitment costs are linked to productivity. Recruitment consultants often charge employers a proportion of the salary paid to the new employee. By extension, employers must face similar costs if they did the recruiting themselves, or else they wouldn't hire recruitment consultants!

We assume that firms will post vacancies until it is no longer profitable to do so. Hence, in equilibrium, the present discounted value of a vacancy (V) will be zero. Using this in **(2.5.7)** and **(2.5.8)** we obtain the job creation condition:

$$\frac{p-w}{r+\lambda}=\frac{pc}{q(\theta)} \qquad \textbf{(2.5.10)}$$

(2.5.10) says that firms will continue to post vacancies until the capitalised value of the per-period profit flow $(p-w)/(r+\lambda)$ just equals the expected total cost of posting the vacancy. To find the wage rate, we use all four value functions and the Nash bargaining condition **(2.5.4)**. After some manipulation we obtain:

$$w=(1-\beta)z+\beta p(1+c\theta) \qquad \textbf{(2.5.11)}$$

(2.5.11) says that the wage rate is a weighted average of the flow utility an individual gets by remaining unemployed (z) and the expression $p+pc\theta$. $p+pc\theta$ can be thought of as the expected social benefit generated by each unemployed worker that is hired. First, he or she produces an amount p per period. Second, he or she makes an expected saving on hiring costs of $pc\theta$ per period.[18] The weights are based on the index of bargaining strength. When the worker is dominant ($\beta=1$) he or she gets the entire surplus generated by the match. When the firm is dominant ($\beta=0$) the worker is forced down to his or her reservation wage.

If we make the simplifying assumption that the flow utility from unemployment is proportional to the wage rate, so that $z=\rho w$ where ρ is the replacement ratio, then we get the following expression for the labour share of income:

$$\frac{w}{p}=\frac{\beta(1+c\theta)}{1-(1-\beta)\rho} \qquad \textbf{(2.5.12)}$$

(2.5.12) says that the labour share of income is increasing in the bargaining strength of workers (β), the cost of posting a vacancy relative to the output from a filled job (c), the replacement ratio (ρ) and the vacancy to unemployment ratio, θ.[19]

---

[18] To see this, note from **(2.3.2)** and **(2.3.3)** that the expected duration of a vacancy is $1/q(\theta)$ while the expected duration of unemployment is $1/\theta q(\theta)$. That means if an unemployed worker pulls out of negotiations once he or she has matched with a vacant job, then the expected per-period cost of re-hiring him or her will be pc times E(duration of vacancy) / E(duration of unemployment) = $pc\theta$.

[19] Although it may be thought that the labour share of income is equal to 1 in a matching model without capital this is not the case. For example, equation **(2.5.12)** shows that provided employers have to pay to post vacancies ($c>0$) then workers will never be able to secure all of the surplus generated from a match.

The system can be analysed graphically with the aid of two diagrams. Chart 6A plots the job creation curve (*JC*) given by **(2.5.10)** and the wage-setting curve (*WC*) given by **(2.5.11)** in vacancy-to-unemployment ratio / real wage space. This chart can be used to solve for the equilibrium vacancy-to-unemployment ratio, θ^*. Chart 6B shows the Beveridge curve **(2.5.9)** and θ^* from Chart 6A.

Matching model: rise in unemployment benefits

Chart 6A
Wages and vacancy-unemployment ratio

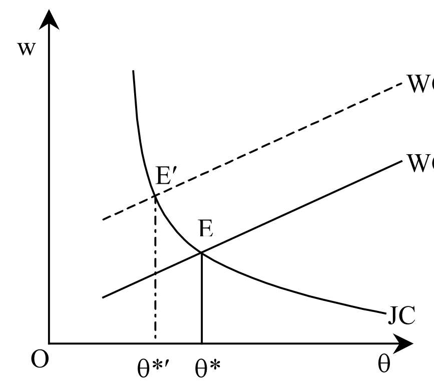
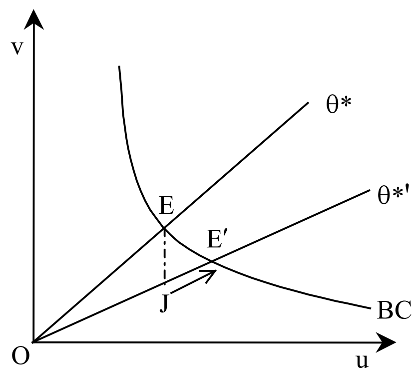


