Seminar Paper No. 424

REMAINING PUZZLES AND NEGLECTED
ISSUES IN MACROECONOMICS

by

Assar Lindbeck

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES

University of Stockholm

ISSN 0347-8769

Seminar Paper No. 424

REMAINING PUZZLES AND NEGLECTED ISSUES IN MACROECONOMICS

by

Assar Lindbeck

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

December, 1988

Institute for International Economic Studies
S-106 91 Stockholm
Sweden

(SJE88n.doc)
Dec 5, 1988

Conference on macroeconomics,
Helsinki, June 1988

REMAINING PUZZLES AND NEGLECTED ISSUES IN MACROECONOMICS*

by Assar Lindbeck

ABSTRACT

This paper analyses the "chain" of transmission mechanisms of economic policy actions to financial markets, output and employment and, finally, unemployment -- in an attempt to identify "remaining puzzles" and "neglected issues" in macroeconomics. The paper emphasizes the consequences of fiscal and monetary policy actions on real variables, including real interest rates and real exchange rates. An attempt is made to explain why product demand shocks often tend to have faster effects on output and employment than on prices. The paper also discusses why demand-induced increases in output and employment are not systematically connected with a "reduction" in the product wage rate. The paper ends with a discussion of various supply-side considerations in macroeconomics.

* I am grateful for comments on an earlier version of this paper from Guiellermo Calvo, Karl Gustav Löfgren, Stephan Marglin, Dennis Snower, Torsten Persson, Lars Svensson, and Jörgen W. Weibull, as well as from the participants of the seminar at the Institute for International Economic Studies, University of Stockholm.

This paper focuses on one basic aspect of macroeconomics, namely the determination of aggregate capacity utilization, employment and unemployment. To limit the scope of the analysis, the issue of capacity growth will not be explicitly addressed, even though capacity utilization and capacity growth should ideally be seen as two aspects of one and the same process. When, as here, capacity utilization and unemployment are emphasized, macroeconomics cannot be just as an aggregate version of traditional market-clearing general equilibrium theory. It is also necessary to highlight the possibilities of both various non-neutralities of money and elements of non-market clearing, at least in the labour market.

It is unavoidable, with this emphasis of the discussion, that the role of expectations has to be raised from time to time. Already Don Patinkin (1956, p. 180) noted that "once the Pandora box of expectations and interest and price uncertainty is opened upon, anything can happen". Thus it is natural that economists have tried to bring about some discipline on their thinking about the role of expectations. It is also reasonable to accept the basic notion of the theory of Rational Expectations that private agents try to use all available information in the best possible manner, and indeed that they learn. However, I agree with the critics of the actual application of that approach that the theory of rational expectations is often too indiscriminately combined with the assumption of market clearing. But even more fundamentally: Rational Expectations Macroeconomics, in my judgment, usually make implausible assumptions about the information-set of individual agent. In fact, the rational expectations macroeconomic models presented so far tend to assume that the information set of the individual agent is the model

itself, so that rational expectations are simply interpreted as "model-consistent expectations".² Agents are assumed to have the same information-set, and indeed the same view of the world, as the author of the article. This means that economists are asserted to have acquired no more insights about the economic system than has the general public. More fundamentally: economic agents are assumed to change their view of the world every time a new macro model is presented. This does not sound as a self-evident constraint on macro models, even though it is often worthwhile to check the performance of a model also in the special case of model-consistent expectations.

In view of our limited knowledge about expectations, it should be easy to agree that empirical research, for instance by survey data, on the formation and role of expectations is a highly neglected issue in macroeconomics, and in economics in general -- including, indeed, expectations about the behaviour of the government and the central bank. Most likely, methods developed by other disciplines, such as psychology, can fruitfully be more systematically applied to these issues.

As a way of organizing the paper I will follow the "chain" of transmission mechanisms of policy actions to financial markets (Section I); product demand (Section II); output and employment (Section III); and, finally, unemployment (Section IV) -- in an attempt to identify "remaining puzzles" and "neglected issues" in macroeconomics along that chain. This analysis is followed by a discussion of various supply-side considerations in macroeconomics (section V). This "sequential" organization of the paper is made only for expositional convenience.

² Indeed, the first macroeconomic model with "rational expectations", developed by Alan Walters (1971), used the term "consistent expectations" for this approach.

I. Transmissions to financial markets

Let us start with the transmission of policy impulses to financial markets, emphasizing the consequences for real rather than nominal interest rates and exchange rates.

To begin with, it is reasonable to refute the "Ricardian Equivalence Hypothesis", as being not only theoretically implausible but probably also empirically counterfactual (see e.g., Bernheim, 1988).³ In general, the basic weakness of this hypothesis, as it has been developed by Robert Barro, is the same as the earlier mentioned weakness of Rational Expectations Macroeconomics in general, i.e. the unrealistic assumption about the information set. Not only is the individual assumed to base his private economic behavior on an unbiased view of the long-term functioning of the entire macroeconomic system (about which private agents in the real world probably have very little, and often contradictory knowledge), but also on how higher government deficits today will, in the distant future, influence the size and the distribution of taxes between his own descendants (his own "dynasty") and the descendants of the rest of the population.

Thus refuting the Ricardian Equivalence theorem, classical interest rate theory tells us that fiscal policy actions, and related changes in public sector real financial saving, systematically influence real interest rates by way of the consequences for the ex ante balance between real aggregate saving and real aggregate investment. However, it remains

³ The latter statement is based not mainly on various formal econometric tests, which are not very conclusive, but rather on the observation that in many countries with increased public budget deficits (such as the US and Sweden) private saving has not increased correspondingly (or even fallen).

to determine the relevant "geographical domain" of such an analysis. For instance, is it reasonable to argue that the world-wide fall in real interest rates in the seventies was basically caused by the increase in the saving propensity of the world economy as a whole at that time, in connection to the redistribution of income and wealth to OPEC countries, with low capacity for absorption (for a number of years)? Can, analogously, the rise of real interest rates in the world in the early eighties be explained by the drastic fall during that period in the OPEC surplus, in combination with the reduction in both government and private ex ante real financial saving in several countries, in particular Japan and the US? If so, these experiences suggest that the classical theory of real interest rate determination, though with the entire world economy rather than an individual nation as the relevant geographical domain, has a considerable analytical power, at least in a medium- and long-term perspective (e.g. a couple of years, or perhaps a decade).

