MACROECONOMIC THEORY AND THE LABOR MARKET

by

Assar Lindbeck

Seminar Papers are preliminary material circulated to stimulate discussion and critical comment.

October 1991

Institute for International Economic Studies
S-106 91 Stockholm
Sweden
MACROECONOMIC THEORY AND THE LABOR MARKET*

by Assar Lindbeck

The purpose of this lecture is to discuss the labor market in the context of macroeconomic theory. The paper is organized around three main issues: first, the sources of unemployment, and the determination of aggregate employment and unemployment (Sections I-III); second, the effects of supply and demand shocks in the product market on aggregate employment and unemployment (Section IV); and third, the persistence of unemployment (Section V). I close with some policy conclusions (Section VI). The lecture, including the formal modelling, is designed to emphasize ideas and respect observed empirical regularities, rather than deal with technicalities.

I. Sources of Unemployment

The trivial but important starting point of the paper is that changes in unemployment in the sense of excess supply of labor, and hence not merely changes in aggregate employment, constitute an important

* Many of the ideas and judgements in this address have been worked out over the years in collaboration with Dennis Snower. This refers, in particular, to the discussion of the insider-outsider theory of unemployment and to various transmission mechanisms of product demand shocks to aggregate employment and unemployment. I am also grateful for comments on an earlier draft by Jonas Agell, Nils Gottfries, Thorvaldur Gylfason, Claes-Henric Siven, Robert Solow and Lars E.O. Svensson. Much of the analysis in this Address is discussed in more detail, and set in a broader context, in my Ohlin Lectures of 1988 (Lindbeck 1992).
aspect of short- and medium term macroeconomic developments. More specifically, 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 for 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. It is also somewhat surprising that economists who strongly emphasize the rationality of private agents in collecting and interpreting information have to rely so heavily on information mistakes ("misunderstandings") to be able to explain observed short-term macroeconomic fluctuations. My inference from all this is that market-clearing approaches to the labor market cannot possibly be appropriate for an analysis of short- and medium term macroeconomic developments.

How, then, should a nonmarket-clearing approach to the labor market, and related excess supply of labor and job rationing, be formulated? More specifically, what are the sources of nonmarket-clearing wages? Four recent attempts to answer these questions may be particularly worth mentioning.

(i) One obvious approach is to refer to social norms that make underbidding of prevailing wages by unemployed workers a socially unacceptable form of behavior. Metaphorically speaking, there may exist an implicit Eleventh Commandment which tells workers: "Thou shalt not steal jobs from thy comrades by underbidding their wages". There may also be an implicit Twelfth Commandment which tells employers: "Thou
shall not encourage, nor accept, job theft by way of underbidding". As scholars, however, we would like to delve deeper, asking how such norms actually emerge, in whose interest they exist, and how they are upheld against the apparent interest of the unemployed.

(ii) Union models are obviously attractive to economists because they regard nonmarket-clearing wages as the outcome of conscious optimizing behavior of economic agents. They also explicitly introduce undeniable institutional features of the real world: labor unions and collective bargaining. But an obvious weakness of union models is that they do not really explain where the market power of unions comes from, and hence why unions are able to push through and sustain nonmarket-clearing wages. More specifically, why is it that unorganized unemployed workers (including workers who have recently left the union) do not try, and succeed, in getting jobs by underbidding existing wages? In other words, union models do not really explain the absence of underbidding. The limitations of union models as a comprehensive explanation of unemployment in the real world are also underscored by the fact that unemployment often exists also in societies, and in historical periods, with very weak unions, or even no unions at all.

There are, however, theories that are explicitly designed to explain the absence of wage underbidding, as in either efficiency wage theories because it is not in the interest of firms to reduce nonmarket-clearing wages, or insider-outsider theories because incumbent workers, so called "insiders", are able to prevent wage underbidding in their own interest.

(iii) It appears distinctly plausible that profit-maximizing firms, as explained by various versions of the efficiency wage theory, may have an interest in keeping wages above the market-clearing level in order to
boost the productivity (or more generally profitability) of their work force. However, an obvious limitation of this set of theories is that workers, and their unions, do not have much of a role in wage formation. The only role of workers is to react, through their choices of jobs and work effort, to the incentives created by the wage setting of firms. This is difficult to reconcile with the apparent sharing of rents between firms and their employees in many production sectors and countries, and with the often observed statistical relation between wages and the competitive situation in the product market.

(iv) While the efficiency wage theory builds on the assumption of a positive relation between wages and the productivity of the work force, and hence postulates a special "production technology" for the use of labor, the culprit of the insider-outsider theory is the existence of labor turnover costs. Because of labor turnover costs, firms are willing to pay rents to incumbent workers, "insiders", in order to avoid the costs associated with firing of insiders and hiring other workers, "outsiders". In other words, labor turnover costs create market power which the insiders may exploit to push up their wages above the potential market-clearing wage and the reservation wage of both insiders and outsiders, without the insiders losing their jobs. Thus, unlike traditional union models, insider-outsider models do explain the source of the market power of workers.

The most obvious type of labor turnover cost is perhaps the traditional costs of hiring and firing labor, including severance pay; Lindbeck and Snower (1988, Chap. 8). A rather different type of labor turnover cost may emerge when frequent firings and layoffs of workers harm insiders' work morale by making them highly uncertain about the future return on present work effort; Lindbeck and Snower, (1988, Chap.
A more sophisticated type of labor turnover cost is related to the ability of insiders to respond to attempted wage underbidding either by refusing to cooperate with the underbidders in the production process, which reduces the productivity of the underbidders, or by threatening to harass them, which raises their reservation wage. Both types of insider responses to attempted underbidding generate labor turnover costs for firms if they would accept wage underbidding, and if the threats were implemented; Lindbeck and Snower (1988, Chap. 5).

Insiders' threats of noncooperation and harassment against underbidding outsiders signal that underbidding is not an acceptable form of behavior, and that insiders are therefore justified in punishing it. This means that the "noncooperation and harassment version" of the insider-outsider theory helps explain the emergence and continuation of the social norms among workers against wage underbidding - the symbolic "Eleventh Commandment".

The insider-outsider theory does not presuppose unions, but it provides a natural role for them. Unions may boost various types of labor turnover costs, both by intensifying noncooperation and harassment activities, and by pursuing political lobbying for job security legislation that raises the firing costs of labor. Unions may also lobby for laws that make wages below those agreed upon in collective bargaining illegal even for unorganized workers.

Labor turnover costs also boost the clout of union actions such as strikes and work-to-rule, as high labor turnover costs make it expensive for firms to fire workers in response to such union actions; Lindbeck and Snower (1988, Chap. 7). Unions may therefore help insiders seize a larger share of the rents that are created by labor turnover costs. Thus, by threatening to shut down, or in other ways harm firms that are
inclined to accept wage underbidding, unions may help create and sustain the symbolic "Twelfth Commandment" concerning wage norms for firms.