Chart 6B
Vacancies and unemployment



It is now quite straightforward to show how, following a shock to one of the exogenous parameters, the labour share will jump immediately to its new steady-state value. Perhaps the simplest shock to consider is an increase in unemployment benefits, z. Such an increase only affects the wage-setting curve, which moves from *WC* to *WC′*. The new equilibrium is at *E′*, with a higher real wage and a lower vacancy-to-unemployment ratio, $\theta^{*\prime}$. Neither w, nor θ are assumed to be sticky (unemployment is sticky, but vacancies can jump). For that reason, the vacancy-unemployment ratio moves immediately to $\theta^{*\prime}$. Looking now at Chart 6B, the equilibrium θ line shifts from θ^* to $\theta^{*\prime}$. In the short run, firms cut vacancies sharply and the economy jumps from *E* to *J*. Over time, unemployment and vacancies both rise as we move back towards the Beveridge curve at *E′*.

Implications for labour market tightness

In the basic search model, the vacancy-to-unemployment ratio is an important determinant of the labour share. Nevertheless, while unemployment adjustment is sluggish (at a rate dependent on the matching technology), the number of vacancies can jump. For that reason, neither the vacancy-to-unemployment ratio nor the labour share are ever away from their steady-state value, and we would argue that the labour market can never be tight nor loose. Modifications could be made to the basic search model that forces the labour share away from its steady-state value for short periods of time. One such modification, the addition of a limited participation monetary sector, due to Cooley and Quadrini (1999) is considered in Section 3. A simpler modification, suggested to us by Chris Pissarides, would be to allow the cost of posting a vacancy to vary procyclically.

2.6 Summary of model implications

This section has considered several popular labour market models with the aim of identifying their implications for the drivers of labour market tightness. Our analysis is based on the idea that a tight (loose) labour market is one in which there is an excess demand for (supply of) labour which causes the labour share of income to rise above (fall below) its steady-state value.

The key point is that one cannot make unambiguous inferences about the degree of labour market tightness by looking at changes in employment, unemployment or the real wage. Consider labour market quantities. In the case of a shock to nominal money balances in the LNJ model, a tightening of the labour market is accompanied by a jump down in unemployment, which then rises over time back to its unchanged steady-state value. In the case of an adverse shock to labour supply in the perfectly competitive or efficiency wage models with sluggish employment adjustment, a tightening of the labour market is accompanied by a rise in unemployment to a higher steady-state value. In the case of an increase in product market competition, unemployment would fall with, we would argue, no implications for labour market tightness. In terms of the response of labour market prices, across all of the models we consider, a tightening of the labour market will always be associated with an increase in real wages. In the case of an adverse shock to labour supply, that increase will be permanent. However, real wages can also increase for reasons that, in our view, are unrelated to changes in labour market tightness, for example following an increase in product market competition.

3. From labour market tightness to inflationary pressure

Labour market tightness, in our terms, describes the balance between the demand for, and the supply of, labour. So it is a real, rather than a nominal, concept. A tightening of the labour market may cause firms to pay more in real terms for a given quality of labour, but it does not follow immediately that there are consequences for the general price level. Consider the matching model. Here an increase in the vacancy to unemployment ratio means that, should a wage negotiation fail, the unemployed worker can expect to find another vacant job more quickly than the firm can expect to find another unemployed worker. This causes the unemployed worker to be offered a higher real wage. If he or she suspected that the higher real wage would then be eroded by a rise in product prices, he or she would not have accepted the job offer in the first place.

While labour market tightness may be a real concept at heart, many economists would accept that a tightening of the labour market also has nominal consequences. This view is certainly implicit in the structure of the Bank of England's macroeconometric model, in which a reduction in the rate of unemployment raises real unit labour costs through the earnings equation.[(20)] In turn, this increase in real unit labour costs causes the price level to error correct towards a new higher equilibrium. This process of gradual adjustment to the price level implies a higher rate of inflation until the new equilibrium price level is reached.

