B) Macroeconomic theories of unemployment and the “European labour market model”

B1) Introduction
We try to understand the structure of unemployment rates across space in relation to the phenomenon of regional agglomeration. We thus need theoretical background from various fields in economics. We need a theoretical framework to analyse the issue of unemployment, with special reference to its regional dimension, a theory of regional agglomeration, and finally we need to combine the two.

We shall at first start with the theory of unemployment in this chapter. In the introduction it was already said that our analysis will build on a specific regional labour market approach that was pioneered by Blanchflower/Oswald (1990, 1996), and that is comprehensively labelled the wage curve. This regional theory is closely linked to one specific macroeconomic approach which is known as the “European labour market model (ELMM)” or the “imperfect competition approach to macroeconomics”.

The purpose of this chapter is to introduce this macroeconomic model, since many essential features of the wage curve theory are built directly on it. In order to demonstrate how the ELMM relates to the long history of economic thought in macroeconomics, we intend to give at first a very brief overview of major developments in macroeconomics over the last century or so, with a special emphasis on theoretical explanations for the phenomenon of unemployment. This historical sketch that is provided in section B2 is by no means an attempt to survey all that has been done in this field. It only tries to present some essential ideas of a few major contributions in a very simplified manner. With this historical information it is easier to see what is actually new about the “imperfect competition approach” to macroeconomics which is then presented afterwards in section B3.

B2) A brief historical overview about macroeconomics

B2.1) The ‘classics’
The very traditional perspective in macroeconomics comes from the so called ‘classics’. In their view, the real and the monetary sphere of an economy can be completely separated from each other. Inflation is interpreted as a purely monetary phenomenon, with the rate of inflation given only by the growth rate of money supply. The real sector on the other hands is seen as a system of perfectly competitive markets which all tend to clear because of the invisible hand of the market.
B2.2) Keynes and the neoclassical synthesis

This consensus view of the economy as a whole, however, was put heavily under strain with the occurrence of the Great Depression in the late 1920s. It was impossible to explain this event with the traditional economic models at hand, since they were not ready to address why there could be persistent and involuntary mass unemployment in an economy. In this historical period, the macroeconomic theory of Keynes (1936) appeared. The contribution of Keynes was to show that economies can be in an equilibrium where full employment is not reached, because the prevailing level of effective demand is insufficient to render full utilization of existing production capacities. In other words, there exists the possibility of involuntary unemployment in the economy.

It is still subject to ongoing and possibly never ending debates what exactly is at the root of involuntary unemployment in the view of Keynes himself. But the most influential interpretation following the General Theory, that came to dominate macroeconomic theory and policy for a very long period of time, was the neoclassical synthesis. This Keynesian theory is crucially based on nominal wage rigidity, i.e. market failures on the labour market. The demand side in this model was modelled in the fashion of the famous IS/LM-model (Hicks, 1937). Economists then added a neoclassical production function and a standard neoclassical labour market to the analysis, but (in the name of Keynes) imposed nominal wages to be downwardly rigid. This was then seen as the major source of unemployment, since real wages were prevented to adjust to market clearing levels by these nominal rigidities. The important implication for economic policy that comes from this well-known model is that the government can effectively achieve increases in production and employment through demand side policies (i.e. expansionary fiscal or monetary policy). As long as workers (or respectively, unions) are only concerned with nominal wages in a situation in which involuntary unemployment exists, the government can stimulate aggregate demand and hereby increase the price level. With nominal wages given, this translates into falling real wages and induces an increase in economic activity, possibly to a level where full employment is restored.

If taken further, an interpretation of the (modified) Phillips-curve in the pure vein of the neoclassical synthesis can be based on this logic: with nominal wage inertia, the government can determine the price level through inflationary policies, thereby the real wage and hence the level of employment in this economy. In other words, the government is subject to a downward sloping schedule in the

---

2 For a comprehensive representation of the model see Jarchow (1998), chapter IV.
3 For this interpretation of the Phillips-curve within the neoclassical synthesis see Carlin/Soscice (1990), ch. I.2; Felderer/Homburg (1993), S. 265 ff.
inflation/unemployment rate-space from which it can choose a desired combination. Although the concept of the Philipps-curve traditionally has been introduced in a different manner into the Keynesian system, 4 this simplifying notion illustrates the main channels of the neoclassical synthesis. Effectively the model predicts that quantity adjustments were needed in this economy to restore equilibrium, because of sluggish nominal variables. Thus, if demand from the private sector is insufficient to render full capacity utilization and prices can not adjust so as to restore equilibrium, the government can step in and take maintenance for low unemployment at the cost of higher inflation.

B2.3) Friedman and the ‘natural rate of unemployment’
This neoclassical synthesis was widely accepted in economic theory and policy during the first 20 to 25 years in the post-war period. However, the Keynesian orthodoxy was gradually challenged by political and economic developments and by advances in economic theory. In the beginning of the 1970s, major industrial countries saw inflation and unemployment rise simultaneously in the aftermath of the oil crisis. This development, called stagflation, was inconsistent with the conventional Phillips-curve. But prior to this, several scholars have criticised the Keynesian orthodoxy also from a theoretical point of view. In particular Milton Friedman (1968) emphasised that the validity of the Keynesian system effectively rested on money illusion by the workers. In the mechanism outlined above, expansionary fiscal or monetary policy drives up the price level, i.e. it causes inflation. This has only positive employment effects as long as nominal wages are rigid, so that real wages fall and restore full employment. But why should nominal wages remain constant in view of positive inflation? This would require that workers do not recognize the inflation pressure brought about by the government policy, and that they base labour supply decisions on nominal rather than on real wages. Friedman’s argument was that this confusion of nominal with real wages can only be true in the short run, because workers might have incomplete information about inflation and might thus do not instantaneously perceive a fall in real wages.