Efficiency wage theories and insider-outsider theories are mutually consistent in the sense that they can be combined in one and the same model. Such a marriage has also been implemented formally; see Lindbeck and Snower (1991). However, it turns out that the efficiency wage mechanism and the insider-outsider mechanism weaken rather than strengthen each other when they are combined in one and the same model, where there is also bargaining over the wage rate. The intuitive reason is that the greater the market power of insiders, and hence the higher the wage due to insider wage pressure, the less necessary is it for firms to strive for high wage rates for efficiency wage reasons, and vice versa.

The conclusion from recent attempts to develop micro-based explanations of nonmarket-clearing wages is, it seems to me, that reasonable explanations do now exist. Needless to say, all available microeconomic theories of nonmarket-clearing wages -- including union theories, efficiency wage theories and insider-outsider theories -- require much more confrontation with empirical observations, and probably also subsequent elaborations and modifications. Moreover, these theories do not, by themselves, explain why some workers choose to remain unemployed rather than work in the "secondary" sector of the economy where jobs are not rationed, i.e. in firms where efficiency wage effects are weak, labor turnover costs are low and unionization insignificant, as compared to the situation in the "primary" sector. These theories therefore have to be amended to explain how workers outside the primary sector are distributed between open unemployment and work in the secondary sector. Such explanations are not difficult to come by, either
by way of Harris-Todaro type models of labor market segmentation, as in Bulow and Summers (1986) or, preferably in my opinion, in terms of the heterogeneity of labor with respect to ability, preferences and wealth. This explanation becomes more profound if the heterogeneity of workers is modelled as an endogenous rather than an exogenous variable, implying, for instance, that ability, preferences and wealth depend on previous work experience; Lindbeck and Snower (1990).

II The Determination of Aggregate Employment and Unemployment

It is important to clarify not only how wage setting may be responsible for unemployment, but also how wage setting behavior itself is influenced by the level of unemployment. In the context of union and insider-outsider theories, it is less "dangerous" for employees to push up wages when the aggregate unemployment rate is low than when it is high, as the probability of being reemployed after losing one's job is relatively high in the former case. In the context of various efficiency wage theories, a lower unemployment rate makes it more necessary for firms to offer high wage rates to limit shirking and quits, and/or have access to a high quality selection of job applicants.

This intuitive reasoning suggests a positive relation between aggregate employment, at given labor supply, and wage setting. In accordance with rather generally accepted terminology, I call this relation a wage setting function. However, such a positive relation cannot be derived unambiguously from all versions of union models, insider-outsider models and efficiency wage models. Cases do exist where the wage setting curve, in real wage and aggregate employment space, is either horizontal or vertical, or even downward sloping. As the "normal" case however, I will assume that the wage setting curve is upward
sloping, in real wage and aggregate employment space, although the possibilities of a horizontal wage setting curve in the very short run and a vertical wage setting curve in the very long run will also be discussed.

Instead of rigorously deriving a wage setting function from one specific micro-based version of union models, efficiency wage models or insider-outsider models -- derivations which have already appeared in the literature -- I prefer here to postulate a "typical" wage setting function for this set of theories. More specifically, let the real wage that is generated by the wage setting process be an increasing function of the ratio between aggregate employment, N, and the labor force, \( \bar{N} \). (No distinction is made here between the (real) product wage and the (real) consumption wage.) In most versions of the wage setting theories discussed above, the real wage rate is also an increasing function of the size of unemployment benefits in real terms, B. Wage setting is also assumed to be a positive function of labor productivity, represented here by the productivity parameter \( b \), as suggested by both the insider-outsider theory and other theories according to which workers share rents with firms. Then the wage setting function simply reads:

\[
(1) \quad w = G(N/\bar{N}, b, B) \quad \text{(WS-function)}
\]

or the inverse function

\[
(1a) \quad N/\bar{N} = H(w, b, B).
\]

In recent years, it has become increasingly common in macroeconomics to model the determination of real wages and aggregate employment as the outcome of the interplay of an aggregate wage setting function of this general type with an aggregate labor demand function,
each one expressed in real terms. Assuming imperfect competition in the product market, the latter function is then, of course, not a strict labor demand curve but rather an optimum relation for firms between the real wage and the employment level. To avoid misunderstandings, I call the curve a labor demand relation rather than a labor demand curve.

Assuming that aggregation over firms generates a macroeconomic labor demand relation of the same basic form as a standard microeconomic labor demand relation, it is simply written:

\[ N_d = F(w, b, m), \]  
\[ (-)(+)(-) \]  

where the parameter \( m \) denotes Lerner's measure of "monopoly power" of a representative firm. The function is illustrated by the LD curve in Figure 1a, where the wage setting function (WS) is also depicted. (The signs of the partial derivates are obvious from traditional microeconomic theory.)

Aggregate employment \( (N) \) and the real wage \( (w) \) are then determined by the condition

\[ N_d = N, \]  

implying that \( w \) is solved out from

\[ N_d = H(w, b, B) = F(w, b, m), \]  
\[ i.e., \quad w = g(b, B, m, N). \]  

The solution is depicted by point \( b \) in Figure 1a, at the intersection of the LD and the WS curves, rather than at point \( a \) as in market clearing models. For simplicity of diagrammatic exposition, it is assumed that labor supply (LS) is vertical. This assumption may be defended by the lack of solid empirical knowledge about the relative strength of the income and substitution effects of changes in the real
wage rate on aggregate labor supply; however, the principles of the analysis do not change if the labor supply curve were instead slightly upward or downward sloping.

Figure 1

Nowadays, this general type of organizational framework for macroeconomic analysis of the labor market is shared by adherents of several different theories of wage setting behavior: union theories, as represented by Layard and Nickell (1986) and Calmfors (1990), efficiency wage theories, e.g. Shapiro and Stiglitz (1984) and insider-outsider theories, cf. Lindbeck and Snower (1990), although the specification of the functions of the model differs considerably among authors. Let us simply call this type of model "the LD-WS model" of real wage and aggregate employment determination. Position b is called the quasi-equilibrium rate of unemployment (QERU) -- "equilibrium" because the point is a solution to equations (1)-(3), and "quasi" because there is aggregate excess supply for labor.¹ Such a quasi-equilibrium position is the only point at which the demand for labor is consistent with real wage setting behavior.

Some clarification of the term "real wage setting" may be in order. No single agent actually sets the real wage. Rather, the WS curve may be said to define the setting of nominal wages, given product prices, while the LD curve defines the nominal prices that are set by firms for alternative labor inputs, given nominal wages. In Layard and Nickell's (1986) useful terminology in the context of a rather similar framework, we may say that the WS curve describes the markup of wages over prices and the LD curve the markup of prices over wages, and that the

¹ A more accurate, but more cumbersome name might be the "non-market clearing equilibrium rate of unemployment" (NERU).
Figure 1
THE DETERMINATION OF UNEMPLOYMENT AND OUTPUT
(a) (b)

(c)
intersection of the curves defines the position where these markups are consistent.