It is now widely recognised that shocks to the real side of the economy, which primarily affect variables like output and employment, can have a secondary and temporary influence on nominal variables, such as the rate of inflation, if there are sufficient nominal rigidities.[(21)] Among the popular labour market models we considered in Section 2, only the one due to Layard, Nickell and Jackman (1991) explicitly incorporates such rigidities in the form of sticky price expectations. More recently, a number of authors have added alternative forms of nominal rigidity to different labour market models with qualitatively similar results.[(22)] Here we consider two frameworks. The first, due to Cooley and Quadrini (1999) is applied to a conventional matching model. The second form of nominal rigidity, due to Gali, Gertler and Lopez-Salido (2000), is agnostic about the structure of the labour market.

[(20)] More precisely, the earnings equation says that the log of nominal earnings is equal to the log of the price level minus the log of productivity minus some multiple of $u - u^*$ in the long run. This is observationally equivalent to world where $u - u^*$ determines real unit labour costs in the long run. See Bank of England (1999) for further details.

[(21)] See, for example, Ball, Mankiw and Romer (1988).

[(22)] Astley and Yates (1999) argue that several popular reduced-form Phillips curves that use different forms of nominal rigidity (ie staggered wage contracts) to link real shocks to inflation are based upon weak identification restrictions and are at odds with optimising behaviour.

Cooley and Quadrini describe a dynamic stochastic general equilibrium model with a labour market characterised by matching frictions. The nominal rigidity emerges from a limited participation model of money, similar to the one described by Christiano, Eichenbaum and Evans (1997) and used by Dhar and Millard (2000a, 2000b). The term 'limited participation' comes from the fact that households cannot participate fully in financial markets. Specifically, they must pay a cost whenever they add to, or subtract from, their deposits held at banks. Consider a monetary expansion brought about by open market operations. Banks' holdings of cash rise, as their holdings of bonds fall. Because it will take time for households to adjust their deposits, there will be some persistent increase in the amount of loanable funds. This, in turn, means there will be some persistent downward effect on the rate of interest. The reduction in the rate of interest makes it cheaper for firms to hire labour.(23) So immediately following the shock, V jumps up. Unemployment is sluggish, so θ rises and the labour market tightens. Eventually, the vacancies are filled, employment and both output rise. Ultimately, the full effect of the monetary expansion comes through in prices and output falls back to its steady-state value. The authors claim that this kind of framework provides micro-foundations for a Phillips curve relationship. For our purposes, it appears to provide a link between labour market tightness (as defined by the vacancy to unemployment ratio), the labour share and inflation.

In a recent paper, Gali, Gertler and Lopez–Salido (2000) derive an equation that relates current inflation to the present discounted value of expected future real marginal costs (which, given Cobb-Douglas technology, are a constant multiple of the labour share).(24) The equation is derived from an optimising model with many imperfectly competitive firms who can each change their price only with some constant probability $(1 - \phi)$ per period. This nominal rigidity, due to Calvo (1983), means that the aggregate price level adjusts slowly towards its long-run equilibrium in such a way that inflation appears to be related to expected real marginal cost.

The reset price (p_t*), defined as the price chosen by those who win the right to change price, is given by:

$$p_t^* = \log\left(\frac{\varepsilon}{\varepsilon - 1}\right) + (1 - \beta\phi)\sum_{k=0}^{\infty} (\beta\phi)^k E_t \left\{\log MC_{t,t+k}^n\right\} \qquad \textbf{(3.1)}$$

(23) By assumption, wages have to be paid before anything is produced, forcing firms to borrow the full amount from financial intermediaries.