4 Usually, the Phillips-curve was introduced in a pure ad-hoc way based on Phillips’s (1958) empirical observation that the level of unemployment and growth rate of money wages was negatively correlated. The theoretical concept of the modified Phillips-curve included mark-up pricing (normal costs pricing) on the firms part. If firms set prices simply as mark-ups over wage costs, and wage increases depend negatively on the rate of unemployment (see Lipsey, 1960), a negative correlation between the unemployment rate and the rate of inflation is implied (Samuelson/Solow, 1960). Thus, the rate of price and the rate of wage increases were identical, but the nominal wages is not necessarily rigid. For the introduction of an ad hoc Phillips-curve in this spirit into the neoclassical synthesis see Carlin/Soscice (1990:69 ff).
Once they do, however, workers will adapt their nominal wage claims so as to restore the level of real wages that was prevailing before the policy intervention. Unemployment will thereby also move back to its old level. Effectively, all that has changed is that prices and nominal wages are now on a higher level after the political intervention.

In the long run, there is no trade-off between inflation and unemployment, but rather a vertical Philipps-curve above some unemployment rate that Friedman called the 'natural rate of unemployment'. This natural rate according to Friedman (1968) is "[...]the level which would be ground out by the Walrasian system of general equilibrium equations, provided that there is imbedded in them the actual structural characteristics of the labour and commodity markets, including market imperfections, stochastic variability in demands and supplies, the cost of gathering information about job vacancies and labour availabilities, the costs of mobility, and so on". In the long-run, unemployment is thus determined in the "classical" way, i.e. on a competitive labour market where unemployment is either voluntary or due to rigidities and the malfunctioning of labour market institutions. With the use of macroeconomic demand management policies, policymakers can keep unemployment below this natural level only at the cost of steadily accelerating inflation, which led to the name of a "non-accelerating inflation rate of unemployment (NAIRU)" as a synonym. Put differently, demand side policies can not be used systematically to reduce unemployment, they are neutral in the long-run and have only short term real effects when agents confuse nominal and real wages.

B2.4) 'New classical macroeconomics' and rational expectations
The short term effects of monetary and fiscal policy in the Friedman-model were due to the workers' temporary misconception of nominal and real wages. This was brought about by systematically wrong expectations about inflation, which in the Friedman-model were formed in an adaptive way. In the 1970s and 1980s, however, the 'new classical macroeconomics' appeared that questioned this construction and urged for a more radical reassessment of macroeconomics.

Economic agents build expectations about future developments. Since the future is by definition uncertain, no agent can make predictions that are always absolutely correct, as there is always the possibility for stochastic and unexpected events. But the proponents of the 'new classical macroeconomics', most notably Robert Lucas, Thomas Sargent, and Edward Prescott, believed that agents will not make systematic errors. Friedman's use of adaptive expectations that left room for short run policy effects was replaced with the construction of rational expectations. This means that agents incorporate all available information that is at their disposal at any moment into their expectation formation.

The consequence is that workers, when faced with inflationary government policies, would not misperceive nominal and real wages, but would correctly expect
falling real wages. Consequently, there is no room for the sort run mechanism à la Friedman, but inflationary policies are offset right away by changes in nominal wages in order to maintain the prevailing level of real wages. Put differently, because of rational expectations, output and employment in the models of new classical macroeconomics are permanently at the natural rate, except maybe for random disturbances (Lucas, 1972). The inherently stable economic system leaves no room for stabilization policy. Any attempt by policymakers to influence the output level via macroeconomic policies would only lead to inflation, or to a complete crowding out of private activities.

The 'new classical macroeconomics' can be viewed at as a modern and mathematically more articulate version of the pre-Keynesian classical economics. Both rest essentially on Walrasian general equilibrium theory with perfect competition and instantaneous price adjustments mandated by an imaginary auctioneer, and the 'new classical' school has added rational expectations to this model. As it is well known, the Walrasian price tâtonnement process leads to the clearing of all markets in the economy, including the labour market.

The principal problem of the 'new classics' is thus more or less the same as it has been for the 'old classics'. General equilibrium theory needs to answer why there can be business cycle fluctuations in a model where output is always at its natural level and agents have rational expectations. How can the observed short-run fluctuations of output and employment be explained in such an environment? Similarly, also the 'new classical macroeconomics' would have to come up with a convincing story why something like the Great Depression can unravel when markets work that perfectly.

One way to deal with business cycle fluctuations in the context of 'new classical macroeconomics' was developed in the early 1980s and became known as "real business cycle theory (RBC)". RBC-models, pioneered by Kydland/Prescott (1982), are market clearing models with agents who form rational expectations. Fluctuations in output and employment are seen as resulting from exogenous technological shocks hitting the production function, and thereby the marginal productivity of labour. Agents in these models decide on intertemporal labour supply paths by equating the obtainable wage rate (the marginal product of labour) with the value of leisure. In view of adverse technological shocks, agents may find it optimal to engage in leisure or non-market activities, as falling wages can fall short of reservation wages. A typical RBC-model thus implies that output is always at its natural level, and this level fluctuates due to technological shocks. These models are therefore often called "equilibrium business cycle models".

Any unemployment is entirely voluntary and simply determined by the optimal intertemporal allocation of time. But the use of the term "voluntary unemploy-

---

5 This puzzle was already noticed by Hayek (1933), p. 33 for the old general equilibrium theory.
6 For this debate see e.g. Lucas (1980), Tobin (1977) and Modigliani (1977).
ment” needs some further classification. Recorded unemployment rates by definition only include such persons who actively seek for jobs, i.e. no persons who have chosen not to belong to the labour force. However, this critique does not necessarily flaw the predictions of RBC-models and all other classical models that view unemployment as “voluntary”. Firstly, these models might refer to a broader measure of joblessness (e.g. the “non-employment rate”) that also includes labour force drop-outs that in principal are interested in accepting a job, but not at going wage rates. And secondly, unemployed workers in reality have some degrees of freedom to signal job search activities, even though they are currently not interested to work. The main implication of classical and RBC models is simply that unemployment can not be viewed at as a phenomenon where individuals would be willing to accept jobs at going wage rates, but receive no job offers and thus remain unwillingly unemployed.