It is more doubtful why real, and not just nominal, interest rates are influenced by monetary policy actions, such as open market operations, as seems indeed to be the case in the real world. For instance, how do we explain the "stylized fact" that real interests rates, in particular after tax, systematically tend to be negatively related to changes in the rate of monetary growth and the rate of inflation?⁴ Lacking reliable information about inflationary expectations, this empirical generalization is then based on ex post rather than ex ante interest rates, or

⁴ For instance, during the periods of price deflation in the early nineteen-twenties' and early nineteen-thirties', real interest increased drastically, which they also did in connection to the disinflation during the first half of the eighties'. By contrast, real interest rates in several countries fell during the inflation surge immediately after the second world war as well as during the inflationary spells in the late sixties and in the seventies. (See statistics in Ibbotson and Sinquefeld (1982). For an econometric study for the US, see Mishkin (1981).)

alternatively on the convention of using actual rates of inflation as proxies for expected rates.⁵

It may be tempting to hypothesize that these non-homogeneities in financial markets are the consequences of prices of financial assets, and hence also nominal interest rates, being more flexible than product prices, presumably because financial assets are usually traded on auction markets, whereas products are usually not. This argument certainly helps explain the empirical observation that increased monetary growth, to begin with, seems to depress real interest rates by way of a reduction in nominal interest rates at approximately unchanged rate of inflation.

However, it does not help explain another important empirical observation, namely that changes in nominal interest rates tend to lag changes in the rate of inflation both during periods of increasing and falling inflation. One conceivable explanation is that inflationary expectations, too, lag actual inflation, so that the latter is not a good proxy for the former. However, another explanation of falling real interest rates during periods of rising inflation may be that central banks often try to resist increased nominal interest rates on government securities during such periods, and that interest rates on these securities, by way of substitution, have a considerable influence on interest rates on other financial assets as well. However, the same type of reasoning cannot possibly explain why real interests tend to rise during periods of disinflation, as it is not plausible to argue that central banks in these cases systematically tend to resist nominal interest rate reductions --

⁵ Maybe the greatest contribution that economists can make to the measurement of real interest rates proper (i.e. nominal rates deflated by expected inflation) is to convince governments to issue indexed government securities in different maturities, beside non-indexed securities, so that both real interest rates and inflationary expectations may be inferred directly from data from financial markets.

except perhaps when a country has balance of payments problems. So there seem, indeed, to be some remaining puzzles in this field. Of course, a "full" explanation would have to be based on a complete macroeconomic model; moreover, as illustrated by the reasoning above, it would be valuable if government and central bank behavior, at least to some extent, could be treated as endogenous variables.

So far I have implicitly assumed that the effects of monetary growth and inflation on real interest rates are "transitional" -- in a world where real interest rates "basically", i.e. in a long-term perspective, are determined by the interplay between ex ante real saving and real investment according to classical interest rate theory. However, we can hardly rule out the possibility of some long-lasting effects of monetary policy on real interest rates as well, and hence some long-term "non-neutrality of money". One conceivable reason is institutionally determined long-term stickiness in some nominal interest rates. Zero-interest rates on money holdings, central bank pegging of interest rates on government securities during some periods, and government interest rate regulations on bank deposits are obvious examples. However, long-term non-neutralities could also arise as a result of the so called "Mundell-Tobin effect", which includes a positive real balance effect of inflation on saving, resulting in increased capital intensity and lower real interest rates in long-term equilibrium. (See, e.g. Fisher, 1988)

The observation that prices of financial assets, by contrast to product prices, are determined on auction markets, is probably a more profound observation when analyzing real exchange rates than real interest rates. However, the difficulties to explain recent behavior of real exchange rates, and related international capital movements, certainly suggest that some basic rethinking is needed on this issue as well (see,

e.g., the papers by Dornbusch and Genberg in this volume.) Not only is it necessary to take a new look at the relation between the "monetary (or financial) portfolio block" which is today usually regarded as the proximate determinant of nominal exchange rates, and the "real block" that is usually regarded as the proximate determinant of trade and real investment. It is also important to reconsider the issue whether it is best to use a "monetary theory" or an "asset theory" as the proximate determinant of nominal exchange rates. As an illustration to this issue, do shifts among asset holders from, say, German treasury bills to US treasury bills increase not only the price of US treasury bills relative to the price of West German treasury bills, but also the dollar/D-mark exchange rate -- beside the indirect influence via the demand for transactions balances for money? This is not a trivial, or artificial, question, for if the answer is "yes", we would infer that not only non-sterilized but also sterilized capital movements may influence exchange rates (as then the proportion between earning assets of different currency denomination in private portfolios would change at a constant quantity of national "highpowered" money).

Another important puzzle in "financial macroeconomics" is the striking instability in financial markets in recent years. I would hypothesize that this instability can fruitfully be seen in connection both to increased instability and inconsistency of government (and central bank) policies in different countries in a world with highly integrated financial markets, and to an increased tendency of private agents to endogenize the behavior of governments when forming their expectations about various macroeconomic variables. The combined effect of these features is that new information among private agents about macroeconomic aggregates -- concerning, e.g. public deficits, current account deficits,

inflation, capacity utilization, unemployment, etc. -- may have drastic consequences for the expectations of private agents about government policies and hence also about future asset prices. (See, for instance, Lindbeck, 1988.) It would seem that this is exactly what has happened during the eighties.

However, many informed observers probably agree that the recent volatility in financial markets also reflects "financial bubbles", i.e. the consequences of self-supporting expectations, which are not founded on fundamentals (the latter concept covering also non-biased expectations about government and central bank policies). As we know, such bubbles may be consistent with "rational expectations" in the sense of non-biased expectations -- "rational expectations sunspot theories". However, we can hardly rule out the possibility that "bubbles" in the real world sometimes also reflect systematic misunderstandings among private agents of the macroeconomic consequences of, for instance, government budget deficits and current account deficits, as well as about government policy reactions to such events. Such expectations may then be in conflict with both "fundamentals" and "rational expectations". Here is certainly an important illustration of the need for more empirical research on the formation and role of expectations.