(24) Real marginal costs are equal to the real wage times the inverse of the marginal product of labour. Under Cobb-Douglas technology this is given by (W/P) * ($1/\alpha$) * (N/Y), which is a constant multiple ($1/\alpha$) of the labour share. Gali, Gertler and Lopez-Salido (2000) do not explicitly model the labour share.

where ε is a parameter from a constant elasticity of substitution demand function, β is a discount rate and $MC^{n}_{t,t+k}$ is nominal marginal cost at time $t + k$ of those who last reset their price at time t. **(3.1)** says that the optimal reset price is equal to the present discounted value of expected future nominal marginal costs. So far, this looks like a conventional, if forward-looking, pricing equation. But, since only a fraction (1 - ϕ) of firms can change their price each period, we know the aggregate price level (p_t) is given by:

$$p_t = (1-\phi)p_t^* + \phi p_{t-1} \qquad \textbf{(3.2)}$$

Taking **(3.1)** and **(3.2)** together, it is possible to derive the following equation for inflation:

$$\pi_t = \lambda \sum_{k=0}^{\infty} \beta^k E_t\{\hat{mc}_{t+k}\} \qquad \textbf{(3.3)}$$

where λ is a constant that depends on the underlying of the model such as the share of firms that can adjust prices each period (ϕ). and $\hat{mc}_{t+k}$ is the log deviation of real marginal costs from steady state at time $t+k$. **(3.3)** says that inflation depends on current and expected future real marginal costs, or equivalently, on current and expected future labour shares.

Why should inflation be a function of this real quantity? The intuition is as follows. When the labour share is above its steady-state value ($\hat{mc}_t$ is positive), each firm tries to claw back some of their 'lost' output by raising their product price. Without the nominal rigidity, they could do this immediately. But we have assumed that only a fraction (1 – ϕ) of firms can change their price each period. Hence the aggregate price level rises slowly, implying a sustained period of inflation, as the labour share is pushed back down to its steady-state value. It *appears* that inflation is caused by the deviation of the labour share from its steady-state value. But this is more apparent than real since **(3.3)** is best viewed as a reduced-form, rather than a behavioural, equation.

Of course this result hinges crucially on the behaviour of the monetary authority. A positive gap between the labour share and its equilibrium value can be eliminated by, either, a fall in nominal wages, a rise in prices, or a rise in productivity, or some combination of all three outcomes. If we assume that nominal wages cannot decline and that the price level cannot jump, then the gap will be eliminated by a gradual increase in the price level (ie a period of inflation) as envisaged by Gali, Gertler and Lopez-Salido (2000), or by a rise in productivity. If the monetary authority does nothing, then inflation is the likely result. However if monetary policy is tightened then output and employment will decline, which, with Cobb-Douglas production, leads to a rise in productivity and a fall in the labour share back to equilibrium.

If we are interested in capturing the implications of labour market tightness for product price inflation, then the paper by Gali, Gerlter and Lopez-Salido (2000), which makes no assumptions about the structure of the labour market, suggests that the labour share may not be a bad place to start. It also lends weight to our belief that the labour market can only be considered tight or loose out of steady state. Permanent shifts in the labour share do not cause firms to seek to alter their product prices, and hence should not be taken to signify a change in labour market tightness.

The purpose of this section has been to highlight that labour market tightness is inherently a real phenomenon. At the same time a change in labour market tightness can have implications for the rate of inflation if there are nominal rigidities in the economy. We considered three such rigidities: sticky price expectations, limited participation of households in financial markets, and Calvo pricing contracts. There is a substantial literature surrounding the empirical plausibility of different kinds of nominal rigidity and we do not intend to enter that debate here.

4. Conclusions

The aim of this paper has been to provide a coherent framework for examining the drivers of labour market tightness, and the relationship between labour market tightness and inflation.

A necessary first step in addressing these issues is to specify what is meant by 'labour market tightness'. We define it in terms of the implications it carries for the labour share. This stems from the idea that a tight (loose) labour market involves an imbalance between the demand for and supply of labour, which will put upward (downward) pressure on real unit labour costs, or equivalently on the labour share of income.[25] One advantage of looking at the labour share is that its response to a change in tightness is unambiguous. One disadvantage is that the labour share will occasionally move for reasons that have nothing to do with labour market frictions.

Because the words tight and loose imply a degree of imbalance, we assert that the labour market can only be considered tight or loose out of steady state. This has two important implications. First, no shock can cause the labour market to become tight or loose unless it pushes the labour market away from its steady state. In practice, this is not too restrictive, since the kinds of rigidity that are present in most popular macro-models are sufficient to do this. Second, any shock which alters the

[25] One possible manifestation of this imbalance is firms find it harder (easier) to recruit and retain staff.

steady-state value of the labour share cannot be said to have made the labour market permanently tighter or permanently looser. This is because movements in the steady state do not involve any change in the balance between the demand for and supply of labour.