A characteristic feature of this organizational framework is the distinction between labor supply, which is provided by individual households, and wage setting, which is performed by either firms, insiders, unions or the "process" of collective bargaining -- depending on which specific theory of wage setting is entertained. This analytical framework has three important advantages as compared to macromodels where the labor market clears, such as New Classical Macroeconomics and Real Business Cycle Models: (1) It allows explicit determination of unemployment, in the sense of excess supply of labor, and not just aggregate employment. (2) The model deals explicitly with wage and price setting behavior, which means that we may retire the "mythical" auctioneer in both the product and the labor market. (3) In this framework, the magnitude of changes in aggregate employment in response to demand and supply shocks is not closely tied to the elasticity of the labor supply function -- by contrast to market-clearing models. This is an advantage, as explanations of large fluctuations in aggregate employment do not have to rely on the dubious assumption of a highly elastic supply of labor by the household sector; what matters instead is the elasticity of the wage setting curve.

It may be noted that this analytical framework is quite different from models that assume, quite unrealistically, monopolistic competition in the labor market; cf. Hart (1982), Benassy (1975) and Blanchard and Kiyotaki (1987). In fact, such models imply that the supply of labor is provided by "syndicates" of workers rather than by individual households. In that type of model, as in traditional competitive models of the labor market, the labor supply curve is indistinguishable from the wage setting
curve -- in the same way as the price setting curve and the product
supply curve of firms are indistinguishable when firms operate under
monopolistic competition in the product market. In models with
monopolistic competition in the labor market, it is rather meaningless to
talk about unemployment, as the suppliers of labor themselves choose the
desired combination of wages and employment. Thus, there is no excess
supply of labor in such models, although the situation is Pareto
inefficient ("underemployment equilibrium") due to the absence of perfect
competition in the labor market and perhaps also the product market.
This illustrates the importance of making a clear distinction between
suboptimal aggregate employment, on one hand, and unemployment in the
sense of excess supply of labor and related job rationing, on the other.

Such models of monopolistic competition in the labor market are
analytically equivalent to union monopoly models in which all workers are
union members, and the union maximizes the expected utility of identical
members, each with the same probability of being unemployed. The union
is then assumed to act in the interest of all workers, and hence to
behave as if all employee households had given the union the right to
supply labor services on their behalf in the labor market.

Even if it provides a useful organizational framework, the LD-WS
model, as presented here, is certainly "unfinished business" as a
foundation for macroeconomic analysis. I therefore discuss three
extensions and modifications of the model which would make it more useful
in the subsequent analysis. First, the effects of product demand shocks
cannot be analyzed without introducing the demand side in the product
market more explicitly, and in this connection, specifying nominal wage
and price dynamics. Second, the model should be modified to make it
consistent with the empirical observation that there does not seem to be
any long-term trendwise relation between the size of the labor force and the level of labor productivity on one hand, and the unemployment rate on the other hand. Third, and most important, is the necessity of introducing mechanisms of unemployment persistence, to be analyzed here as the dependence of the LD or WS curves on the previous development of aggregate employment or unemployment.

III. Extensions and Modifications of the Model

Before specifying the demand side in the product market, it may be noted that a product supply function, of course, is implicit in the labor demand relation, and that it has the same properties. Hence, the aggregate product supply function (the PS curve in Figure 1b) may be written

\[ Q = f(N_d) = \sum_{(w, b, m)}. \quad \text{(PS-function)} \]

The equilibrium aggregate output level \( Q_0 \) is then read off at point \( \ddot{d} \) in Figure 1b on the product supply curve (PS), after the real wage \( w_0 \) has been determined in the labor market subset.

I use a rather traditional product demand function. It is assumed to include four variables. First, there are two "monetary" variables, real money balances \( M/P \) and the real exchange rate \( e \cdot P^*/P \), where \( M \) is the money stock, \( P \) and \( P^* \) the product price of domestically and foreign produced goods, respectively, and \( e \) is the exchange rate. \( M, e, \) and \( P^* \) are assumed to be exogenous variables in the model. Both empirical analysis and experience from various countries suggest that the effects of changes in the price level on aggregate demand for domestic output are much stronger via the real exchange rate (i.e., by way of overvalued or undervalued exchange rates) than via the real value of money balances.
For the purpose of focusing on the role of the labor market in macroeconomics, capital markets and interest rates are not included in this analysis. If that is done the realism of the analysis would certainly have been enhanced by explicitly introducing also debt/equity considerations of firms and credit rationing, and allowing such factors to influence aggregate demand and aggregate supply. In particular, if firms are concerned with their debt positions ("reluctance towards debt") and if rationing occurs in capital markets, changes in the cash flow and in the net wealth position of firms will influence investment and production decisions; Lindbeck (1963, Chaps. III and VII) and Greenwald and Stiglitz (1988). (In the present model, we may vaguely let changes in M also reflect such considerations.)

As there is unemployment, and assuming that employed workers are constrained in their hours of work, the product demand function will also include as arguments real aggregate labor income, \( w \cdot N \), and profits, \( Q - w \cdot N \); see Lindbeck, (1963, pp. 33-34 and 42-44) and Clower (1965). However, as I do not intend to emphasize the role of redistributions of national income between labor and capital, these types of income are introduced in the product demand functions as an aggregate, i.e., \( Q \). A "real" shift parameter, \( \Delta \), is also assumed to enter the aggregate product demand function (PD-function), which then reads

\[
Q_D = k(M/P, e \cdot P^*/P, Q, A),
\]

\[
(+) \quad (+) \quad (+) \quad (+) \quad (+)
\]

(PD-function)

depicted in Figure 1b.

As \( Q \) is a function of \( w, b \) and \( m \) by eq. (4), and \( w \) is a function of \( b, B, m \) and \( \dot{N} \) by eq.(3'), equation (5) may alternatively be written as the semi-reduced form
(5') \( Q_D = K(M/P, e \cdot P^*/P, b, m, \bar{N}, B, A).2 \\
\quad (+) \quad (+) \quad (+)(-) \quad ? \quad (+)(+)

If \( Q \) had instead been replaced by separate arguments for aggregate labor income, \( w \cdot N \), and profits, \( Q - w \cdot N \), the PD curve in Figure 1b would not, in general, have been vertical; this would not, however, change the principles of the subsequent analysis.3

It remains to be decided whether or not the product market should be assumed to clear. Let us, to begin with, assume that the product market does clear, in the sense that there is no (informal) rationing of either buyers or sellers in the product market. In an aggregate context this assumption is presumably much more reasonable than the assumption that the labor market clears.4 The equilibrium requirement for the product market, then, determines \( P \) after \( w, N, \) and hence also \( Q \), have been determined in the labor market. Thus \( P \) is determined by equation

---

2 The reason for the predicted signs of the partials with respect to \( b, B \) and \( m \) in the reduced form (5') is that a rise in \( b \) and a fall in \( B \) or \( m \) raises \( N, Q \) and \( w \) according to equations (1)-(4). The partial with respect to \( \bar{N} \) is not signed in equation (5'). Unemployment benefits are not introduced in the product demand function (5) because the product demand effects of financing the benefits is assumed to cancel the demand effects of the benefits themselves.