B2.5) The Keynesian response

The ‘new classical’ model of macroeconomics as well as RBC-models make extreme assumptions with respect to the rationality of agents and their state of knowledge, but also with respect to the functioning of markets. In particular, the models rest very heavily on the instantaneous price adjustments brought about by the Walrasian auctioneer.

The first wave of Keynesian criticism against the new classical model was mainly concerned with the artificiality of the price tatonnement. What are the implications for the economy if this process is not functioning? Recall that in theory the Walrasian auctioneering process works such that some initial price vector for all commodities in the economy is announced, all market participants signal the quantities they are willing to buy and sell at these prices, the auctioneer calculates excess demands and excess supplies on all markets, and calls out a new pricing vector that brings quantitative supply and demand closer together. This process is repeated until the market clearing price vector is found. Only then do transactions take place, at the equilibrium prices that equilibrate supply and demand on all markets.

Criticism against the artificiality of this theoretical construct was articulated by Clower (1965) and Leijonhufvud (1967, 1968) even before the emergence of the ‘new classical’ macroeconomics, with the original intention of “re-interpreting Keynes” in opposition against the neoclassical synthesis. Their works, however, where the precursors to much of the Neo-Keynesian critique in the 1970s, e.g. by Malinveaud (1977). Central to this literature was the analysis of economic systems when an auctioneer is absent and transactions take place at “wrong” prices.7

---

7 For a comprehensive overview of this school of Neo-Keynesian economics see Felderer/Homburg (1993:287ff.) or Carlin/Sosicice (1990:106ff.)
Clower (1965) has emphasised that it is highly artificial to assume that transactions only take place once the market clearing price vector is announced. In reality, there is continuous trading also at non-market clearing prices. Because markets do not instantaneously clear, agents face quantity constraints on some markets, which via a budget constraint also affects economic plans with respect to other markets. For example, if an initial price vector exists where the wage rate is higher than the theoretical market clearing level, agents can not sell the desired amount of labour, but rather face a quantity constraint on the labour market. Nevertheless, trading already starts at this pricing vectors and agents do not wait for an imaginary auctioneer to calculate market-clearing prices. Because agents supply less labour than intended, they can also not realize their desired consumption plans. They are forced to demand a lower quantity of goods, because actual income is below desired income. This is the essence of the “dual decision hypothesis”. Agents first form desired economic plans. After having perceived quantity constraints, they have to adjust to market conditions and choose their actual economic plans. These actual plans are the optimal household decisions subject to the quantity constraints at the given non market-clearing price vector.

The critical question is then whether there exist market forces that realign prices such that agents can gradually come to realize their desired economic plans. As shown by Dreze (1975), this is difficult or even impossible if there is no auctioneer and all agents are atomistic. In other words, an economy might be trapped in a sub-optimal equilibrium where markets do not clear, because the Pareto-optimal Walrasian equilibrium is unknown to single agents and an adjustment process can not be triggered.

Leijonhufvud (1967) emphasised that at best it takes a considerable amount of time for an economy to converge from a Keynesian equilibrium with quantity constraints towards the Walrasian general equilibrium. During the transition, there is room for government action, including the exploitation of multiplier effects through stabilization policy.

To sum up, this first wave of Neo-Keynesian economics with the so-called “fixed-price models” was concerned with the analysis of non-Walrasian equilibria. In particular, involuntary unemployment can follow in these models, just as there is room for macroeconomic policy under some circumstances.
atomistic agents and perfect competition on goods and labour markets. The bottom-line message was that within this very conventional market system, economies can still end up in sub-optimal equilibria with unemployment, because price adjustments do not work well.

**B3) The “European labour market model (ELMM)”**

During the late 1980s, a different macroeconomic approach was developed mainly in Europe that was popularised under the name ELMM, or “imperfect competition approach to macroeconomics”. It is generally considered to be an approach in the Keynesian tradition, which is supported by the following citation that comes from one of its most famous adherents

“My interpretation of the empirical evidence is that the magnitude and persistence of changes in statistically recorded unemployment are too large to be explained as variations in search or frictional unemployment, intertemporal substitution of leisure or a misinterpretation among economic agents regarding inflation or relative price and wage changes in the context of market-clearing models. The apparent unhappiness of many unemployed workers do not suggest that they have simply, in an optimal fashion, reallocated leisure in response to perceived temporal or intertemporal wage changes. [...] My inference from all this is that market-clearing approaches to the labour market cannot possibly be appropriate for an analysis of short- and medium-term macroeconomic developments.”

Lindbeck (1992:209f.)

Probably all Neo-Keynesians would subscribe to this viewpoint. But the new school departed from the “fixed-price model” described above, since there was a growing dissatisfaction with its micro-foundations. It appeared as if this literature has examined the properties of Walrasian systems with price rigidities, described non-Walrasian equilibria and classified 'new classical macroeconomics' as a very special case, where the prevailing price vector happens to be the market-clearing one.

But this theory did not point to the origins and sources of price stickiness.\(^\text{10}\)

Moreover and more importantly, the use of perfect competition as the reference system of markets became increasingly unsatisfactory to economists. In particular for European countries it seemed much more appropriate to acknowledge that

\(^\text{10}\) This issue was later developed in much more detail in a different string of “Neo Keynesianism” that was concerned to formulate rigid micro-foundations for price stickiness in the context of Walrasian market-clearing models (menu costs etc.). For an overview of this literature see Gordon (1990).
goods markets and especially labour markets were not working perfectly, but were rather characterised by substantial and systematic imperfections. Thus, theoretical effort of the ELMM was devoted to the formulation of an alternative economic model that was based on the concept of imperfect competition. The explicit departure from the concept of perfect competition is the important feature that distinguishes the ELMM from other theoretical macro approaches. Working with imperfect competition has the important advantage that it explicitly involves wage and price setting of economic agents. The Walrasian auctioneer is no longer needed, nor are discussions about possible deficiencies of the tâtonnement process: it does no longer exist. Furthermore, it is much more straightforward in this environment to discuss the sources of price stickiness, which might result from the market power of agents operating in imperfectly competitive markets.