With our definition of labour market tightness and its steady-state properties in mind, we then investigate the implications for labour market tightness of several popular labour market models. We start with the basic competitive model, and then work through models of efficiency wages, insider power, skill mismatch and matching frictions. One key message is that the implications for a given set of conventional wage and employment indicators for changes in labour market tightness depend on the underlying economic shock, and on the nature of the rigidities assumed. For example, in the case of a shock to nominal money balances in the LNJ model, a tightening of the labour market is accompanied by a jump down in unemployment, which then rises over time back to its unchanged steady-state value. In the case of an adverse shock to labour supply in the perfectly competitive or efficiency wage models, a tightening of the labour market is accompanied by a rise in unemployment to a higher steady-state value. In the case of an increase in product market competition we would argue that unemployment would fall, without implications for labour market tightness.

The discussion thus far has been in terms of real economic magnitudes. Other things equal, a tightening of the labour market will cause the labour share of income to rise. But what about inflation? The key point here is that, for labour market tightness to have nominal consequences, some kind of nominal rigidity must be imposed. Examples of such rigidities that appear in the literature include sticky price expectations, and restrictions on the frequency with which firms can change prices. With such rigidities, it is possible for out of steady-state movements in the labour share to affect inflation.

There are several areas for further research. One potential drawback of all the models considered is that they only allow workers to be either employed or unemployed. Extending the models to allow for inactivity would facilitate an examination of the tightness impacts of recent initiatives aimed at increasing activity rates, such as the Working Families Tax Credit. Evidence on the empirical determinants of the labour share would also be useful. If movements in the share are primarily driven by shocks that originate in the labour market, then it is arguable that the labour share can be usefully employed as a tightness indicator. However, if shifts in the labour share largely reflect non-labour market factors, such as shifts in import prices (Bentolila and Saint-Paul (1999)), then its link to movements in labour market tightness may be weak.

References

Astley, M S and Yates, T (1999), 'Inflation and real disequilibria', *Bank of England Working Paper no. 103*

Ball, L, Mankiw, N G and Romer, D (1988), 'The new Keynesian economics and the output-inflation trade-off', *Brookings Papers on Economic Activity*, Vol. 1, pages 1-65.

Bank of England (1999), *Economic models at the Bank of England*, Bank of England, London.

Batini, N, Jackson, B and Nickell, S (2000), 'Inflation dynamics and the labour share in the UK', *External MPC Unit Discussion Paper no. 2.*

Bentolila, S and Bertola, G (1990), 'Firing costs and labour demand: how bad is Eurosclerosis?', *Review of Economic Studies*, Vol. 57, pages 381-402.

Bentolila, S and Saint-Paul, G (1999), 'Explaining movements in the labour share', *CEMFI Working Paper, no. 9905.*

Blanchard, O (1997), *Macroeconomics*, Prentice-Hall, New Jersey.

Blanchard, O and Katz, L (1997), 'What we know and do not know about the natural rate of employment', *Journal of Economic Perspectives*, Vol. 11, pages 51-72.

Calvo, G A (1983), 'Staggered prices in a utility maximising framework', *Journal of Monetary Economics*, Vol. 12, pages 383-98.

Christiano, L J, Eichenbaum, M and Evans, C (1997), 'Sticky price and limited participation models of money: a comparison', *European Economic Review*, Vol. 41, pages 1,201-49.

Coles, M and Smith, E (1996), 'Cross-section estimation of the matching function: evidence from England and Wales', *Economica*, Vol. 63, pages 589-98.

Cooley, T F and Quadrini, V (1999), 'A neoclassical model of the Phillips curve relation', *Journal of Monetary Economics*, Vol. 44, pages 165-93.

Dhar, S and Millard, S P (2000a), 'A limited participation model of the monetary transmission mechanism in the United Kingdom', *Bank of England Working Paper no. 117.*

Dhar, S and Millard, S P (2000b), 'How well does a limited participation model of the monetary transmission mechanism match UK data?', *Bank of England Working Paper no. 118*.