3 Perhaps in a very short-term perspective, the PD curve in Figure 1b should realistically be modelled as upward sloping, as aggregate product demand in that perspective may be stimulated by higher real wages (at constant \( Q \)). This is because the marginal propensity to consume is probably higher for wage income than for profits, in particular retained earnings. However, in a somewhat longer time perspective, when investment spending would be expected to go down along with falling profits, the relation between \( Q_D \) and \( w \) may be exactly the opposite.

4 As there is "notional" excess supply of labor, it follows from Walras' Law that there is notional excess demand in some other market. Indeed, at prevailing real wages, households would have liked to spend more, but are unable to do so, as they are income constrained. Thus, there is notional excess demand for products.

A natural extension of the model would be to introduce inventories, and to allow "unintended" inventory accumulation as an element of temporary "disequilibrium" in the product market.
(6) \( Q_D = Q \), i.e.

(6a) \( K(M/P, e \cdot P^* / P, Q, A) = J(w, b, m) \).

The nominal wage rate is then, of course, \( W = w \cdot P \).

It should be noted that "the extended LD-WS model" presented so far implies a traditional dichotomy between the determination of real and nominal variables, and related neutrality of money, even though the labor market does not clear. This follows, of course, from basing the labor demand relation on traditional microeconomic theory of the firm, and from modelling wage setting behavior in real terms, implying that all behavior functions are homogeneous in degree zero in absolute prices. The real variables are in fact determined in a labor market subset of the model, just as in Classical Economics, though the labor market is not clearing in our model, unlike in Classical and Neo Classical models.

While this model assumes perfectly flexible product prices, it will later on also be used as a benchmark for an analysis of the effects of demand shocks when the price level adjusts only gradually.

The concept of quasi-equilibrium (QERU), as defined above, is obviously somewhat similar to, but not identical with "the natural rate" and the NAIRU. The latter concepts are usually defined as requiring not only that the real wage rate is constant, i.e., that the rates of change in the nominal wage rate and price level coincide, but also that both rates of change are constant. By contrast, the QERU only requires these rates of change to be equal, but not necessarily constant. Thus, denoting nominal values by capital letters, and the proportional rate of change by a "hat": while the NAIRU requires that \( \hat{W} = \hat{P} \ and \hat{W} \) and \( \hat{P} \) are constant (if productivity growth is zero), the QERU, as defined so far, only requires that \( \hat{W} = \hat{P} \). We also require, however, that \( \hat{M} = e \cdot \hat{P^*} = \hat{P} \), so that all arguments in the product demand
function, and not just the arguments in the labor market functions, are constant over time.

A more definite modification of (or even break with) the notion of a natural rate and the NAIRU will occur later on, when I introduce various mechanisms of unemployment persistence.

Suppose now that the economy, for some reason, has wound up outside the QERU. Dynamic stability then requires that the real wage falls to the left, and rises to the right, of the QERU. In other words, the nominal wage rate has to increase less (fall more) than the nominal price to the left of $N_0$, and increases more (falls less) to the right of $N_0$. ($\hat{W} < \hat{P}$ when $N/N < N_0/N$, and $\hat{W} > \hat{P}$ when $N/N > N_0/N$.)

This assumed relation between the rate of change in the real wage rate, $w$, and deviations from the QERU is depicted in Fig.1c. The assumption is reasonable, as it makes sense to assume that wage inflation tends to rise along with falling unemployment rates; and that price inflation tends to rise along with higher excess demand for products. In conformity with standard Phillips curve assumptions, we may also assume that wage inflation rises at a higher trend of price inflation (or expectations of higher inflation), and that price inflation rises at a higher trend of wage inflation (or expectations of higher wage inflation). If such Phillips curve type wage and price dynamics are formally added to our extended LD-WS model, it is easy to show that the quasi-equilibrium position $b$ is a stable node, under the reasonable
assumption that the cross-market effects on nominal wages and prices are weaker than the own-market effects; see Lindbeck (1992, Appendix). ⁵

Such nominal wage and price dynamics of the Phillips-curve type build, of course, largely on "empirical generalizations"; the microeconomic foundations for all types of price and wage dynamics are, as we know, quite weak. Macroeconomics cannot, however, "afford" to abstain from exploiting whatever common sense and empirical generalizations that are available -- and generalizations about nominal wage and price dynamics are available. We should not deny what we see simply because we do not understand it -- or because existing microeconomic theories cannot explain it. In spite of much scorn of wage and price Phillips' curves during the last few decades, there is, in my view, strong empirical support for this conventional macroeconomic wisdom of the 1960s concerning wage and price dynamics, as exploited in the stability discussion above; cf. Tobin (1972). Indeed, I will subsequently try to provide some rationale for such traditional aggregate price and wage dynamics.

Even though, as we shall see, the extended LD-WS model has many advantages, two long-term predictions have to be regarded as "artifacts". First, a rise in the size of work force would, according to the model,

⁵ Starting from the quasi-equilibrium point Ŵ in Figure 1a, where Ŵ = Ŵ, it is assumed not only that Ŵ falls by a higher unemployment rate (i.e. by a lower N/N) but also that Ŵ tends to increase when excess demand for products emerges due to the fall in output (i.e., when Qd - Q > 0). This tendency of Ŵ to rise, however, is counteracted by the influence of the fall in Ŵ on Ŵ, and vice versa.
generate a trendwise rise in the unemployment rate.\textsuperscript{6} Second, the model does not rule out that gradually rising labor productivity generates a long-term trendwise change in the unemployment rate; there is no guarantee that rising labor productivity in the long run shifts the WS curve by the same amount as the LD curve, which would be required to avoid a long term trend of the unemployment rate. These are certainly weaknesses of the model, as in the real world we hardly notice any long-term trend of the unemployment rate.

However, the model can be modified so as to remove these dubious long-run properties. For instance, in the context of the insider-outsider theory, with explicitly modelled labor market flows and wage bargaining, the wage setting function becomes vertical, and the long term unemployment rate independent of both labor productivity and the size of the labor force, provided the retention ratio of the unemployment benefit system (B/wN) is constant, and there is free entry of firms; Lindbeck and Snower (1991.). These may be rather realistic assumptions in the long run.\textsuperscript{7}

Indeed, it is, easily seen from Figure 1a, and the underlying equations, that if the WS curve happens to be vertical, the unemployment rate is independent not only of labor productivity, but also of the size of the labor force. A vertical WS curve shifts, for any level of

\textsuperscript{6} According to eq. (1) a shift in $\bar{N}$ that is followed by a proportional shift in $N$ leaves $w$ unchanged. Hence, at the wage rate $w_0$, the WS curve in Figure 1a shifts horizontally in the same proportions as $\bar{N}$. It follows that the new intersection point along the LD curve implies a smaller rise in quasi-equilibrium employment than in the labor force.

