ON THE CONCEPT OF EFFICIENCY
Prabhat Patnaik

Concepts which have to be handled with great care are bandied about irresponsibly today. It would not be inappropriate therefore to dwell on this issue and to illustrate it with reference to the concept of efficiency, which is widely misused.

I

A particularly common example of misuse is the statement: "the private sector is more efficient than the public sector". If the evidence offered in support relates to overall profitability of the two sectors then the baselessness of the claim is obvious. If two sectors producing non-identical commodity vectors and facing arbitrary prices, as opposed to `shadow prices' derived from an optimisation exercise, have different rates of profit, then this fact says absolutely nothing about their respective `efficiencies' no matter how the latter is defined (unless it is defined to be identical with such empirical profitability in which case the proposition is trivial). Even when we compare the two sectors producing apparently the same commodity, say steel, there are important differences in product-mix because of which profitability comparisons are no index of relative efficiency. In fact such a comparison between the public and the private sectors is exceedingly complex. A possible way out is to take an engineering notion of efficiency, examining the use of some key inputs per unit of comparable output, e.g. consumption of power per unit of pig iron production. On such a comparison however the public sector does not come out badly at all (Bagchi et al 1985).

I shall be concerned here, however, not with such sectoral comparisons but with the macro level, where the concept of efficiency commonly used is equally flawed. One often comes across the view that "the Japanese economy is more efficient that the Indian economy". The proof of this statement is supposed to lie in Japan's superior export performance. As an explanation of the relative export performances of the two countries this is a tautology reminiscent of the following anecdote. When Paul Samuelson was asked why doctors earned so much more than sweepers, his reply was that the former's contribution to society was greater than the latter's, and when asked on what basis he could assert this his reply was: "look at their respective earnings"! If relative `efficiencies', assessed in terms of relative export performances, are used as an explanation of the latter, then we have an `explanation' of the same genre. On the other hand relative export performances as an indicator of relative efficiencies have nothing to recommend them. Like relative profitabilities between the public and the private sectors in the example above, they have no rationale as an index of relative efficiencies. And it would force us into absurd conclusions, e.g. that colonial India prior to the First World War was a shining example of `efficiency' (since it witnessed rapid export growth), that it was more `efficient' than inter-war India (when the Depression arrested export growth) and so on.

Fortunately however a direct meaning is given to the concept of efficiency in the economic literature, namely, if with the resource endowments available in an economy in
any given situation (in the sense of being offered by their owners), the output obtainable is not in a vector-wise sense larger than what is produced, then the situation is one of ‘efficient’ production. Putting it differently, if the resource endowments of an economy are given, then a ‘production possibility frontier’ can be drawn with respect to those endowments showing the alternative bundles of maximal amounts of commodities that can be produced with those endowments. Any point on this frontier is a point of efficient production.

It by no means follows that any point on this frontier is ‘better’ or ‘socially preferable’ to a point inside the frontier, i.e. that a movement from a situation of ‘inefficient production’ to a situation of ‘efficient production’ is necessarily a good thing. Welfare economics spent decades looking for arguments that could establish that a point of efficient production, no matter where it is located on the frontier, is preferable to a point of inefficient production. But a comparison between any two points necessarily requires a judgment about income distribution. It follows then that not only can we not compare two points, one inside the frontier and one on the frontier but producing vector-wise non-comparable commodity bundles, without making some judgment about income distribution, but we cannot even compare two points, one involving a larger output (vector-wise) than another, without such a judgment. In other words even a move from one point to another where the latter produces more of all commodities cannot be defended on welfare grounds unless there is a judgment that it has not worsened income distribution. Thus, even according to conventional Welfare Economics, an improvement in the ‘efficiency’ of production is not per se desirable. So, all the talk one hears these days about the need for improving the ‘efficiency’ of production, which says nothing at all about income distribution, lacks any theoretical rationale, even according to conventional economics.

But I shall not be concerned about this aspect either. My concern is with the concept of ‘efficiency’ itself, not with whether a move towards efficiency as conventionally defined is a good thing or a bad thing.