II. Transmissions from financial markets to product demand

In general, there has been much research over the years, indeed from the late thirties', on the transmission of impulses from financial markets to aggregate demand in the product market; so this is hardly a "neglected issue" in macroeconomics. However, there are certainly a number of remaining puzzles in this area as well. For instance, how do we reconcile

the earlier observation that real interest rates tend to fall (often, in fact, considerably) during periods of rising inflation with the empirical evidence that real investment tends to be hurt by inflation? (Fisher, 1981.) Is the explanation to be found in inflation-induced distortions of investment decisions in connection to the tax-system, as suggested by Feldstein and Summers (1978), or are other explanations more important, for instance inflation-induced expectations of future restrictive policies?

Moreover, we do not really know how robust the "real" part of the macroeconomy is to the earlier discussed volatility of asset prices, for instance, in connections to deviations of real interest rates, real exchange rates and the real value of other assets from "fundamentals". However, I would hypothesize that changes in wealth in connection to bubbles, and the bursting of bubbles, have rather small effects on spending -- as compared to the effects of changes of wealth that are not regarded as reflecting bubbles. One reason would be the uncertainty about the prices at which private agents will be able to "come out" of an asset market in connection to a bubble.⁶ Another reason is that a large fraction of the total value of financial assets in society is held by institutions like insurance companies or pensions funds. A "bubble capital gain" is, then, quite difficult to base increased spending on by

⁶ More specifically, an asset holder may be quite sure that a high market price of an asset reflects a "bubble", and that the bubble will burst, but he is uncertain as to when it will burst. As the price may continue to increase before the market bursts, he may decide to "hang on" to the asset, hoping that he will get out of the market "in time". In other words, agents operating in financial markets may run into a "chicken-race situation", where many agents try to keep an asset just until the burst actually occurs, each agent hoping to come out of the market ahead of other asset holders. Though this type of behavior implies that asset prices deviate from fundamentals, and that the behavior of agents probably violates the principles of rational expectations (as agents, on the average, may loose on not deciding to sell earlier), this does not necessarily imply that this view of the world is unrealistic.

contrast to "non-bubble" capital gains, which create expectations of improved future insurance conditions or higher pensions.⁷

It is also an open question as to how important short-term instability in real interest rates and real exchange rates is for product demand, and more generally for the real part of the economy. Of course, prolonged periods (such as five to ten years) of high (or low) real interest rates would be expected to have considerable effects on real investment and real consumption. By contrast, we do not really know to what extent short-term fluctuations (with a reversal within, say, one or two years) in real interest rates influence such decisions, or to what extent they rather influence the timing of the financing of spending, such as the timing of the issue of shares or debt instruments by corporations.

Similarly, long periods of overvalued exchange rates (in the sense of lower profitability in the tradable than in the nontradable sector, or lower profitability in the tradable sector in one specific country than in other countries) certainly harm production, employment and investment in the tradable sector of the country in question. However, it is today not clear if short-term instability (e.g. over a year or two) in real exchange rates have much effects on aggregate production and investment. It is also not obvious that they lead to significant shifts, back and forth, in the allocation of resources between the tradable and the nontradable sectors, or between countries, considering the costs of such shifts. It is likely, though, that increased instability of this type keeps down the

⁷ Economists may be able to think of ingenious ways for the individual to "cash" such capital gains, for instance, by purchasing sales-options for the same type of assets as those that are held by the institution in question. But if many agents simultaneously regard the high asset price as a bubble, such transactions may be very costly to pursue, and hence not necessarily worthwhile. More fundamentally: very few households in the real world seem to operate, at all, in futures' markets of this type.

trend-value of the size of the tradable sector, relative to the nontradable sector, due to the risks that are connected with fluctuations in real exchange rates.

There has, of course, been a long-standing controversy as to whether devaluations are expansionary (due to the stimulation of production in the tradable sector) or depressive (due to negative consequences for domestic demand for consumer goods by way of real balance effects, and possibly also redistributions of income and wealth). However, this controversy cannot possibly be relevant for the question whether devaluations should be used to stimulate capacity utilization and employment. The Tinbergen theory of economic policy should have taught us that one instrument, as a rule, is not adequate to influence two variables in a desirable way. If a devaluation that stimulates production in the tradable sector, at the same time, is expected to create too strong negative effects on domestic aggregate demand, actions to stimulate domestic aggregate demand is, of course, called for. Thus, research on this issue should perhaps concentrate more on how various combinations of devaluations and aggregate demand management may influence both production in the tradable sector and domestic aggregate demand in a desired way.

Another interesting issue, which is perhaps worth some research, is whether increased instability of real exchange rates will induce firms to diversify not only their financial portfolios, but also their real capital assets, i.e. the distribution of production plants between countries, for the purpose of limiting the variability over time of profits, output and

employment. If so, here is another driving force for the "internationalization" of the production activities of firms.⁸

Recent events should give much material for research on these various issues of the consequences of financial market instability for aggregate product demand -- and indeed for the real side of national economies in general.

III. Transmissions from product demand to output and employment

Even more fundamental problems are connected with the transmission mechanisms from product demand to output and employment. For instance, how do we explain the "stylized fact" that an aggregate product demand increase, for a considerable period of time, tends to result in higher aggregate output and employment, while most of the price effects are delayed?⁹ And is it possible to break out of a situation of "unemployment persistence" by the help of a government-induced increase in aggregate

⁸ A more provocative hypothesis, which refers to capacity growth rather than capacity utilization, would be that a certain amount of instability of real exchange rate is favorable for productivity growth in the tradable sector. The reason would be that periods of exceptionally low profitability (in connection with overvalued exchange rates) will push some of the least efficient firms out of business, as well as force surviving firms to remove slack, while periods of exceptionally high profits (due to undervalued exchange rates) may stimulate investment, the entry of firms, and innovations. (This argument is, of course, rather similar to the idea in some old business cycle theories, including Schumpeter's, that the existence of business cycles is favorable for long-term productivity growth). Perhaps microeconomic studies of firms could throw some light on this hypothesis.