\textsuperscript{7} In their model, Layard and Nickell (1986) make the long-run unemployment rate independent of labor productivity by assuming that the elasticity of the long-run demand for labor with respect to the wage rate is constant.
w, in proportion to the labor force, \( \bar{N} \), according to eq. (1). (See also footnote 6.)

IV. Supply and Demand Shocks

It is fairly straightforward to analyze supply shocks using this model. For instance, exogenous reductions in the marginal product of labor or, more realistically, exogenous increases in the price of imported intermediate inputs (such as oil and raw materials), may be represented as (at least temporary) downward shifts of both the labor demand relation and the wage setting curve -- with a reduction in the real wage rate as the result. There will also be a fall in aggregate employment and output if the LD curve shifts more than the WS curve, resulting in a procyclical pattern for the real wage rate. If, as suggested earlier, the long-run wage setting curve is vertical, only the real wage will change in that time perspective.

In the real world, of course, there may be short-term resistance to such a reduction in the real wage. In the extreme, but not necessarily unrealistic case where the real wage rate is unchanged for a while, the economy may wind up at point \( h \) in Figure 1a, with a gap between the

---

8 While exogenous increases in the price of imported intermediate inputs seem to have been an important macroeconomic disturbance on some occasions, exogenous macroeconomic productivity shocks due to sudden economywide shifts in technology, as assumed in the Real Business Cycle Theory, do not seem plausible as an explanation for fluctuations in aggregate output, and even less so in aggregate employment. The law of large numbers, which presumably is also relevant for changes in technologies among millions of firms producing millions of different products, speaks against the theory. Moreover, studies of the process of diffusion of new technology do not indicate that technology is suddenly, in one or a few years, spread among the bulk of firms in a national economy. Equally implausible is the interpretation, in the application of this theory, of short-term fluctuations in "the Solow residual" as a measure of such exogenous technological shocks. As we know, there are other, more plausible, interpretations of the Solow residual.
actual real wage rate ($w_0$) and the new "quasi-equilibrium" real wage rate (at point $g$ if the WS curve is unchanged). This is recognizable as the celebrated, or notorious, "wage gap", which was frequently, and I believe realistically, referred to as an explanation for the rise in unemployment in Western Europe after the negative supply shocks in the mid- and late 1970s.

By contrast, "exogenous", i.e. unexplained, nominal wage explosions (generated, for instance, by political forces), may instead be depicted as (at least temporary) upward shifts in the WS curve, with countercyclical changes in the real wage as a result. Increases in the payroll tax rate or the income and consumption tax rates may be depicted in the same way, if there is resistance to a fall in the real consumption wage. (It is then assumed, of course, that the LD curve is held constant in the experiment.) Higher unemployment benefits would also be predicted to generate an upward shift of the WS curve -- as well as a leftward shift of the LS curve; these shifts would be expected to be permanent, as there is a permanent change in the payoff between work and nonwork.

Several empirical studies have argued that such nominal wage explosions and increases in taxes and unemployment benefits contributed to the rise in unemployment in Western Europe during the 1970s; see Bean, Layard, Nickell (1986).

The extended LD-WS model encounters, however, the same difficulties as Classical and New Classical macro models in explaining real effects of aggregate demand shocks. Real wage rigidities, and related excess supply of labor and job rationing, do not by themselves assure a more important role for demand management policies than in models where the labor market clears.
Keynesians have often made matters too simple for themselves by assuming that aggregate product demand shocks more or less automatically shift the aggregate labor demand curve, even without sticky product prices and related "spillover effects" as in Barro-Grossman type models. This habit is reflected in the tendency among Keynesians to write the aggregate labor demand curve using not only the real wage rate, but also an aggregate product demand parameter as arguments, i.e., in our notation $N_d = F(w, b, A)$ rather than as $N_d = F(w, b, m)$ as in eq. (2); cf. Layard and Nickell (1986) and Solow (1986). The basic weakness of this specification is that microeconomic labor demand curves, in the context of a product market that clears, reflect parameters on the "production side" of the economy but not parameters that define the level of aggregate product demand; thus, it is difficult to see why the parameter $A$ would enter the labor demand function, except by some rather arbitrary assumption that $m$ is a negative monotone function of $A$.

The reason why the level of product demand does not, by itself, influence the position of the labor demand relation is, of course, that the conventional labor demand function of an individual firm simply reads $n_d = h'^{-1}(w/(1-m))$, where $h(n)$ is the production function, and $m$ is Lerner's measure of "monopoly power", where $m = 1/\eta_F = \mu/\eta \cdot s$, and where $\eta_F$ = the elasticity of product demand of the individual firm, $\eta$ = the price elasticity of aggregate product demand, $s$ = the number of firms, and $\mu$ = the conjecture coefficient expressing the expected change in aggregate output by an individual firm when it changes its own output by one unit; Lindbeck and Snower (1987). In the simple case of identical firms, the aggregate labor demand curve may then be written

$$(2') \quad N_d = s \cdot h'^{-1}(w/(1-m)),$$
which is a more explicit formulation of the previously assumed LD curve, i.e., equation (2) above.

These trivial observations have the important, but often neglected, consequence that, in a model with flexible product prices, a product demand shock will shift the aggregate labor demand curve (in \( w, N \) space), only if (at least) one of the variables \( h'(\cdot) \), \( \eta \), \( \mu \) or \( s \) changes. Systematic transmission mechanisms via \( \eta \) and \( \mu \) are very weak "pegs" on which to hang a macroeconomic theory of transmission mechanisms.

Systematic transmission by way of \( h'(n) \) or \( s \) is more plausible. For instance, during the first phase of a business upswing, there are often underutilized capital assets within firms -- unmanned machines and empty assembly lines. Therefore, the individual firm may temporarily be able to expand labor inputs and capital services from earlier installed capital goods simultaneously, without a fall in the marginal product of labor. Here, then, is a conceivable mechanism for overcoming "the tyranny of the downward-sloping labor demand relation" -- by way of a horizontal, or possibly even upward-sloping, segment of the labor demand relation, and hence also of the LD curve in Figure 1a.

However, a temporarily horizontal or even upward-sloping LD curve is not sufficient for explaining why volume effects of product demand shocks will arise, without a fall in the real wage rate. It is also necessary to explain how movements along that curve may be generated by product demand shocks, i.e. why the wage setting curve (in the context of Figure 1a) would tend to shift down (to the right) in response to product demand shocks. It is, no doubt, possible to identify some plausible mechanism for such effects, most easily in open economies, where the prices of imported goods are components of the consumer price index, and therefore a constant real consumption wage does not require nominal wages
to rise in proportion to prices of domestically produced output. However, the generality, quantitative importance, and permanence of such effects may be somewhat doubtful.