II

The proposition that a vector-wise increase in output from a given endowment of inputs constitutes an improvement in productive efficiency is an unexceptionable proposition. But it is invariably taken to be synonymous with another quite distinct proposition, which can be expressed as follows.

Consider any actual position of production. It is characterised by an activity-set that transforms a bundle of inputs into a bundle of outputs. One can think of marginal rates of transformation/substitution of inputs into outputs, of outputs into outputs and of inputs into inputs associated with this activity-set. Now, if there exists some activity outside of this set whose inclusion at the expense of some activity inside the set "improves" these rates of transformation/substitution, in the sense that one more unit of any input either gives more extra output in the alternative set than currently or saves more of some other
input than currently, or that one more unit of any output necessitates a lower fall in some other output than currently, then the use of this alternative set instead of the current one entails an improvement in efficiency.

This second proposition can be stated in a different way: associated with any activity set is a set of implicit prices (whose ratios are nothing else but these rates of transformation/substitution we have been talking about). If at these prices some other activity-set, formed by replacing a current activity by an activity outside the set, yields a profit, then a move to this set represents an improvement of efficiency.

These two propositions, one relating to production outcome and the other relating to activity choice, which every economics student has been taught over and over again to take to be synonymous are however not so.

The classic case of an activity through which the rate of transformation of inputs into some output gets enhanced is of course trade. The second proposition therefore is reflected in statements such as "Trade is efficient production" which occurs in Samuelson's textbook on economics or in such well-known theorems as "Free trade is better than no trade" or "Free trade is better than restricted trade", "better" being defined in terms of a potential vector-wise improvement in the availability of commodities for a given endowment. All these propositions are based on the belief that the production possibility frontier with trade (i.e. treating trade as a method of "producing") lies outside that for the no-trade situation (except for a point of tangency which denotes actual post-trade production). Consequently through trade the country can move to a situation where the vector-wise availability of goods is larger than in the pre-trade situation. And since efficiency is defined in terms of an increase in the vector-wise availability of goods, trade is efficient production. (The fact that the actual post-trade consumption point may not be vector-wise comparable to the pre-trade consumption point does raise some complications but these, not being germane to the present argument, need not detain us).

All these propositions however are wrong. And the reason is obvious. In saying that with trade we move from a lower to a higher 'production possibility frontier' we are implicitly assuming that trade leaves the magnitude of resource use unchanged, i.e. it does not result in unemployment or idle capacity. If trade did give rise to unemployment or unutilised endowments, then even though the full-endowment-use production possibility frontier with trade would lie outside the corresponding frontier without trade, the actual point of post-trade availability may well be vector-wise inferior to the point of pre-trade availability. The entire corpus of literature proclaiming trade to be efficiency-augmenting implicitly assumes therefore that the economy is always on the post-trade availability frontier (or what we have called 'production frontier with trade'). Not only however is there no reason for this assumption to hold, but indeed precisely the opposite is the likely scenario. These propositions in other words are based on assumptions which, as a rule, do not get fulfilled.

Putting it differently the conclusion about the efficiency-augmenting effect of trade is derived from an implicit assumption of full employment in the post-trade situation. But
this is never explicitly stated since the whole of neo-classical economics is based on a full employment assumption anyway. In other words whether we are talking about the pre-trade or the post-trade scenario, the presumption is that there are spontaneous forces in the economy keeping the economy at full employment; the question then is simply comparing the alternative output bundles that could be obtained from the given, fully-employed endowments, through alternative activity sets, and the set whose output bundle can be improved upon in a vector-wise sense must be inefficient.

The claimed synonymity of the two propositions that we mentioned earlier, one defining efficiency improvement in terms of a higher output bundle from given endowments and the other defining it in terms of a movement to an activity set giving a "better" rate of transformation holds only on the assumption of full employment. Once we drop this assumption as being utterly unreal (we shall discuss this lack of realism in greater detail in the next section) the first concept of efficiency would of course survive but not the second. And with the collapse of the second concept of efficiency, all conventional arguments for trade liberalisation on efficiency grounds also collapse.