⁹ The assertion that inflation usually lags changes in output during the business cycle is easily confirmed from charts over the rate of inflation and changes in output in various issues of the OECD Outlook; indeed, inflation usually seems to peak about twelve to eighteen months after the peak of the business cycle, as measured by the rate of capacity utilization. (Ball, Mankiw and Romer (1988) argue that this type of empirical regularity is more characteristic for situations of low, and moderate, rather than high, inflation).

demand for product, or will such policy actions rather generate more inflation? As an analysis of the pricing policy of firms is crucial for an understanding of these issues, it is obvious that further progress in this field requires that the analysis is pursued in the context of firms that are "price setters", such as firms operating under conditions of imperfect competition.

However, a shift from macro models with perfect competition to macro models with imperfect competition does not, by itself, change the standard macroeconomic conclusion that nominal demand shocks (without redistributational implications) leave equilibrium output and employment unaffected -- with prices and wages increasing in proportion to the nominal shock. (Blanchard and Kiyotaki, 1987). However, various types of non-neutralities may, of course, occur during the course of the adjustment to the new equilibrium position. Macroeconomic policy actions during the course of such an adjustment process may then help speed up the movement towards equilibrium. This point is, of course, of particular importance if the macroeconomy is in a situation far below full utilization of labor, and perhaps also capital, and if the tendencies to "automatic" recovery are weak -- which I understand is the meaning of the term "unemployment persistence".

Let us look, heuristically, at the microeconomics of these issues in the context of firms in monopolistic competition. Figure 1 depicts traditional partial equilibrium microeconomic analysis of monopolistic competition. In the context of this figure, we would have to explain why a shift of the firm-specific product demand curve, for instance from D_1 to D_2 , to begin with would induce a move from point A to point B, rather than to the new optimum point C. Thus, we would have to explain why the latter point would be approached only gradually, for instance, along the

price-output trajectory that is denoted by the "arrowed" path in the figure, assuming unchanged factor prices for the time being. However, in the context of traditional (general equilibrium) macroeconomic analysis, where input prices, including wage rates, are endogenous, the marginal cost curve would, of course, shift. For instance, in the context of macromodels with homogeneous behavior functions, the marginal cost curve for the representative firm would shift to MC_2 , so that the final equilibrium point is D, denoting a return to the initial level of equilibrium output, q_1 . Thus, schematically speaking, we would have to explain both why there is a lag in the adjustment of product prices at given factor prices, and why the changes in factor prices, and the adjustment to these, are delayed.

The most popular explanation has certainly been to assume that firms are not short-term profit maximizers, but that prices are set by way of some "rule of thumb" e.g., with fixed mark-ups on long-run marginal or average costs. The firm is then assumed to be willing to sell and produce whatever is demanded at that "predetermined" price. As the product price in the case of monopolistic competition is usually higher than short-run marginal cost, the firm would then have an incentive to react to higher product demand by increased output at the predetermined product price. This hypothesis of price and output behavior, if realistic, would then also help explain why output and employment may increase, in response to higher product demand, without a reduction in the product wage.

However, even though this type of price and output behavior is perhaps fairly consistent with the earlier mentioned "stylized macroeconomic facts", we cannot really say that macroeconomic behavior has been given a solid micro-economic foundation, as there is not much theory involved. However, attempts to build such a foundation have, of

Fig. 1

The microeconomics of pricing and output with monopolistic competition.

course, been made. The most important ones are perhaps to modify the theory of monopolistic competition with the assumption (i) that the firm does not have good enough information about the marginal cost and marginal revenue functions to shift immediately to the new optimum (profit-maximizing) price-output combination (hence to point C or point D in fig. 1), and that the firm therefore chooses just to change output to get better information on these matters (Kalecki, 1943); or (ii) that the firm for good-will reasons does not change the price except in response to a cost increase (Okun, 1981); or (iii) that the firm uses optimum combinations of "price and quantity rationing", with the implication that the price is not raised to the market-clearing level in booms (Carlton, 1986).

There may very well be some truth in each one of these three assertions. However, with respect to hypothesis (i), it is far from clear why a firm would choose to change only output, rather than both price and output, when it lacks information about the marginal cost and marginal revenue functions. And, with respect to hypotheses (ii) and (iii), it is really an open question if quantity rationing, and variations of such rationing, is important enough empirically in market economics to explain the delay of macroeconomic price changes in response to aggregate product demand increases. Moreover, the empirical observation of fixed mark-ups is perhaps less convincing today than, say, twenty years ago, as we have recently noticed considerable variability in mark-ups (profit margins) over time, in particular in the tradable sector, for instance in connection to "wage-cost explosions" (at constant exchange rates) and to fluctuations in exchange rates.

The "menu costs" theory, emphasizing the costs of price changes, is another recent attempt to explain why firms react to product demand

increases just by higher output, rather than by a combination of higher prices and output as suggested by the traditional theory of monopolistic competition. However, it is hard to believe that the "price rigidity" that is explained by the menu cost argument provides a realistic explanation of quantity adjustments for periods long enough to explain business upswings (downswings) that extend for years. In particular, it does not really sound plausible to assert that the transactions costs of changing prices (even with a "liberal" interpretation of the concept of "costs") would be larger than the transactions costs of changing output, and in particular labor inputs, for which transactions costs are often very large. Already this observation casts serious doubt on the empirical importance of the menu cost hypothesis. Moreover, in an inflationary world, prices are in fact often changed one or even several times a year. If aggregate demand suddenly expands for a firm that is a "price-setter", why would the firm then, the next time the price is changed, not simply make the price change larger than it would have otherwise? For, it is hardly likely that a price rise by, say, five per cent, would be more costly than a price rise by, say, two per cent. We may also ask if the "menu costs" argument does not unduly exploit the "envelope properties" of calculus -- with the partial derivative of the profit function with respect to the price being zero at the profit maximum point. For instance, if a firm, to begin with, is outside the profit maximum point, which would be rather normal if the "menu costs" hypothesis makes sense, it is not really possible to exploit these "envelope properties". (The latter criticism has, in fact, been put forward also by some advocates of the menu cost approach, such as O. Blanchard and N. Kiyotaki, 1988).