B3.1) Why is there insider market power in the labour market?
The labour market does not clear at a given price level, because nominal wages (and hence real wages) do not adjust to the market-clearing level. The outsiders in the labour market are involuntarily unemployed, because it is not possible for them to replace the employed insiders by underbidding wages. Why is that? The general answer is that insiders have some degree of market power that they can use to defend their position against outsiders. Lindbeck (1992) distinguishes four broad channels why employed insiders have the power to maintain wages at higher than market clearing levels and prevent underbidding by outsiders.

a) Social norms: A first channel might be that wage underbidding is regarded by society as an unacceptable form of behaviour. The unemployed might refrain from offering labour at lower than current wages, because they fear social stigmatisation for having “stolen jobs”. Similarly, firms might also not want to accept low wage offers by outsiders, because they fear negative drawbacks on their reputation. If consumers regard the firm’s recruitment behaviour as unacceptable, they might become negatively conditioned against the products of the particular firm. Lindbeck (1992) calls these considerations the implicit eleventh and twelfth commandment, whereby individuals are told “thou shalt not steal jobs from thy comrades by underbidding their wages” and respectively firms are urged “not to encourage nor accept job theft by way of underbidding”. This view of human behaviour is of course inconsistent with the assumption of a purely selfish homo oeconomicus, since otherwise outsiders would have no problem whatsoever to underbid wages. Recent evidence from experimental economics (see e.g. Fehr/Fischbacher, 2002), however, points to the fact that individuals in the real world are also driven by non-selfish motives like social norms.
Nevertheless, it is surely insufficient to explain nominal wage rigidity, insider market-power and hence unemployment only through the honourable social behaviour of the unemployed who would rather suffer than cause wage decreases for somebody else. There must be other sources of insider market-power.

b) Union wage setting: Wages are not determined by market forces, but rather are set by trade unions, or respectively unions and employer associations bargain over wages. This type of wage determination is specifically relevant for continental Europe, where union wages typically apply to more than 80% of all employees (Ochel, 2000). Unions represent the interests and bargain on behalf of their members. It is the primary concern of unions to maximize their members’ utility function, which includes their wage income subject to the interrelated constraints on employment. Typically this optimisation process will lead unions to set wages above the market clearing level (Layard/Nickell/Jackman, 1991:83ff.). But this alone does not explain why there is no wage underbidding in the economy. Even though unions might bargain high wages for their members, non-union members in principle have the freedom to offer labour to employers at lower salaries. This has led Lindbeck (1992) to infer that union models alone do not explain why labour markets do not clear. However, this underestimates the role that unions play as political pressure groups. Unions do not only bargain about wages on behalf of their members, they also seek to prevent underbidding through political lobbying activities. In Germany e.g., bargaining results may under certain circumstances be declared as generally binding by the government for all firms and workers in a specific sector, which legally enforces union wages and prevents underbidding (Kirsch/Bispink, 2002).

In general, the high rates of bargaining coverage in continental Europe that are far greater than union membership rates indicate that union wages have a high factual compulsoriness. Outsiders are both legally and factually restrained from underbidding in many important sectors of the economy.

c) Turnover costs: A natural complement to union models is the insider-outsider-theory by Lindbeck/Snower (1986). This theory specifically stresses the source of insider market power. Employed insiders know that they can not easily be replaced with outsiders, because of various costs associated with the turnover. The most direct form of turnover costs are hiring and firing costs as

---

11 It is commonly believed that the deviation of union wages from market-clearing wages particularly depends on the level of aggregation at which bargaining takes place, because of the different abilities of a union to exter-nalise the employment consequences of its wage claim (Calmfors/Driffill, 1988).
well as training costs. But also more subtle costs might accrue to firms, if e.g. the productivity of prospective entrants is lower because old insiders engage in harassment behaviour and refuse to cooperate with new entrants. Unions, as the interest group of insiders, again might lobby for various political steps to increase turnover costs and thereby the market power of employed workers, e.g. through advocating employment protection legislation. The existence of these costs grants insiders with market power, as they can claim higher wages than outsiders. Of course this market power is not unlimited. But within a feasible range, insiders can carry through wages which on an aggregate level lead to non-market clearing on the labour market.

d) Efficiency wages: Firms might also consciously decide to pay wages above the market clearing level in order to boost the productivity or profitability of their incumbent workforce. This might be so for various reasons, highlighted by various different efficiency wage models. All models have in common that they take labour to be a very special input in the production process, hardly comparable to capital and machinery. Since workers are human beings, they are very specialized with respect to their characteristics and skills. A firm might thus look at its incumbent workforce as an asset portfolio that is worth to retain and to develop. Moreover, workers are able to vary their personal work performance. The possible range starts at zero performance, and goes up to the maximum performance that a worker can supply with given skill level and with the given capital stock at hand. Unlike machinery, workers will consciously adjust their effort in response to firm internal and external incentives. Firms can only incompletely monitor the personal performance of single workers. Within this “human” environment, the wage paid by a firm becomes an instrument to govern the performance and the motivation of the incumbent workforce (Stiglitz, 2002).

As Layard/Nickell/Jackman (1991:150ff.) put it, firms use efficiency wages in order to “recruit, retain and motivate” its workers, which refers to different particular efficiency wage models as e.g. Schlicht (1978); Salop (1979); Akerlof (1982); Shapiro/Stiglitz (1984) or Akerlof/Yellen (1990). We do not want to introduce nor survey the details of the different efficiency wage models.\(^{12}\) For the issue of wage underbidding, however, all existing models share qualitatively similar predictions: Since employers pay higher than market wages out of free will, they are not interested in accepting low wage offers by outsiders. They can use outside unemployment as a pressure device, but motivation, fairness or screening considerations (depending on the specific model) induce the firm to keep on paying higher wages. Consequently, outsiders can not enter jobs at these firms and must stay unemployed.