It follows too that on the basis of the first definition, which is the only possible definition of production efficiency available in economics, the advanced capitalist countries are extraordinarily inefficient, with Europe having in recent times recorded an unemployment rate of around 12 percent, the US about half of that (though its official rate is a gross understatement), and Japan's unemployment rate creeping up. The difference between the developed countries and the pre-structural adjustment underdeveloped countries like India lay in the fact they had different types of inefficiency, not in the fact that the former were efficient while the latter were not. The fact that the former were better entrenched in world trade than the latter (if we exclude the four East Asian economies) is a fact sui generis. One can use some other term to characterise it, one can talk in terms of more or less 'export-successful' economies, but one should not confuse 'export-successfulness' with 'efficiency'.

Looking at it differently, the only meaningful definition of efficiency available in economics really means the absence of two very dissimilar types of 'waste', one, recognised by neo-classical economics, arising from a wrong activity-set, the other, not recognised by neo-classical economics but perhaps far more important in practice, arising from the forced idleness, or involuntary unemployment, of resources. The lack of recognition of this latter phenomenon by neo-classical economics would not matter if it did not make policy prescriptions based on this lack of recognition. But because it does, a critique becomes urgent.

III

The fact that there is no spontaneous tendency towards full employment in a capitalist economy hardly needs reiteration; and this of course undermines both the theoretical structure as well as the policy prescriptions of neo-classical economics (and hence of the Bretton Woods institutions). But what is not usually recognised is that there are two separate and simultaneously operating reasons (in the sense of output being the lower of
what either of these considerations determines) why full employment of resources is the exception rather than the rule under capitalism. Since either of these would nullify neoclassical conclusions, the existence of both of them makes the case against such conclusions far stronger.

The first of these is the Keynesian-Kaleckian reason. This states, to borrow Janos Kornai's (1979) terminology, that classical capitalism is a demand-constrained system. Kalecki, however, had gone even further: he had argued (1943 reprinted in 1971), with remarkable prescience, that even with State intervention through demand management, capitalism would still be characterised by unemployment because of the political opposition of capitalists to a state of full employment. True, full employment would bring the capitalists larger profits but "discipline in the factories' and 'political stability' are more appreciated by business leaders than profits" (1971, p.141): "lasting full employment is not at all to their liking" since the "workers would 'get out of hand' and the 'captains of industry' would be anxious to 'teach them a lesson'" (p.144). Writing in 1971, even before the protracted downturn in the advanced capitalist economies (which some have identified as a Kondratiev downswing) Kalecki had argued that "even contemporary capitalism where deep depressions are avoided as a result of Government intervention, is in general still fairly remote from … a state of full utilisation of resources. This is best shown by the fact that prices of finished goods are fixed on a cost basis rather than determined by demand" (p.164).

The second reason consists of the fact that, even in a world where there is no demand constraint, output would still be limited by 'disproportionality', of which the commonest form in an underdeveloped capitalist economy would be the constraint imposed on employment by the size of the wage goods sector in a situation where the real wage rate is downward inflexible below a certain level. If the production capacity in the wage goods sector and the methods of production in that sector as well as elsewhere are given, then the downward-inflexible real wage rate makes the maximal employment and hence the maximal overall output in the economy dependent upon the full capacity output of the wage goods sector. The actual output in both the wage goods and the non-wage goods sectors may well be less than this, when there is a demand constraint, but it cannot be higher. Even if the demand constraint is overcome, there would still be unemployment in the economy and unutilised capacity in the non-wage goods sector since the constraint upon employment and capacity utilisation would arise in such a case by the size of the wage goods output.