These points are not meant to deny that various explanations of "price rigidity" tell part of the story why output often change without

price changes and variations in the product wage. However, considering, the difficulties of explaining the transmission mechanisms from product demand to output and employment by reference to "sticky prices", let us inquire whether the observed phenomena could possibly, alternatively or perhaps rather as a complement, be explained in some other way. After all, as we live in a highly inflationary world, with the rate of inflation often fluctuating between, say, two per cent and ten per cent (and sometimes more), the need to explain stickiness of the price level, or the rate of inflation, is perhaps not overwhelming.

It is then useful to conduct the analysis in the context of the relation between the product wage and labor demand, a relation that will be called "the labor demand relation", for short. It is in this context convenient to distinguish between three types of transmission mechanisms: (1) movements along the "traditional" downward sloping labor demand relation; (2) movements along a horizontal, or even upward-sloping labour demand relation (due to inter-relations between inputs); and (3) shifts of a downward-sloping labour demand relation. As we shall see, each of these three cases represent rather different channels of transmission.

III:1 Movements along a downward-sloping labour demand relation

Keynesian text-book models have for decades analyzed fluctuations in employment (and unemployment) as movement along a downward-sloping labor demand relation (in employment and product wage space), with the implication that the product wage would move countercyclically to employment. While in the case of perfect competition this relation is a demand curve for labor, it is, of course, a semi-reduced form in the case

of monopolistic competition,

$$(1) \quad n = \ell \left(\frac{w}{P} \cdot \frac{1}{1-\eta} \right),$$

which is downward-sloping when the second derivation of the production function (with regard to labor) is negative and the elasticity of product demand, η , is constant; (n is labor input, w the wage rate and P the product price). (Lindbeck and Snower 1987.)

Why would then expansionary aggregate demand management in the product market systematically create downward movements along the downward-sloping labor demand relation? Long-term nominal wage contracts, in particular if they are staggered, are, of course, often asserted to contribute to such countercyclical movements of the product wage in response to product demand shocks, as then nominal wages may be less responsive to demand shocks than are nominal product prices. However, it is equally obvious that wage-indexation works against such tendencies. Moreover, in highly open economies, there are even forces that may make wages react stronger than prices to domestic product demand shocks, as domestic product prices in such economies are to a considerable extent tied to world market prices, while the proximate determinants of wages are to be found in the domestic economy. Hence, a positive product demand shock may very well result in a rise in the product wage in the tradable sector in highly open economies, and therefore induce a movement up along the downward-sloping labor demand relation, with a fall in employment as result.

The US economy is often regarded as being a rather closed economy with long-term, staggered, non-indexed wage contracts. Nevertheless there is particularly weak evidence just in that country that fluctuations in employment has been systematically related to countercyclical variations

in the product wage (even in situations when the disturbances seem to have emerged from the demand side in the product market). Somewhat paradoxically, it is in some of the open European countries, with much indexation of wages, that countercyclical movements of the product wage is to be found. The explanation is probably that variations in product wage rates in various European countries have often been caused by "exogenous" wage hikes (for instance in connection to union activities) or by exchange rate fluctuations, rather than by product demand shocks.

My preliminary conclusion is that movements along a down-ward sloping labor demand relation cannot possibly be the whole, or even the main, story of how product demand shocks are transmitted to changes in output and employment. In other words, it is important to break out of "the tyranny" of the downward-sloping demand relation for labour, according to which only negatively correlated movements in product wages and employment can occur in response to products market demand shocks -- except possibly in the case of expectational errors, or systematic shifts in the elasticity of product demand in connection to aggregate demand changes. (The analytical background for this "tyranny" is, of course, that the traditional labour demand relation is derived from the marginal product function for labour, which is independent of product demand, for a constant elasticity of product demand and given inputs of factors other than labor).

III:2 Movements along a horizontal or upward-sloping labour demand relation

One way of coming out of this analytical "tyranny" would, of course, be to assume that the marginal product for labor, in some interval, does not fall, or even rises, when production goes up. In a short-term

perspective, one possible microeconomic rationale for such an outcome may be that there is idle capital within the firm when a positive product demand shock occurs. For then, increased input of labour could be combined with increased, "free", capital services from the existing capital stock within the firm. Labour is simply "called back" to man idle machines or idle assembly lines. For instance, if the "free" capital services are increased in proportion to labor inputs, we would encounter a situation of constant returns to scale, and hence no fall in the marginal product of labor. Indeed, if capital services and labor inputs are sufficiently complementary ("Edgeworth complements") it is easy to show that the marginal product of labor may even rise in connection to higher production, and that the labor demand relation (equ. (1)) will be upward-sloping, so that increased output and employment is consistent with a rise in the product wage rate.¹⁰ (Lindbeck and Snower 1987) Personally, I believe that this mechanism, i.e. variations in the utilization of existing capital goods within firms, and complementarity of capital services and labor inputs, is an important explanation why, in the real world, output often expands cyclically without a simultaneous fall in the product wage. However, we do not really know how often firms do operate in such an interval of a horizontal,¹¹ or even an upward-sloping labor demand relation.

¹⁰ Geometrically speaking, the production volume is in this case determined by the intersection between a falling marginal revenue curve and a downwards-sloping marginal cost curve, and shifts in product demand is translated into increased output and employment without requiring an accompanying fall in the product wage rate.

¹¹ Clearly, in models with monopolistic competition, a horizontal marginal product of labor curve, and related horizontal marginal costs curve, is not sufficient to account for increased output at a constant price in response to a positive product demand shock; it is also necessary to assume that the elasticity of product demand is constant.

III:3 Shifts in the labour demand relation

Another possibility to explain a positive correlation over the business cycle between product wages and employment would, of course, arise if product demand shocks generate shifts in a downward-sloping labor demand relation for individual firms. This may occur if either the elasticity of product demand or the marginal product of labor shifts (exogenously from the point of view of the individual firm) in connection to higher product demand. However it would be rather curious if increases (reductions) in aggregate product demand would systematically result in positive (negative) shifts also in the elasticity of product demand. For such shifts must then depend on either changes in the distribution of demand between different types of agents (public versus private agents, foreign versus domestic agents, different categories of households, etc.), or on changes in to the composition of demand on different products. The conclusion is presumably that the assumption of systematic, positive shifts of the elasticity of product demand in response to aggregate demand shocks is a very weak peg to hang a macroeconomic transmission mechanism on! (see Lindbeck and Snower, 1987.)