\(^{12}\) For an overview see e.g. Yellen (1984).
It is quite conceivable that the single arguments a)-d) are not mutually exclusive, but rather complementary. The sources of insider market-power highlighted above are at the core of the aggregate description of the labour market under imperfect competition, because they clarify how wage setting in imperfect labour markets can cause unemployment. It is now important to go also the other way around and show how the level of unemployment in turn influences the wage setting behaviour. To this issue we turn next.

B3.2) The aggregate wage setting curve
The aggregate description of the labour market in the ELMM does no longer consist only of an aggregate labour supply and an labour demand function as under perfect competition. A central relation in the ELMM is the *wage setting curve* (WS)\(^\text{13}\), illustrated in fig. B1. The curve WS is an upward sloping curve in the real wage/employment rate space. It shows all combinations of real wages and employment rate that are consistent with equilibrium in the imperfectly competitive labour market in an economy. The combination \((w/p)_1\) and \((1-U_1)\) is e.g. one such combination.

**Figure B1: The wage setting curve in the ELMM**

For simplicity we have assumed that labour supply (LS) is perfectly inelastic, so that the LS curve is vertical above the full employment situation with \((1-U_1)=1\). The unemployment rate for any given wage rate is thus given by the vertical distance between the WS and the LS curve.

\(^{13}\) There is no unique terminology for labelling the WS curve. Carlin/Soscice (1990) use the term “bargained real wage curve”. Lindbeck (1992) calls WS also simply a “wage setting curve”. Blanchard/Katz (1997) call WS a “supply wage relation” (WS). Layard/Nickell/Jackman (1991) also use the most simple terminology, wage setting curve (WS).
There are various ways to provide economic intuition for this positive slope, which correspond to the intuition about imperfect competition in the labour market that was given in section B3.1.). Two "stories" have become most common in recent years, union models and efficiency wages, although several authors (e.g. Blanchard/Katz, 1997) also base WS curves on different micro-foundations.

The first explanation for the positive slope of the WS curve rests on union wage setting and insider-outsider theory (points b and c from section B3.1.). Suppose the labour market at stake is heavily unionised. The union sets nominal wages in relation to the given or expected price level, which in the short run implies a real wage. It is quite intuitive that the union’s bargaining power and thereby the bargained real wage is a positive function of the employment rate. With high unemployment, insider power is low. Outsiders are willing to underbid wages more aggressively, and insiders can rely on turnover costs to a smaller extent. In tight labour markets, unions can negotiate higher real wages. Strike announcements are more credible, the pool of remaining outsiders is characterised by unfavourable personal characteristics. And thus insiders can rely more heavily on their market power.\(^\text{14}\)

The second common foundation for the positive slope of the WS curve comes from efficiency wage theory (approach d in B3.1.). It is also straightforward to see why employers have to increasingly rely on efficiency wages in tight labour markets. If unemployment is high, it alone acts as a disciplining or motivating device for workers who fear to loose their jobs.\(^\text{15}\) Employers do not have to use additional instruments. If the unemployment rate is low and the outflow rate from unemployment is high, joblessness can be perceived to be a minor threat by workers. They might consequently feel induced to engage in shirking behaviour, low commitment to employers, low investments in firm-specific human capital etc. If a particular job pays above the market level, however, the commitment and the performance of workers can be maintained through efficiency wages, because workers fear to loose their privileged position at their current employer. The WS curve then represents the level of real wages that firms are willing to pay in order to achieve their motivation or screening objectives for any given employment rate.

All in all, there exist plausible economic explanations for a WS curve as depicted in fig. B1. In this section we have presented only intuitive arguments rather than presenting the details of one particular model for a WS curve. In the next chapter, when we discuss regional labour markets and introduce the wage curve, we will

\(^{14}\) For a more complete discussion of this approach, and for a formal derivation of the WS curve under union wage setting see Layard/Nickell/Jackman (1991:83ff.) or Carlin/Soskice (1990:387 ff.).

\(^{15}\) Blanchard/Katz (1997:53 f.) rightly notice, it is really the outflow rate from unemployment that determines the strength the perceived penalty, not so much the overall unemployment rate. However, the overall unemployment rate is commonly used as a proxy for the labour market prospects of the unemployed.
use one specific theoretical approach to explicitly derive a wage curve. This will be one particular efficiency wage model, namely the *shirking approach* by Shapiro/Stiglitz (1984). It is worth mentioning that the same shirking approach could also be used to justify a macroeconomic WS curve. And one should also keep in mind that a WS curve (and a wage curve) can be based on more than one plausible economic argument.

B3.3) The aggregate price setting curve
The WS curve represents equilibrium combinations of (w/p) and U stemming from the labour market. The second central relation in the ELMM is the aggregate representation of product market equilibrium, which in our terminology will be called the *price setting curve* (PS). It is a distinguishing feature of the ELMM to assume imperfect competition in the labour market and derive an upward sloping WS curve. For the product market, however, there exist approaches both with perfect and with imperfect competition.

a) Perfect competition in the product market
With perfect competition in the product market, there is again price taking behaviour of atomistic firms, as well as the requirement that firms make zero profits. It is well known that under these circumstances competitive firms pay labour according to its marginal product.

Any firms maximizes profits \( \pi = p \cdot Y - w \cdot N - r \cdot K \), where \( Y \) is output, \( N \) is labour and \( K \) is capital. It has to take factor prices \( r \) and \( w \) as given. Suppose that capital is internationally traded and thus capital costs \( r \) are determined on world markets. The nominal wage rate \( w \) on the other hand is dictated by the imperfectly competitive labour market. Output prices \( p \) can not be changed by any single firm, but they ground out from the Walrasian tâtonnement process.