Disproportionality of course need not arise exclusively from the size of the wage goods sector. Indeed at least two other forms of disproportionality have figured prominently in the development literature, the first relating to foreign exchange shortage, and the second, which formed the basis of the celebrated Mahalanobis model, to the size of the capital goods sector. But the disproportionality arising from the shortage of foreign exchange is a 'derived disproportionality', since foreign exchange represents command over goods and its shortage can impose a constraint on employment and capacity use only via a shortage of some particular good. And if, in general, goods are assumed to be domestically producible (or their domestic use restrainable), then the foreign exchange constraint
simply boils down to one form of capacity constraint or the other, whether of wage goods, or of capital goods, or of intermediate goods (as in the Raj-Sen model 1961). Of these the most potent, especially in an unreformed agrarian structure, is, as Kalecki (1972) emphasised, the wage goods constraint.

If there is this constraint then the standard neo-classical prescriptions such as trade liberalisation could well become counterproductive. If for example wage goods output is exported and non-wage goods (e.g. manufacturing) output is imported in exchange, then domestic activity and employment shrinks as a result (Patnaik 1996). Trade, far from being a means of augmenting efficiency, actually reduces efficiency in the sense of reducing the output vector for given endowments (quite distinct from its adverse income distribution effects in this case which would be additional to this).

Looking at it differently, the existence of any fixity of price makes output determination subject to multipliers (Kaldor 1978). Both the aggregate demand problem as well as the disproportionality problem arise from fixity of prices which disrupts the possibility of a Walrasian equilibrium. The aggregate demand, or the Keynesian, problem arises, as Kaldor had pointed out, because of the inflexibility of the interest rate. The disproportionality problem that we have been talking about arises because of the inflexibility of the real wage rate and this makes overall output equal to some multiplier times the wage goods output. When output determination is subject to multipliers, all neo-classical propositions about efficiency break down and in fact the prescriptions may become counterproductive. And since the theory underlying the programmes advocated by the Bretton Woods institutions is the standard neo-classical theory with its assumption of full employment of all resources, any inflexibility of prices (which incidentally has nothing necessarily to do with controls over the 'free' functioning of markets since both the inflexibilities we have mentioned arise under 'free' markets) and its consequence in terms of a multiplier determination of output, completely undermines the theoretical rationale of these programmes.

In discussing disproportionality above I assumed, realistically in my view, non-shiftable capital across sectors and fixed production methods in the short-period. It may be thought that it is these assumptions rather than the downward inflexibility of the real wage which are responsible for the conclusion that trade liberalisation, if it leads to wage goods export and non-wage goods import, would reduce domestic employment and output. This however is wrong. Even in a neo-classical model of shiftable capital and production functions it remains true that if the wage rate reckoned in terms, say, of corn (but not necessarily exclusively spent on corn) remains downward-inflexible, then trade liberalisation resulting in corn export and non-corn import would cause a reduction in employment and output in the economy, provided the corn sector has a higher capital-labour ratio over the relevant range than the non-corn sector. This condition which appears restrictive at first sight is however required if the initial pre-trade equilibrium is to be locally stable, and hence for a coherent neo-classical story in the first place (this is true whether or not real wages are flexible). It follows then that even if a neo-classical story is told for a world with inflexible real wages, the opening up of trade in such a world, involving the export of wage goods, must, according to this story itself result in a
reduction in output and employment.

It is intriguing that while the existence of a `floor' to real wages would be conceded by most development economists, including those of neo-classical persuasion (Arthur Lewis 1954, after all used this assumption in the context of an otherwise neo-classical universe and his celebrated model still constitutes a staple for development economics), the fact that the consequent absence of full resource utilisation makes neo-classical policy-prescriptions infructuous is scarcely perceived by many.

IV

Until now I have argued that the neo-classical prescriptions for promoting efficiency are based on the assumption that the economy spontaneously achieves full employment of all resources and that the violation of this assumption in reality, owing to problems of aggregate demand or of disproportionality, makes these prescriptions not only infructuous, but possibly even counterproductive. Now I wish to advance the stronger proposition, namely that in the concrete context of underdeveloped economies neo-classical efficiency-promotion measures, which come to us in the guise of prescriptions by the Bretton Woods institutions, are actually counterproductive, that far from increasing output for given endowments they have the very opposite effect ceteris paribus of reducing output. The most significant elements of these prescriptions relate to trade and financial liberalisation, which exert an output-contracting effect through three simultaneously operating mechanisms.