Entry of new firms, and hence an increase in $s$, may be another important transmission mechanism of aggregate product demand shocks, in this case through a shift of the aggregate labor demand relation to the right, rather than, as above, movements along the curve. The mechanism through which new firms enter is not modelled in the present exposition; see, however Lindbeck and Snower (1987). Going outside the model we may speculate on whether a temporary drop in the product wage rate would be necessary in order to generate the higher profits that may be required to induce the entry of firms (i.e., to compensate them for various fixed entry costs). The importance of the entry of firms for higher employment is further highlighted by the insider-outsider theory, as new firms have no insiders who may exploit a business upswing to push up their own wages rather than allow hiring (see the persistence analysis below). The latter argument, however, builds on the assumption that existing wage contracts do not automatically cover also employees in newly established firms.

It may be noted that these transmission mechanisms of product demand shocks at flexible product prices hinge on the operation of various supply-side mechanisms, such as the entry of firms and some special short-term properties of the marginal product curve of labor. It is an open question how important such transmission mechanisms via supply-side mechanisms are, as compared to other transmission mechanisms, such as supply side effects due to various peculiarities in the capital market, cf. Greenwald and Stiglitz (1986), and pure demand-side mechanisms that rely on sluggish nominal product prices -- so highly emphasized in New Keynesian Macroeconomics during the 1980s -- according
to which output passively adjusts to product demand at sluggish product prices. The strongest evidence that the latter type of transmission mechanism is also important is that the aggregate price level does not usually "jump" to a new equilibrium position after an aggregate demand shock, but rises gradually, which is exactly what price-Phillips curves are supposed to reflect.

In a time perspective when nominal prices and wages do not react (much) to a positive product demand shock, and firms simply supply whatever is demanded at the (approximately) unchanged product price, there will, in the context of Figure 1, be a rightward shift of the labor demand and the product supply relations. This is the way New Keynesians tend to introduce nonneutralities of money. The effects are qualitatively similar, in a short-term perspective, if the product demand shock is instead generated by shocks of \( e, P^* \) or \( A \). If, as Keynes assumed in The General Theory, and as Keynesian textbooks have argued for many decades, nominal prices react more than nominal wages to product demand shocks, there would also be a downward movement along the labor demand curve, i.e. to a point southeast of the new quasi-equilibrium position.

Most attempts during the 1980s to explain such sluggishness of aggregate prices in response to demand shocks have relied on asserted microeconomic costs and uncertainties in connection with price changes, including "menu costs", asymmetric information and uncertainties regarding the effects of price changes on sales, as well as "good will" considerations à la Okun. Another explanation, which has been suggested and formalized by Olivier Blanchard (1986), involves combining staggered wage setting, à la Stanley Fisher (1977) and John Taylor (1979), with staggered price setting.
I would prefer to argue that the main reason for the observed sluggishness of the aggregate price level should rather be sought in inter-firm production relations in the context of a complex input-output system. Along such lines, Robert Gordon (1991) has suggested that firms systematically underestimate future price changes of intermediate inputs after an increase in aggregate product demand, and that this underestimation delays the ultimate effect of product demand shocks on the aggregate price level.

I basically agree with Gordon about the importance of delays in passing on price changes of intermediary products through the complex input-output system. However, it is not necessary to assume, as Gordon in fact does, that the difficulties in predicting the ultimate effects of demand shocks lead to systematic mistakes, i.e., to a bias, in agents' expectations about the consequences for the prices of intermediate inputs. An alternative mechanism of delays in price changes is production lags in individual firms between intermediate inputs and the output of the firm.

As long as their input prices have not changed much, some firms may adjust their output prices quite modestly in response to a product demand shock, others not at all. One reason may simply be that the marginal cost curve for many firms is rather flat immediately after a recession and that the elasticity of product demand may be approximately constant. However, some prices, for instance of raw materials and intermediate inputs, as well as nominal wages, are likely to go up and subsequently, after time lags in the production process in individual firms, shift the marginal and average cost curves for other firms upwards. As a result, price increases of individual firms will "wander" gradually through the input-output system. Such a production lag model has, in fact, recently
been formalized; see Lindbeck and Snower (1991). By using such models, with production lags in the context of complex input-output systems, to explain sluggish aggregate prices in response to aggregate product demand shocks, we do not have to rely much on asserted costs or uncertainties regarding price changes in individual firms. We can adhere to the traditional assumption that firms choose optimum combinations of price and output as in the ordinary theory of imperfect (monopolistic) competition, although with a short lag before prices are adjusted by individual firms to product demand shocks.

V. Unemployment Persistence

The discussion above suggests that in a short- and medium term perspective, there are a number of potentially important mechanisms by which both supply and demand shocks in the product market may be transmitted to aggregate output and employment. This holds for models with both flexible and sluggish nominal prices and wages. The apparent, and empirically confirmed, persistence of unemployment, in the sense that today’s unemployment depends positively on yesterday’s unemployment (positive autocorrelation) remains to be explained. Persistence may be understood either as slow dynamic adjustment toward an unchanged (quasi-) equilibrium level of unemployment after a disturbance, or as a change in the (quasi-)equilibrium rate itself under the influence of the previous path of unemployment. In empirical analysis, it is difficult, perhaps impossible, to distinguish between these alternative interpretations of persistence.

9 In this specific model, firms set prices for a short period of time before demand shocks are realized, and the production lags then generate a leverage on the short-term rigidities of prices in individual firms that follows from this type of price setting.
There is no lack of attempted explanations of unemployment persistence. Explanations in terms of capital shortage, cf. Sneessens and Drèze (1986), would imply that the labor demand relation tends to be stuck in a position to the left of the prerecession level also after the negative shock has been reversed. In the context of Figure 1a, the labor demand relation, for instance, would be \( L \bar{D}' \) rather than move to the prerecession level \( L \bar{D} \) in connection with a business upswing after a prolonged recession with low capital accumulation. Unemployment persistence as the result of discouraged worker effects, as emphasized by Layard and Nickell (1986), may instead be depicted as an upward shift of the WS curve when business conditions pick up after a recession, as the long-term unemployed are assumed to have largely stopped searching, and hence exert very little downward pressure on wages. This persistence effect may be accentuated by the development of an "unemployment culture" among unemployed workers.

Alternative, or complementary explanations of unemployment persistence build on labor turnover costs. Ceteris paribus, such costs, have ambiguous effects on the average level of employment over the cycle in many models. However, they tend to create employment inertia at the historically existing employment level, and hence generate unemployment persistence in the sense of high serial correlation of both aggregate employment and unemployment; Bentolila and Bertola (1990) and Lindbeck and Snower (1988, Chaps. 9 and 10). In terms of Figure 1a, the labor demand relation would shift only very "reluctantly" to the right in connection with changes in circumstances that would "normally" raise labor demand in models without labor turnover costs.