In equilibrium, there is only one price level that realigns real wages such that there is equilibrium in the product market. This real wage must be consistent with efficient production along the minimum cost schedules, profit maximization and zero profits for all firms. Under perfect competition this real wage is given by

\[
\frac{w}{p} = \frac{\partial Y}{\partial N} = MPL
\]

Equation (B.1) is the equilibrium condition for the product market and describes the *PS curve*, which for the case of perfect competition can also simply be labelled
the labour demand curve. Its shape is determined entirely by the marginal product of labour (MPL), i.e. by the properties of the underlying production function. If the capital stock is fixed and the production function exhibits neoclassical features, one would typically expect that labour faces diminishing marginal returns. If this is so, the PS curve under perfect competition becomes a downward sloping curve in the real wage/employment rate-space. With a low employment rate, the marginal product of labour is high. Product market equilibrium is obtained at a high real wage level. As the economy comes closer to full employment, labour faces diminishing returns, and product market equilibrium is only compatible with a lower real wage.

However, it is subject to considerable dispute if marginal costs are indeed rising on an aggregate level (e.g. Blanchard/Fisher, 1989:463 ff.). For several reasons it is also conceivable that the MPL is actually constant. This will particularly be the case if the production function is such that capital and labour are used in fixed proportions. If this is the case, as in Blanchard/Katz (1997:55), the real wage consistent with product market equilibrium is independent of the employment rate. The PS curve is then simply a horizontal line in figure B1. The curve would also be horizontal if labour is seen as the only variable output, and constant marginal returns are assumed.

But no matter if the PS curve is flat or downward sloping, the general equilibrium in this economy would lie at the intersection point of the PS curve with the upward sloping WS curve (at point B in figure B2). At this point, a combination of the real wage and the employment rate prevails that is consistent with equilibrium in both the product and the labour market. We will come back to the case with perfect competition when discussing the wage curve approach of Blachflower/Oswald (1996). But now we turn to the (probably) empirically more relevant and the theoretically more interesting case with imperfect product market competition.

16 Under perfect competition the terminology PS curve it not fully appropriate, since the firm sector has no market power to set prices. However, the labelling is chosen in order to highlight the analogy with the imperfect competition case.

17 The issue of market clearing is this more is a bit more complex and requires some discussion. Since there is excess supply in the labour market, by Walras’ law there must be excess demand on some other market. As Lindbeck (1992:192) has put it, there is indeed ‘notional’ excess demand in the product market, since workers are income constrained because of unemployment. They can not demand as much goods as they would desire based on full employment considerations, i.e. at the Walrasian equilibrium represented by point C in fig. B2. The ‘dual decision’ hypothesis of Clower (1965) thus pops up again. However, perfect competition in the product market here simply means that there is no additional quantity rationing stemming from market power in the goods market.
b) **Imperfect competition in the product market**

With imperfect competition, firms have market power and can actively set prices. Suppose that there is a large number of firms in the economy, each producing a distinct but symmetrical good under monopolistic competition. Each firm now maximizes profits subject to a downward sloping demand curve for its specific product. For simplicity we assume that labour is the only variable input of the firm. Profits are given by

\[ \pi = p(Y(N))Y(N) - wN \] (B.2)

where \( Y(N) \) describes the short run production function for any firm, depending only on employment \( N \). The downward sloping demand function is given by \( p(Y(N)) \). Maximizing (B.2) with respect to \( N \) yields the familiar rule that prices are a mark-up over marginal costs, which by definition is equal to the nominal wage divided by the MPL

\[ p = \frac{1}{1-1/\sigma} \frac{w}{MPL} \] (B.3)

\( \sigma \) is simply the absolute value of the price elasticity of demand. To arrive at an expression for the PS curve under imperfect competition, we simply have to rearrange (B.3) in order to obtain

\[ \frac{w}{p} = (1-1/\sigma)MPL \] (B.4)

The slope of the PS curve in the real wage/employment rate-space depends on two factors: a) whether the MPL is constant or declining in the aggregate employment rate, and b) whether the demand elasticity is a function of aggregate employment or not. As far as the MPL is concerned, the same consideration apply as for the case with perfect competition. Labour might face diminishing returns if the capital stock becomes a binding factor, but the evidence for this proposition in aggregate data seems scant. With respect to the demand elasticity \( \sigma \), matters are also controversial. It has become common to work with iso-elastic demand functions e.g. of a CES-type. With these functions, the slope of the PS curve is entirely determined by the slope of the MPL. But some authors have argued that demand elasticity tends to be a pro-cyclical variable (Bils 1987, 1989). If this is so, the mark-up \((1-1/\sigma)^{-1}\) is decreasing in the employment rate and not constant.
Neither theory nor empirical evidence are fully decisive on the slope of the PS curve under imperfect competition, since neither the slope of the marginal cost curve nor the behaviour of the mark-up is known on an aggregate level. For most insights of the ELMM it is, however, not essential whether the PS curve is horizontal or downward sloping.

**B3.4) Equilibrium in the ELMM**

For simplicity and expositional purposes we will work with a flat PS curve. This slope can follow either because both the MPL and the demand elasticity σ are constant, or if the MPL is declining whereas σ is rising in (1-U). We will consider here the former possibility. Figure B2 graphically summarizes the ELMM, and highlights the differences between a competitive Walrasian equilibrium (point C), and the equilibrium points with imperfect competition in both product and labour markets (A), or only in the labour market (B).

The Walrasian equilibrium with perfect competition in goods and labour markets is simply at the intersection of labour supply (LS) and labour demand (MPL). Of course, full employment is rendered under this competitive general equilibrium. But we have argued above that this Pareto-optimal situation might not develop, because of systematic market imperfections.

**Figure B2: Equilibrium in the ELMM**

Product market equilibrium with perfect competition in the goods market is graphically described by the labour demand curve (MPL). For the imperfect competition case, the pricing behaviour of firms must be taken into account that is described by equation (B.4). The PS curve, the “labour demand determined real wage”, will lie below the MPL. Equilibrium in the imperfectly competitive labour market is described by the upward sloping WS curve. For a general equilibrium, both the labour and the product market need to be in equilibrium.
With imperfections in both markets this is the case at point A. The real wage \((w/p)^*\) must be equal to the level that is determined by the PS curve. The associated unemployment rate \(U^*\) can be called the "equilibrium rate of unemployment" (Carlin/Soskice, 1990), or, in reminiscence to Friedman's expression, even the "natural rate of unemployment" (Blanchard/Katz, 1997).\(^{18}\)

It is important to understand in what respect \(U^*\) is an equilibrium rate of unemployment. Probably the best way to think about this issue has been proposed by Layard/Nickell/Jackman (1991:8 ff.). They put strongest emphasis on the fact that equilibrium in the imperfect competition approach does not imply market clearing. It only means that private economic plans are compatible with each other, and that a system will return to the equilibrium configuration in case of a random disturbance.