There are, however, examples of "demand management policies" that may shift the aggregate labor demand relation. The most obvious example is, of course, changes in government demand for labor.¹² A somewhat more profound example is increased government spending on infrastructure

¹² For instance, since the late sixties public sector employment (excluding public works programs) has increased from 20 per cent of the labor force to about 38 per cent in Sweden in terms of the number of persons, (and 33 per cent in terms of hours). This is one reason why the unemployment figures have been so low in Sweden during the last decade. (Aggressive devaluations, extensive public works programs and public training programs, as well as liberal use of early retirement of workers, are other explanations).

investment. Not only is aggregate labor demand raised by increased demand for labour by the government to build the infrastructure. Also private sector demand for labour would be stimulated if the infrastructure investment raises the marginal product of labour in the private sector, which may often occur if private production is a technical complement to the infrastructure services. It is perhaps tempting to argue that these effects take longer time to emerge than the length of the ordinary business cycle. However, infrastructure investment may very well stimulate private investment in advance of the infrastructure investments being completed, as already expectations that such infrastructure will be available in the future may stimulate the build-up of private capital stocks already today, with increased private demand for labor of the building industry as a result. (Clearly government spending that stimulates investment in human capital may also shift the marginal product curves in the private sector, though in this case the effects would be expected to occur after quite long lags).

However, in a medium-term perspective (a couple of years), an expansion of aggregate product demand would, of course, be expected to generate a rightward shift in the aggregate labor demand curve also by way of the entry of new firms, though a temporary fall in the product wage may then be needed to create the necessary "dynamic incentives" for their entry. This means that, easy entry of firms may be an important supply-side factor for successful demand management policies

We know, of course, that shifts of a vertical labor demand relation may be brought about by demand management policy in the context of "disequilibrium macro models" à la Barro-Grossman (1971), where both wages and prices are fixed at levels above equilibrium, with excess supply in both markets as a result. (More general formulations are found in

Benassy, (1975); and Grandmont and Laroque (1976)). The shifts of a vertical (derived) demand curve for labour then occurs via "spillover effect" from the product market to the demand for labour. However from a micro-economic point of view, this is a rather curious model. Firms are assumed to be price takers even though they cannot sell what they want. They are assumed to be unable to optimize prices in spite of their ability to make quite elaborate calculations in other respects.¹³

As we know, there have been some attempts to remove the analytical deficiencies of the Barro-Grossman model by reformulations in the context of monopolistic competition. (e.g. Benassy, 1976, 1987) However, then we are really back to the earlier discussion of transmission mechanisms in the context of monopolistic competition. But monopolistic competition cannot possibly be a rationale for Barro-Grossman type "spill-over effects", as monopolistic competition implies movements along a downward-sloping labour demand relation (if there is not a systematic increase also in the elasticity of product demand, as explained above), rather than shifts of a vertical (effective) labor demand curve. To provide a microfoundation for Barro-Grossman type macro models, monopolistic competition would have to be combined with some rationale for predetermined ("sticky") prices, such as the "menu costs" assumption (as in Blanchard and Kiyotaki (1987)).

Transmission mechanisms from product demand to labour demand in search models also encounter difficulties. An example is models with assumed positive externalities of search, where the intensity of job search by workers depends positively on the level of search for labor by

¹³ The reasoning by Okun and Carlton, referred to above, according to which firms engage in combinations of price and quantity rationing, cannot be used as rationale for the Barro-Grossman mechanism, as the former deals with rationing of buyers in booms, whereas the latter deals with rationing of sellers in recessions.

firms, and vice versa, with the possibility of multiple equilibria -- some with low and others with high search activity (e.g. Drazen, 1985). However, it is not self-evident even that the externality of search in such models has the required sign. There is certainly a positive substitution effect on search on one side of the market on the search on the other side, in the sense that the return on search on one side is higher the more intensive is the level of search on the other side. But there is also a negative income effect that works in the opposite direction. For the more intensively agents search on one side of the market, the less do agents on the other side have to search to reach a given probability of getting a job.

IV. Unemployment

It remains to find out the consequences for unemployment. It was, in my judgment, an appropriate approach of Keynesian economists, from the very beginning, to treat unemployment as a disequilibrium phenomenon in the labour market, in the sense that the wage rate is above the market-clearing level. On that matter the Keynesians did, in fact, agree with pre-keynesian macroeconomists, who also asserted that unemployment (beside frictional unemployment) is best characterized as a situation of non-market-clearing wages.

However, a basic weakness of both pre-Keynesian and Keynesian explanations of labour market disequilibria is that they do not tell a reasonable story of (a) where the non-market clearing wage comes from; (b) why laid-off workers are not willing or able to keep their jobs by offering to work for less; and (c) why unemployed workers are not willing

or able to get jobs by underbidding the wages of those who are already employed.

Both the union-monopoly model (Oswald 1982) and models with rivalry between unions (Oswald 1979; Gylfason and Lindbeck 1984) take us some way in explaining the existence of non-market clearing wages. However, they do not really address the basic question of the absence of underbidding by laid-off and unemployed workers, in particular non-unionized workers. By contrast, both the efficiency-wage theory (e.g. Shapiro and Stiglitz, 1984) and the insider-outsider theory (Lindbeck and Snower 1989) try to do just that, though in quite different ways: while the efficiency wage theory derives a non-market clearing wage from the incentives of profit-maximizing firms to set wages so as to minimize wage costs in efficiency units, the insider-outsider theory derives a non-market clearing wage from the market powers which incumbent workers, "insiders", obtain from the turnover costs for labor that firms encounter if they want to replace incumbent workers by others, so called "outsiders".

However, neither theory explains why unemployed workers do not get jobs in the "informal" sector of the economy. (In the context of the efficiency wage theory this sector may be defined as the set of firms for which the efficiency wage effect is weak, while in the context of the insider-outsider theory it may be defined as the set of firms for which the labor turnover costs are low). In other words, neither one of these theories tries to explain why, for a number of unemployed workers, the reservation wage for work in the informal sector is higher than the actual wage rate in that sector. Perhaps one explanation is simply that many individuals have strong preferences for types of job for which they have got specialized education and training, and that the informal sector does not offer such job opportunities for them.