In addition to such "direct" consequences of labor turnover costs for unemployment persistence, the insider-outsider theory suggests some
"indirect" consequences via wage formation. This is illustrated in Figure 2, where IDC denotes the insider labor demand curve and EDC the entrant labor demand curve. The vertical difference between the curves represents the marginal labor turnover costs. Let us assume that these consist of marginal hiring costs, $H'$, and marginal firing costs, $F'$. Let us further assume that the entrant wage is $W_E$. In the context of the "non-cooperation and harassment" version of the insider-outsider model, we may instead say that the entrant demand curve for underbidders has been pushed down to EDC and that $W_E$, interpreted as the reservation wage, has been pushed up by threats of harassment by underbidders.

However, it is also important to add another horizontal line, RPC in the figure, the so-called "relative profitability constraint". This defines the highest wage rate that insiders can obtain without being replaced by outsiders. It is intuitively obvious, and easy to show formally, that this line constitutes the sum of the entrant wage ($W_E$) and the marginal turnover costs: $W_E + H' + F'$.

Let us also assume that the incumbent work force in the firm under consideration is $\bar{n}$, and, without loss of generality, that workers have been able to push up the wage rate to point $A$, i.e. to the highest possible level of that work force.

**Figure 2**

Suppose, now, that the labor demand curve shifts down (to the left). It is useful to discuss the consequences for real wages and employment in cases of two extreme, alternative responses of wages. One extreme is that the insiders are anxious for all of them to keep their jobs, and that they are therefore willing to accept the wage reduction required to reach that goal. In the context of Figure 2, the real wage rate would then have to fall from $w_1$ to $w_2$, and the wage-employment point
would move from A to B. This outcome is particularly likely if there is no seniority system among workers, so that all employed workers are faced with the same risk of being fired if the wage rate is not reduced.

The other extreme is that insiders insist on keeping the earlier wage, \( w_1 \), and that some workers are then fired (or laid off), so that the wage/employment point instead moves to \( C \). This outcome is particularly likely if there is a well defined seniority system, so that a majority of workers are confident that they will not be fired even if the product wage is unchanged. This mechanism builds, of course, on the assumption that the remaining workers care less about the employment possibilities of the workers who risk being fired than about their own employment opportunities, implying, in fact, that there are different degrees of "insidedness" among incumbent workers.

Thus, while the insider-outsider theory predicts that shocks of labor demand result in fluctuations in the real wage when there is no seniority system, the theory predicts fluctuations in employment instead when there is a well-defined seniority system -- and, indeed, also when the labor demand shock is not foreseen when wages are set. One reason why high-seniority workers are able to prevent wage underbidding by laid-off workers is that remaining insiders may threaten not to cooperate with, or to harass, fired workers if the latter would attempt to keep their jobs by underbidding prevailing wages. The insider-outsider theory, then, helps explain not only the existence of unemployment, but also the existence of layoffs.

What happens if the shock turns out to be temporary, so that after a while the insider labor demand curve returns to the initial position (IDC)? If the insiders care enough about the employment opportunities of previously fired (laid-off) workers -- another example of a "social
norm"? -- they may abstain from exploiting the opportunity to raise the real wage rate, and hence instead "allow" the firm to move back to the initial employment position, i.e., point A. This is perhaps particularly likely if a period of "bad times" has been short, and the workers fired earlier are still unemployed and living in the same community as the employed workers. We may say that fired workers in this case have the status of "near-insiders" in the eyes of senior insiders.

Another possible type of wage behavior is that after the return of the labor demand curve to the initial position, the remaining insiders try to exploit the improved labor demand situation to push up their wages, say, for simplicity, to the highest possible level without being fired, i.e., to the wage $w_3$. This means that the firm will choose a position, point D, with only a modest increase in employment. The reason why the system does not return to the initial position is, then, that the group of optimizers, and hence the maximand, has shrunk. This means that a temporary negative labor demand shock tends in this case to keep the employment level below the initial level even after the shock has been reversed, with "persistent" effects of a temporary shock; for a more complete analysis, see Lindbeck (1992.)\(^\text{10}\)

This type of wage behavior of remaining insiders towards outsiders is perhaps particularly likely in the case of long-term unemployed workers, who may have lost contact with the insiders and perhaps even

\(^{10}\) It is sometimes suggested that the insider-outsider theory unambiguously predicts that the wage rate is a negative function of the size of the incumbent work force in the immediately preceding period. This is not a correct inference, however. First of all, higher aggregate employment may result in lower unemployment, which would make insiders more willing to push up wages, as the expected fallback income would rise -- as also reflected in the upward sloping wage setting curve in Figure la. Second, the relation between the lagged incumbent work force and wages depends on whether or not the distribution of expected labor market shocks has changed.
moved to other communities. In other words, we may refer not only to
different degrees of "insidedness", but also to different degrees of
"outsidedness". Here, then, is an alternative, or perhaps rather a
complement, to the "discouraged worker's effect" for explaining the
difficulties of the long-term unemployed in getting jobs. This also
underscores the fact that the insider-outsider theory is not a "pure"
insider theory, according to which only conditions inside the firm
influence wages; conditions outside the firm are also important. Not
only the discouraged worker approach, but also the insider-outsider
approach means that the situation of outsiders, as manifest in the size
and duration of unemployment, is important for wage setting.

This latter type of persistence may be characterized as asymmetric
in the sense that positive and negative shocks of equal magnitudes have
quantitatively different effects. In order to return to the initial
employment level, the subsequent positive shock has to be larger than the
initial negative shock, i.e., the new demand curve has to wind up to the
right of the initial labor demand curve.

It should be noted, however, that this type of insider-outsider
model does not predict "full hysteresis" after temporary shocks, as the
term "hysteresis" has been used recently in the macroeconomic literature.
The employment level does not stay at the lower level (point C) after the
labor demand curve has shifted back to the initial position; rather it
winds up at a level (in the figure at point D) between the initial
position (point A) and the position after the negative shock (point C).

VI. Policy Conclusions

In the context of a nonmarket-clearing labor market, it is
certainly reasonable to regard unemployment, in particular highly
persistent unemployment, as a major macroeconomic distortion. There is therefore a potential case for policy actions, provided such actions do not create more problems than they solve. Experience in many countries suggests that the latter reservation is not trivial. Time lags and uncertainties regarding the effects of policy intervention (as emphasized in particular by Milton Friedman), are certainly serious obstacles to welfare-improving stabilization policies. So are conceivable adjustments of individual behavior to expected policies (as emphasized by Robert Lucas), not to mention opportunistic policy actions by governments, i.e., actions induced by party politics or attempts by individual politicians to gain short-term personal popularity.