Gali, J, Gertler, M and Lopez-Salido, J (2000), 'European inflation dynamics', *mimeo,* New York University.

Jackman, R, Layard, R, Manacorda, M and Petrongolo, B (1999), 'European versus US unemployment: different responses to increased demand for skill?', in Layard, R (ed), *Tackling unemployment,* MacMillan Press Ltd, London.

Layard, R, Nickell, S and Jackman, R (1991), *Unemployment: macroeconomic performance and the labour market*, Oxford University Press, Oxford.

Lilien, D (1982), 'Sectoral shifts and cyclical unemployment', *Journal of Political Economy*, Vol. 90, pages 777-93.

Manacorda, M and Petrongolo, B (1999), 'Skill mismatch and unemployment in OECD countries', *Economica*, Vol. 66, pages 181-207.

Manning, A (1993), 'Wage bargaining and the Phillips curve: the identification and specification of aggregate wage equations', *The Economic Journal*, Vol. 103, pages 98-118.

Nickell, S (1986), 'Dynamic models of labour demand', in Ashenfelter, O and Layard, R (eds), *Handbook of labor economics*, Elsevier Science, Amsterdam.

Nickell, S and Bell, B (1995), 'The collapse in the demand for the unskilled and unemployment across the OECD', *Oxford Review of Economic Policy*, Vol. 11, pages 40-62.

Pissarides, C A (1998), 'The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and the tax structure', *European Economic Review*, Vol. 42, pages 155-83.

Pissarides, C A (2000), *Equilibrium unemployment theory*, The MIT Press, London.

Shapiro, C and Stiglitz, J (1984), 'Equilibrium unemployment as a worker discipline device', *American Economic Review*, Vol. 74, pages 433-44.

Appendix

The wage bargain

The wage in firm i, W_i, is determined by a Nash bargaining process. More precisely, W_i is the wage rate that solves:

$$\begin{aligned} \max \Omega &= \left(V_i(W_i) - \overline{V}_i\right)^{\beta} \Pi_{o,i}(W_i) \\ \text{where } V_i(W_i) &= W_i S_i + (1 - S_i)A \\ \text{and } A &= (1 - \varphi u)W^e + \varphi u B \end{aligned} \qquad \textbf{(A1.1)}$$

The net gain to the insiders from reaching agreement with the employer is the difference between the value of being employed at wage W_i, which we write as $V_i(W_i)$, and his or her strike income, which we write as $\overline{V}_i$. The value of being employed at a wage W_i is equal to that wage multiplied by the probability that the insider continues to be employed at that wage (S_i), plus the value of alternative income (A) multiplied by the probability that the insider is laid off. A in turn is a weighted average of unemployment benefits, B, and the expected outside wage, W^e. The weights depend on the probability that an individual is rehired, and hence on the unemployment rate, u.

Let $\Pi_{o,i}$ denote firm i's operating profits, defined as revenue net of variable costs *(PY - WN)*. This definition of operating profit is convenient, since we know that $\Pi_{o,i}$ will be zero in the event of a strike (whereas revenue net of *all* costs would be equal to $-rK$). Hence the net gain to the employer of reaching agreement is all of the profit that is made, namely $\Pi_{o,i}(W_i)$. β is a measure of the relative bargaining power of the insiders.

We follow Layard, Nickell and Jackman (1991) in assuming that strike income ($\overline{V}_i$) is the same as alternative income *(A)*. Using this in **(A1.1)** our problem becomes:

$$\max \Omega = (W_i - A)^{\beta} S(W_i)^{\beta} \Pi_{o,i}(W_i) \qquad \textbf{(A1.2)}$$

Taking logs of the objective function and differentiating with respect to the wage we obtain the first-order condition:

$$\frac{\partial \ln \Omega}{\partial W_i} \equiv \frac{\beta}{W_i - A} + \frac{\beta}{S_i} \cdot \frac{\partial S_i}{\partial W_i} - \frac{N_i}{\Pi_{o,i}} = 0 \qquad \textbf{(A1.3)}$$