Another puzzle is why new firms do not absorb all unemployed labour.

To explain this, it is certainly necessary to know more about various obstacles to the entry and expansion of new firms. For instance, are the most important limitations for the entry of firms to be found in "natural" obstacles like returns to scale (in connection to various fixed costs) and the scarcity of entrepreneurs (such as people with specific abilities and/or low risk aversion)? Or are the most important explanations rather connected with institutional obstacles, such as (a) unions not allowing new firms to pay lower wages than the wages in existing firms; (b) government regulations; (c) imperfections in credit and capital markets; (d) limitations in the availability of equity capital; or (e) privately created entry restrictions in product markets, for instance by way of producer's cartels? More empirical research on these various matters is certainly highly relevant for an understanding of why unemployed workers are not absorbed by new firms entering the market, including firms that are started by unemployed workers themselves.

Inter-country differences in the ease of the entry of firms could possibly be an important explanation for the more favorable development of employment and unemployment in the US than in Europe during the last decade, assuming that entrepreneurship is more vigorous, and less restrained, in the US than in most European countries. There is indeed some casual evidence to that effect (OECD, 1987, 1988).

V. The supply side

It is not much of an exaggeration to say that Keynesian macroeconomics turned Say's Law on its head by, in fact, implying that

"demand creates its own supply". Of course, the supply side has never been completely neglected in macroeconomics. In those Keynesian models where the demand curve for labour is derived from the marginal product of labor, the conventional production function is an important part of the analysis. Moreover, Keynesian macroeconometric models have gradually, in particular from the early sixties, included both the capacity effects of investment and an exogenous productivity trend, so that aggregate demand has been compared with a growing rather than constant level of full capacity output. And "The New Classical Macroeconomists" have emphasized the consequences for fluctuations in output and employment of both intertemporal substitution of factor supply and exogenous productivity shocks, i.e. an example of so called "real business cycles" (which may be an important factor for macroeconomic fluctuations at a number of occasions). However, macroeconomic analysis has, I believe, much to gain by introducing other more, and perhaps more important supply-side aspects.

Some supply-side factors have already been discussed in previous sections, such as the ease of entry of firms, public infrastructure investment, and the reintroduction in the production process of unused physical capital within the firm. There are, of course, also other potentially important supply-side considerations. Two of these may be worth particular attention, namely job security legislation and the system of unemployment benefits.

The consequences for employment and unemployment of job security legislation would be expected to differ drastically depending on the type of business conditions that exist. In the case of ordinary business cycles, such legislation by raising the turnover costs of labor induces firms to hoard labour during recessions, in particular if these are expected to be short-lived. This tends to stabilize employment, and it

may be an important explanation of why aggregate employment was much more stable in Europe than in the US during the period of "ordinary" business cycles in the fifties and sixties. However, after long recessions, high turnover costs would be expected to make firms quite reluctant to hire more workers in the next business upswing, in particular if firms have little confidence in the sustainability of that upswing. As a consequence, a country with high turnover costs may encounter severe difficulties of expanding employment again after prolonged recessions with "feeble" business upswings. This may be another reason for the higher persistence of unemployment in Europe than in the US in the eighties.

Thus the different levels of turnover costs of labor in the US and Europe may help explain both the greater stability of employment in Europe during the fifties and sixties and the greater persistence of unemployment in the former "continent" during the eighties.

The size and ease of getting unemployment benefits has also been asserted to be an important explanation of unemployment persistence. It is, in my judgement, reasonable to hypothesize that the level of registered unemployment is kept down if either the unemployment benefits are quite low (as in the US), or if people (as in Sweden) are cut off from unemployment benefits (regardless whether they are high or low) if workers turn down offers of jobs or retraining. Thus, to understand the consequences for the level of unemployment of different unemployment benefits systems, it is important to study carefully the combined effects of the level and administration of the system. Both the European continent and the UK, have probably "chosen" the pessimim system of high and liberally provided unemployment benefits (hence without the requirement of the worker to accept offered jobs or offers of retraining programs).

Moreover, the enormous differences in the unemployment rates between different regions within national economies suggest that aggregate unemployment within a country cannot be well understood without a realistic analysis of the reasons why unemployed people do not move to areas with better unemployment prospects. To what extent are the explanations for the limited labor mobility to be found in rigid relative wages, the system of unemployment benefits, rent control, council housing, high marginal tax rates, etc.? And to what extent can the mobility-reducing effects of such circumstances be mitigated by "active labor mobility policies", such as labor training programs, nation-wide market exchange systems, and the subsidization of geographical labor mobility? Our knowledge on these issues is still quite fragmentary, and partly contradictory, also in countries where such policies have been practiced and studied for a long time (such as Sweden).

The possibilities persistence of high unemployment, and of long-lasting real effects of demand and supply shocks, including demand management policies, suggests that the Natural Rate Hypothesis, (or the NAIRU), may have been interpreted much too rigidly by economists. Moreover, as recently pointed out by various authors, the natural rate may not be invariant to the previous path of the economy -- due either to the decumulation of physical or human capital during prolonged recessions¹⁴ or to a tendency among incumbent workers ("insiders") to exploit a business upturn in their own interest, by pushing up their wages so much that new hiring is discouraged.¹⁵ In other words, whereas the traditional natural rate theory gives little scope for demand management policy, except in the very short run, because unemployment tends towards a unique long-term

¹⁴ See Sneesens and Drèze 1986 and Malinvaud 1982

¹⁵ See Blanchard and Summers 1986; Lindbeck and Snower, 1987; Gottfries and Horn 1988.

equilibrium rate, the above considerations certainly widen the scope for such actions because current policies may influence the future long-run equilibrium position in the labor market. (Of course, in the context of a sufficiently long-run perspective, there may still be a unique stationary rate of unemployment.)