For reasons such as these, I am perfectly willing to subscribe to the prevailing scepticism about "fine tuning" of monetary and fiscal policy actions, if such policies are interpreted as very ambitious attempts to counteract even modest fluctuations in aggregate employment and unemployment. However, these objections to stabilization policy activism are less important when there are what may be called major macroeconomic "level problems", i.e., situations where the economy is either clearly overheated, as for instance in the U.S. in the mid-1960s, Norway in the mid-1980s and Sweden in the late 1980s, or where unemployment is far above the long-run trend, as in several EC countries in the mid-1980s. In cases like these, it does not matter much if the size and timing of policy actions are far from perfect from a cyclical point of view. Thus, even though it may be advisable to avoid "fine tuning" in monetary and fiscal policy, there may be a strong case for "coarse tuning", e.g. policies designed to avoid serious macroeconomic "level problems".
If this "economic policy philosophy" is accepted, it is certainly more important to analyze the consequences of huge macroeconomic shocks, which are often related to unique historical events, and to the persistence of the effects of those shocks, than to study short-term fluctuations in unemployment in connection with "ordinary" business cycles. Policies that avoid huge increases in aggregate unemployment to begin with, or that speed up the adjustment to the initial (or desired) quasi-equilibrium after a huge rise in unemployment, before persistence mechanisms have had much time to operate, may then be regarded as "investment" in low unemployment in the future. This is perhaps the most important lesson from both recent empirical experience and contemporary developments of the theory of employment and unemployment.

However, the problem is more complex than this. Higher unemployment today, when starting from a very low level, certainly tends to reduce inflation, which means that the future international competitiveness of the domestic economy would improve, with tendencies toward reduced unemployment in the future as a result. The fall in inflation also reduces the political pressure to pursue a restrictive economic policy in the future. The previously discussed intertemporal aspects of unemployment in connection with unemployment persistence therefore have to be combined with concern for inflation persistence, and for the future unemployment consequences of such inflation persistence, in particular when the initial level of unemployment is quite low.

These complex intertemporal interactions between unemployment and inflation and the complications of demand management policies raise the issue as to whether institutional reforms can improve macroeconomic performance and hence reduce the burden for demand management policies. Rather than considering various types of "income policies", which have
seldom been successful except in the very short term, and which may severely damage the market system, I will limit the discussion to two aspects that are directly related to the preceding analysis. I refer to the potential importance of institutional reforms that either reduce the market power of insiders, "power-reducing policies" for short, or make outsiders more attractive for firms, "enfranchising policies" for short; see Lindbeck and Snower (1989 and 1990) and Solow (1990).

Among theoretically conceivable power-reducing policies, the most obvious perhaps are attempts to reduce the turnover costs of labor. Examples include the softening of job security legislation in order to reduce the costs of hiring and firing labor, as such reforms would reduce the "gap" between the insider and the entrant demand curves, and hence reduce the market power of insiders. Other examples are laws that punish illegal strikes, as well as strikes and blockades against "third parties", including strikes and blockades against firms that use non-organized labor. Another conceivable reform to reduce insiders' market power is to force incumbent workers, more directly than at the present time (for instance through earmarked taxes on wage incomes), to finance the unemployment benefit system, and possibly also various types of labor market policies that are induced by unemployment-creating activities of insiders. In terms of Figure 1a, the expected result of such reforms is to generate a downward shift of the WS curve.11

It should be emphasized, however, that power-reducing policies are not Pareto improving. They benefit outsiders at the expense of reducing insiders' real wages and job security. For this reason, current insiders have an incentive to resist such policies, for instance by engaging more

11 I disregard the controversial question of whether the government could influence the unemployment rate by encouraging either centralized or decentralized bargaining.
energetically in rent-creating activities. Of course, the social usefulness of power-reducing policies for insiders also has to be judged on the basis of considerations not encompassed by insider-outsider model. For instance, high job security for insiders, by way of legislation, has been a social target in itself. Moreover, the power position of insiders and their unions may have other values, outside the realm of wage formation, for instance in terms of improving the work environment and protecting employees against arbitrary treatment.

Enfranchising policies are probably less controversial. The most obvious examples of such policies, which are basically designed to raise the marginal product of outsiders, are perhaps improved vocational training of unemployed workers, possibly in connection with apprenticeship systems (as in Germany), which also tend to keep down the entrant wage for unskilled youth, and presumably allow more hiring of young outsiders. Improvements in nationwide labor market exchange systems, and reforms that increase the mobility of labor, are also potentially important policy actions to help outsiders get jobs. (Some of these policies have been tried in Sweden, although the results are controversial.) A better functioning housing market is another obvious example.

Yet another instance of enfranchising policies for outsiders is reform that stimulates firms and employees to introduce productivity-related wage contracts, possibly in the form of profit-sharing arrangements, so that wages would fall automatically in response to negative profitability shocks. Such arrangements would be expected both to limit layoffs of workers without much seniority in business downturns, and to reduce the marginal costs of employing outsiders. This means that the insider-outsider theory gives some support for Weitzman's (1987)
suggestion to reduce unemployment by way of profit sharing. It may be noted, however, that such policies (like the power-reducing policies) may make senior insiders themselves worse off and provoke more rent-creating activities from them.

Changes in the general construction of the unemployment benefit system may also reduce the tendencies for unemployed workers to become disenfranchized in the labor market. There seem to be two main types of unemployment benefit systems that satisfy this requirement. One is the US system, with low benefit levels and short duration of benefits. The other is the Swedish-Norwegian system, with rather high benefits of fixed duration, but where the authorities have the right to terminate benefit payments if the worker repeatedly refuses to accept offered jobs or retraining. The benefit system that is most likely to create long spells of unemployment, and hence also to disenfranchise unemployed workers, is the system that prevails in most EC countries, i.e. long and in reality perhaps even indefinite duration of benefits, and without compelling unemployed workers to accept offered jobs or retraining.

Another conceivable measure to improve the employment prospects of outsiders is to remove barriers to the entry of firms, such as excessive regulations, red tape, tax systems and regulations and imperfections in capital markets that discriminate against new and small firms in many countries. Thus, from the point of view of the employment level, it may be important to fight restrictions on competition not only in the labor market but also in the market for products and capital. Prohibiting wage contracts from also covering new firms (i.e., firms which did not exist when the contracts were signed) may also facilitate the entry of firms. Policies that boost the entry of firms would stimulate aggregate employment both directly by increasing the number of firms in the
economy, and indirectly by reducing the monopoly power of individual firms.

A more controversial method for minimizing the discrimination of outsiders is to stimulate work sharing during periods of unemployment. The advantage could be that the number of insiders would then be larger than otherwise after a fall in aggregate demand for labor; as we know, however, there are also a number of important objections to worksharing.\textsuperscript{12}

Some of these institutional reforms may mainly have the effect of reducing the QERU, as defined above, while other reforms may mainly speed up the return to the quasi-equilibrium level of unemployment after a disturbance. Needless to say, institutional reforms such as these are best regarded as complements to rather than as substitutes for demand management policies, in particular demand management policies for the purpose of avoiding serious "level problems", rather than "fine tuning". For instance, if due to poor sales prospects, there are hardly any vacancies in firms, enfranchizing policies are probably not very successful.

I have tried to show that demand management is indeed important, as there are potentially powerful transmission mechanisms of product demand shocks to the labor market -- at both fully flexible and sluggish nominal prices. I have also indicated that sluggishness of nominal prices can be explained without heavy reliance on highly controversial assumptions of costs and uncertainties of price changes in individual firms.

\textsuperscript{12} See, for instance, Calmfors (1988).
References