The WS and the PS curves represent such private plans. The WS curve represents how nominal wages are set in relation to goods prices in the labour market. The PS curve shows how firms on goods markets set output prices in relation to costs (i.e. nominal wages). Both price- and wage setters make their economic plans in real terms, and accordingly use their nominal action parameters. Suppose the WS curve is based on a monopoly union model. The union's bargaining power, and hence the real wage claim, positively depends on the employment rate. The union claims a nominal wage so as to achieve the desired level of the real wage, taking into account (or forming expectations about) the prevailing price level. Equally, price setters are only concerned with the real value of their profits. They set prices in relation to the nominal wage in pursue of their economic plans in real terms. Only at the equilibrium level \((w/p)^*\) are the private economic plans compatible with each other. The adjustment mechanism that brings about the consistency of the real claims is the level of unemployment.

The working of the ELMM can be seen best when considering an example. In figure B3 the general equilibrium is given by point A. This point represents the constellation of real wage and (un)employment, where output claims of firms and union are compatible. Suppose the union expects output prices to rise by 2% in the next year. Since the economy is in its long-run steady state in point A, the unions will claim nominal wage increases also equal to 2%. The real wage is thus constant, as well as the real level of profits and the unemployment rate.

Now suppose that a shock pulls the economy out of the equilibrium constellation A and to some point D that is associated with an unemployment rate \(U_D\) below the equilibrium level. The low unemployment positively affects the union's bargaining power. With adaptive expectations, the union expects inflation to remain at 2% in the next period. But at \(U_D\) the union will want to increase nominal wage by

---

\(^{18}\) One can see that the equilibrium rate of unemployment is lower with perfect than with imperfect product market competition. This is of course due to the fact that imperfect competition in the product market imply quantity rationing from firms, and thus lower labour demand.
more than 2%, say by 5%, in order to reach the desired real wage level \((w/p)_D\). However, this real wage level would put the real profits under strain. Recall that firms set prices as a mark-up over wages, and the equilibrium conditions for the product market are only satisfied at the real wage level \((w/p)^*\).

At point D there is an inconsistency of economic plans. Put differently, the desired real wage income and the desired real profit level do exceed real output. Since this is not possible, an adjustment needs to occur that brings the competing claims of firms and unions for real output shares in line again. This will occur at first through a change in output prices. Typically the timing of the pricing decisions is assumed to be the following: first the union make the nominal wage claim based on expectations about future inflation and unemployment. Afterwards, pricing decisions of firms are made on the basis of the nominal wage claim, which then determine the actual rate inflation.\(^{19}\) Since the firms seeks to maintain their real level of profits, they react to the nominal wage claim and increase prices also by 5% instead of 2%. With inflation equal to 5%, the real wage level is consistent again with product market equilibrium. Graphically this situation is given at point E in fig. B3, where the equilibrium real wage \((w/p)^*\) prevails, but on a higher nominal level than before.

**Figure B3: Adjustment in the ELMM**

![Figure B3: Adjustment in the ELMM](image)

At point E, unemployment is still at the level \(U_D\) below the natural rate \(U^*\). If in the next period the union still seeks to increase real wages, it would now have to claim nominal wage increases of more than 5%. Again, firms would respond by

\(^{19}\) In case the WS curve is based on efficiency wages rather than on a union model, both nominal wages and output prices are set simultaneously by firms that just need to bring together goods market considerations and motivation objectives. The timing of pricing decisions is discussed more intensively in Carlin/Soskice (1990:163).
price increases proportional to the nominal wage increase. The ELMM thus predicts that unemployment can be kept below the natural rate only at the costs of accelerating inflation. Similarly, unemployment rates can remain above $U^*$ only in association with permanently falling inflation. Only at the equilibrium $A$ are real wages claimed by unions and firms compatible with each other, and there is no need for a change in the rate of inflation. The unemployment rate $U^*$ can therefore be understood as the non-accelerating inflation rate of unemployment (NAIRU). However, note that inflation at $U^*$ does not necessarily have to be zero. It is just implied that the rate of change of nominal wages and output prices is identical.

But will the system be endogenously driven back to the equilibrium value $U^*$ after an exogenous shock? In other words, are there endogenous forces that bring unemployment back from $U_D$ to $U^*$? So far we have only shown that maintaining unemployment at $U_D$ imposes costs in form of an accelerating wage-price spiral. But unemployment dynamics have not been explicitly spelled out. Many authors such as Carlin/Soskice (1990) or Lindbeck (1992) are in fact not very explicit about this issue. However, a long-run convergence of unemployment to its natural level $U^*$ is the only economically plausible possibility. Otherwise, as pointed out by Layard/Nickell/Jackman (1991:12f.), there would be an everlasting wage-price spiral. The intuition for the transition from $U_D$ to $U^*$ is easier to provide in the context of efficiency wage models at the core of the WS curve. Since here both wage and price setting is actually done by firms, real wages and (un)employment are determined simultaneously. The combination of an unemployment rate $U_D$ and a real wage $(w/p)^*$ as in point $E$ is not sustainable, because unemployment is too low for given real wages to assure the optimal level of worker morale and effort. The firm sector will thus fire workers until the economy is at the equilibrium rate of joblessness, $U^*$.