A realistic analysis of fluctuations and persistence, respectively, of aggregate capacity utilization and unemployment should, of course, be consistent with the observation that, in a long-term perspective, labour markets have usually been able to absorb a growing labour force -- indeed at gradually rising product wages in economies with technological progress. But then we would enter the domain of macroeconomic growth theory. Indeed, a worthwhile task for future research in macroeconomics is certainly to bring about a better integration of the theory of cyclical fluctuations not only with the analysis of unemployment persistence, but also with theories of capacity growth. Macroeconomic growth models (à la Solow) have certainly helped clarify that an economy may be able to absorb a growing labour force, regardless of the rates of growth of both the labor force and total factor productivity. However, I would argue that macroeconomic (i.e. highly aggregate) approaches are basically more conducive to the analysis of capacity utilization than to the analysis of capacity growth. The latter, I largely regard as an issue of whether the general economic, social and political environment, including economic incentives and property rights, are conducive to the efficiency of factor supply, the allocation of resources, work intensity, technological improvements, innovation and entrepreneurship. In my judgment, analyses of such issues require more concern both for institutional arrangements and for the heterogeneity of the microeconomic agents, and indeed concern for

unique agents such as entrepreneurs, than does the analysis of variations in capacity utilization and employment.

References

- Ball, L., Mankiw, G., and Romer, D., "The New Keynesian Economics and the Output-Inflation Trade-Off", Brookings Papers on Economic Activity, 1, 1988, 11-82.
- Barro, R.J. and Grossman, H.I., "A General Disequilibrium Model of Income and Employment", American Economic Review, 61, 1971, 82-93.
- Benassy, J.P., "Neo-Keynesian Disequilibrium Theory in a Monetary Economy", Review of Economic Studies, 1975, 503-23.
- Benassy, J.P., "The disequilibrium approach to monopolistic price setting and general monopolistic equilibrium", Review of Economic Studies, 43, 69-81, 1976.
- Benassy, J.P., "Imperfect Competition, Unemployment and Policy", European Economic Review, Papers and Proceedings, 31, 1987, 417-26.
- Bernheim, B.D. "Ricardian Equivalence: An Evaluation of Theory and Evidence", NBER, Macroeconomic Annual, 1988.
- Blanchard, O. and N. Kiyotaki, "Monopolistic Competition and the Effects of Aggregate Demand", American Economic Review, 77, 647-66, 1987.
- Blanchard, O. and L. Summers, "Hysteresis and the European Unemployment Problem", National Bureau of Economic Research, Macroeconomic Annual, 15-78, 1986.
- Carlton, "The Rigidity of Prices", American Economic Review, 76, 637-58, 1986.
- Dornbusch, R., "Exchange Rates and Macroeconomics: A selective Survey", this volume, 1989.
- Drazen, A., "Involuntary Unemployment and aggregate Demand Spillovers in an optimum Search Model", mimeo, 1985.
- Dixit, A and Stiglitz, J., "Monopolistic Competition and Optimum Product Diversity", American Economic Review, 67, 1977, 297-308.
- Feldstein, M. and Summers, L., "Inflation, Tax Rules, and the Long-Term Interest Rate", Brookings Papers on Economic Activity, 1, 1978, 61-99.
- Fisher, S., "Towards an Understanding of the Costs of Inflation:II", in K. Brunner and A. Meltzer (eds.), "The Costs and Consequences of Inflation", Carnegie-Rochester Conference on Public Policy, vol. 15, 1981, 5-42.
- Fisher, S., "Recent Developments of Macroeconomics", forthcoming Economic Journal.
- Genberg, H., "Exchange Rate Management and Macroeconomic Policy: A National Perspective", this volume, 1989.

- Gottfries, N. and Horn, H. "Wage formation and the Persistence of Unemployment", The economic Journal, 97, 484-497, 1988.
- Grandmont, J.M. and Laroque, G., "On Temporary Keynesian Equilibria", Review of Economic Studies, XLIII, 1976, 53-67.
- Gylfason, T. and A., Lindbeck, "Union Rivalry and wages: An Oligopolistic Approach", Economica, 51, 129-39, 1984.
- Ibbotson R. and Sinquefield A., Stocks, Bonds Bills and Inflation: The Past and the Future, 1982.
- Kalecki, M., "Costs and Prices", in Studies in Economic Dynamics, Allen and Unwing, London, 1943.
- Layard R. and Nickel S., "Unemployment in Britain" Economica, vol. 53, 1986 Supplement, S121-S169.
- Lindbeck, A., "Swedish Industry in a National and International Perspective", Skandinaviska Banken Quarterly Review, no 3., 1988.
-
- Lindbeck, A. and Snower D., "Union Activity and Wage Employment Movements", European Economic Review Proceedings, 31, 157-67, 1987.
- Lindbeck A. and Snower D., "Transmission Mechanism from Product to the Labor Market", Institute for International Economic Studies, Seminar Paper No 403, 1987.
- Lindbeck A. and Snower D., The insider-Outsider Theory of Employment and Unemployment, MIT Press 1989.
- Lindbeck, A. and Weibull, J.W., "Intergenerational Aspects of Public Transfers, Borrowing and debt", Scandinavian Journal of Economics, 88, 1986, 239-267.
- Malinvaud, E. "Wages and Unemployment", Economic Journal, 92, 1982.
- OECD, Labor Market Outlook, 1987 and 1988.
- Okun, A., Prices and Quantities, Basil Blackwell, Oxford, 1981.
- Oswald, A.J., "The Microeconomic Theory of the Trade Union", Economic Journal, 92, 1982, 576-595.
- Oswald, A., "Wage Determination in an Economy With Many Trade Unions", Oxford Economic Papers, 31, 369-85, 1979.
- Patinkin, D. Money, Interest and Prices, Row, Peterson and Company, Evanstone, 1956.
- Shapiro C. and Stiglitz J.E., "Equilibrium Unemployment as a Worker Discipline Device", American Economic Review, 74, 1984, 433-444.
- Sneesens, H. and J. Dréze, "A discussion of Belgian Unemployment, Combining Traditional Concepts and Disequilibrium Econometrics", Economica, (Supplement), 89-119, 1986.

Walters, A., "Consistent Expectations, Distributed Lags and the Quantity Theory", Economic Journal, 81, 1971.