Rearranging gives:

$$\frac{W_i - A}{W_i} = \frac{1}{\varepsilon_{sw}(W_i) + [\beta\gamma(W_i)]^{-1}} \tag{A1.4}$$

where the function $\gamma(W_i) = \Pi_{o,i}/(W_iN_i)$ gives the ratio of operating profits to workers' income, and ε_{sw} is the absolute value of the elasticity of the survival productivity with respect to the wage. We can obtain an expression for operating profit and hence the function γ from the price-setting rule **(2.2.6)** in the main text.

$$\begin{aligned}\Pi_{o,i} &= P_iY_i - W_iN_i \\ &= \left(\frac{1}{\alpha\kappa}\right)W_i\left(\frac{N_i}{Y_i}\right)Y_i - W_iN_i \\ &= \left(\frac{1-\alpha\kappa}{\alpha\kappa}\right)W_iN_i\end{aligned} \tag{A1.5}$$

$$\text{Hence } \gamma(W_i) = \left(\frac{1-\alpha\kappa}{\alpha\kappa}\right)$$

It is convenient to write ε_{sn} as the product of the elasticity of the survival probability with respect to employment (ε_{sn}), and the elasticity of employment with respect to the wage rate (ε_{nw}). It can be shown that ε_{sn} is a function of the ratio of the number of insiders in firm *i* (N_{Ii}) to expected employment in firm *i* (N_i^e). To derive ε_{nw} we must first find an expression for the optimal quantity of labour (N_i^*) as a function of the wage rate (W_i). By substituting the supply and demand functions — **(2.2.1)** and **(2.2.2)** in the main text — into an expression for operating profit we obtain:

$$\Pi_{o,i} = P \cdot N_i^{\alpha(1-1/\eta)} K_i^{(1-\alpha)(1-1/\eta)} Y_d^{1/\eta} - W_iN_i \tag{A1.6}$$

Differentiating with respect to N_i and setting the resulting expression equal to zero gives:

$$N_i* = \frac{1}{\alpha(1-1/\eta)}\left[W_iY_d^{-1/\eta}K_i^{-(1-\alpha)(1-1/\eta)}P^{-1}\right]^{-1/(1-\alpha\kappa)} \tag{A1.7}$$

Using **(A1.7)**, the elasticity of employment with respect to the wage is given by:

$$\varepsilon_{nw} = \left| \frac{\partial N_i}{\partial W_i} \cdot \frac{W_i}{N_i} \right| = \frac{1}{1-\alpha\kappa} \qquad \textbf{(A1.8)}$$

We are now in a position to derive an expression for wages as a function of the unemployment rate and structural parameters of the model. To do this we take the expression for the wage mark-up over alternative income in **(A1.4)** substituting for A from **(A1.1)**, for the function γ from **(A1.5)** and for ε_{nw} from **(A1.8)** we obtain:

$$\frac{W-B}{W} = \frac{1-\alpha\kappa}{\left[\varepsilon_{sn}\left(N_I / N_i^e\right) + \alpha\kappa/\beta\right]\varphi u} - \left(\frac{1}{\varphi u} - 1\right)\left(\frac{W-W^e}{W}\right) \qquad \textbf{(A1.9)}$$

To move from **(A1.9)** to the wage-setting function **(2.2.1)** we note that **(A1.9)** can be rewritten:

$$\frac{B}{W} = 1 - \frac{1-\alpha\kappa}{\left[\varepsilon_{sn}\left(N_I / N_i^e\right) + \alpha\kappa/\beta\right]\varphi u} + \left(\frac{1}{\varphi u} - 1\right)\left(\frac{W-W^e}{W}\right) \qquad \textbf{(A1.10)}$$

Taking logs of both sides, and using the log approximation that log(1+z) ˜ z when z is small, we can derive:

$$w - p = b - p + \frac{1-\alpha\kappa}{\left[\varepsilon_{sn}\left(N_I / N_i^e\right) + \alpha\kappa/\beta\right]\varphi u} - \left(\frac{1}{\varphi u} - 1\right)\left(\frac{W-W^e}{W}\right) \qquad \textbf{(A1.11)}$$

(A1.11) says that the bargained real wage will be a mark-up over real benefits. This mark-up is decreasing in the rate of unemployment (u) and product market competitiveness (κ) but increasing in trade union power (β).