In the context of a union model, firms do not set nominal wages. They set nominal prices in relation to the unions’ wage claims. But they also choose the level of employment. Because we have assumed a flat MPL curve, firms can not speculate on a higher marginal productivity of workers at a lower employment rate. If the PS curve were to slope downwards, it would again be easy to see how the economy would converge to its long-run steady state at point $A$. But the basic trade-off implied by a downward-sloping labour demand curve, that plays a very prominent role in other collective bargaining models such as McDonald/Solow (1981), is absent with this linear technology. However, there is no reason to believe that firms will keep on reacting to excessive nominal wage increases only with rising output prices. There will also be a reduction in employment to the equilibrium level $(1-U^*)$ in order to avoid the wage-price-spiral.

The equilibrium combination of unemployment $U^*$ and real wages $(w/p)^*$ is thus going to prevail in the long-run. In the short-run, unemployment can differ from its natural level. But changes in inflation and subsequent changes in employment
will gradually restore the general equilibrium, which is characterised by a consistency of claims for real output shares of wage and price setters.

**B3.5.) Some further issues of the ELMM**

Is the ELMM really a Keynesian model? Since we have placed the ELMM in the tradition of Neo-Keynesianism in the brief history of economic thought that was presented in section B2, one might somehow be biased to say that the ELMM is Keynesian. This is correct insofar as it is by nature a non-market clearing approach with involuntary unemployment. However, the model also has a NAIRU and a long-run vertical Phillips-curve as an integral part, which sounds pretty familiar from Friedman’s (1968) classical model. It would thus probably be most appropriate to view the ELMM as a model with both Keynesian and classical features. For the remainder of this chapter, we want to illustrate this claim a little bit further by comparing the ELMM directly with the Friedman-model and other essential approaches introduced in this chapter.

Even though both the ELMM and the Friedman-model imply the existence of a “natural” rate of unemployment at which inflation is remaining constant, there are still some notable differences. Most notably, these are the underlying microfoundations and the general concept about how the labour market works.

This can be seen best by an example that highlights the differences: Both the Friedman-model and the ELMM share the policy implication that a deprivation of union bargaining power would lead to a lower equilibrium rate of unemployment. In the Friedman-model this would directly reduce the NAIRU, which is reflecting structural characteristics of the labour market. In the ELMM, lower union bargaining power ceteris paribus shifts down the WS curve, so that the intersection with the PS curve occurs at a lower level of U. However, the ELMM would still not imply that the deprivation of union power would transform the labour market into a perfectly competitive ideal. As argued in section B3.1., there are additional arguments that are inherent to the labour market itself why wage underbidding and perfect competition among workers will not occur: efficiency wage considerations as well as inevitable turnover costs on which insiders can rely also without unions. The ELMM derives from a fundamentally different view with respect to the working of labour markets. Friedman’s reference model is a Walrasian system with perfect product and labour markets. The equilibrium rate of unemployment is thus seen as a measure of market malfunctioning, or as the degree of deviation

---

20 This does by the way not imply that all observed unemployment in industrialized economies is entirely involuntary. Lindbeck (1992) points out, that the ELMM might just describe the rationing process of in a primary labour market. Individuals who are kept away from ‘good jobs’ still have to decide whether they accept jobs in an unregulated and usually flexible secondary labour market that can more realistically be described in the usual fashion as a market with perfect competition. It is easy to see how institutions and welfare state arrangements play a role for this individual decision problem.
from the desirable ideal of perfect competition. The ELMM on the other hand acknowledges that the labour market is systematically, by its very nature, characterised by imperfect competition and has no endogenous tendency to clear.

Differences between the ELMM and the Friedman-model also concern the endogenous inflation-unemployment dynamics. Recall that the initial inflationary impulse in Friedman (1968) comes from a monetary injection that stimulates aggregate demand. Short run real effects arise because workers on instance misperceive higher nominal wages with higher real wages. The lower unemployment can only be sustained if authorities keep on injecting money into the economy and systematically keep on confusing the individuals’ perceptions.

The mechanisms of the short-run trade-off between unemployment and inflation in the ELMM are fundamentally different. If unemployment is below its natural level (e.g. because of a demand stimulus), inflationary pressure arises because the bargaining positions of wage and price setters have shifted. Through inflation and employment changes, the real claims on output shares are brought back in line again. Monetary accommodation is not at all needed for this inflation mechanism, nor is the explicit introduction of money. Inflation in the ELMM is not a monetary phenomenon, but is rather stemming from the “battle of the mark-ups” (Layard/Nickell, 1986) between wage and price setters.

The similarities between Friedman and the ELMM are more or less exhausted with the common result that unemployment can be kept below the equilibrium rate only at the cost of accelerating inflation, and with the use of adaptive expectations. Both approaches do not work with the extreme construction of rational expectations as ‘new classical macroeconomics’. If this were the case in the ELMM, the inflation-unemployment dynamics we have just spelled out would not develop. The economy would rather always be at the equilibrium constellation U* and (w/p)*. Note, however, that the use of rational expectations does not imply that the emergence of equilibrium unemployment vanishes. The underlying causes, the market imperfections in labour and goods markets, are more general and do not hinge on the type of expectation formation.

It should be mentioned that the ELMM can be used as a framework to analyse the impact of various exogenous shocks or policy interventions on real wages and unemployment. Among the issues that have been analysed in the vein of the ELMM are structural change, a change in labour productivity, employment subsidies, taxation, profit-sharing schemes, monetary and fiscal policy etc.\(^{21}\)

We have argued in this chapter that the main motivation for developing the ELMM has been the insight that labour markets in the real world can not, or at least should not be modelled as a perfectly competitive market. There is “some-

\(^{21}\) For an extensive coverage of policy implications of the ELMM, see: Layard/Nickell/Jackman (1991); Blanchard/Katz (1997); Bean (1994); and various others.
thing special” about the labour market (Solow, 1990). Approaches with imperfect competition seemingly are more appropriate for describing labour markets from a theoretical point of view. Stated in more technical terms, the major insight of the ELMM has been the replacement of a standard competitive labour supply function with an upward sloping WS curve, which has required micro-foundations radically different from the usual neoclassical ones. The economic arguments that are at the root of the WS curve also apply to the wage curve on the basis of regions, to which we turn now.