First, there is an enlargement of the trade deficit, and hence of the current account deficit, owing to import liberalisation. This is because in some instances foreign goods are cheaper given their quality, in some instances (e.g.investment goods) they come together with foreign credit which takes care of financing problems, and more generally because a preference for foreign goods characterises the elite in third world societies. It follows that even if import liberalisation is accompanied by a domestic consumption splurge the increase in trade deficit is not fully offset by the increase in autonomous components of expenditure, such as government expenditure, investment or autonomous consumption expenditure. As a result the expansion of the trade deficit has a contractionary effect on the economy. In other words the increase in net imports stimulated by import liberalisation entails, in part at least, a shift of demand away from some domestically-produced commodities, and it is this which results in contraction taking the form of deindustrialisation.

On the other hand, the consequent expansion of the current account deficit requires larger external borrowing on the part of the country. In effect therefore the country borrows from abroad to finance a destruction of its own industry, or undertakes what might be called `a debt-financed deindustrialisation'. Debt-financed deindustrialiation is what capitalist countries tried to impose on one another during the Great Depression as a part of `beggar-my-neighbour' policies (these were usually accompanied by exchange rate depreciations by the inflicting economies) for sustaining their own levels of domestic activity. What we are witnessing today is a similar imposition of debt-financed
deindustrialisation on third world countries by the advanced capitalist economies through the conditionalities of the Bretton Woods institutions. This is justified in the name of achieving efficiency of resource use but it results in fact in greater inefficiency of resource use since domestic output and employment shrinks and capacity is rendered idle.

The second mechanism has already been alluded to. Trade liberalisation results in an increase in exports of primary commodities whose domestic prices are often held below their world prices (at the prevailing exchange rate) in the pre-liberalisation period. Enlarged exports however come through a reduction of the domestic availability of wage goods, and this lowers the maximal industrial output compatible with a given real wage rate, giving rise to an overall contraction in the economy even in the absence of any demand constraint. The reduced domestic availability of wage goods arises either because such goods are directly exported and their output cannot be augmented or because land hitherto devoted to them is shifted to the production of exportables and this is not compensated by any offsetting increase in their yields. In either case, however, underlying reduced domestic availability is the fact that output does not increase sufficiently to compensate for the increase in exports, and this results from the fact that public investment in overheads such as irrigation, power and extension which is essential for an increase in output in the agricultural sector is not stepped up appropriately. Indeed the liberalisation package entails on the contrary a curtailment of such investment.

This second mechanism however may operate in a somewhat different manner. Primary production for export typically tends to be less labour-using than the production for domestic use which it replaces. As a result, trade liberalisation, even when there is no expansion in the trade deficit, i.e. even when the first mechanism mentioned above does not operate, gives rise to a lower level of employment for this reason. A lower employment of this sort of course need not mean a contractionary effect on output. But it does because it is associated with a reduction in the share of wages and hence in the marginal (and average) propensity to consume in the economy. Trade liberalisation in other words, through its impact on the composition of output, can have exactly the same contractionary effect on output as a rise in the 'degree of monopoly' a la Kalecki. The initial reduction in employment owing to the changed composition of output, with export-oriented primary production replacing production for domestic use, is followed by a further reduction in employment owing to the contractionary consequences of the associated shift in income shares, even when there is no increase in the trade deficit. This contractionary effect on output of course may be counteracted through larger involuntary stock-holding (as in the case of foodgrains in India until recently), but this does not negate the fact that a shift to primary exports following trade liberalisation has a contractionary effect via lower employment in export production.

It follows then that the shift to primary exports has a contractionary effect both through the demand side, i.e. via lower labour-use in such export goods (e.g. orchards and prawn fisheries), and through the supply-side, via a reduced domestic availability of wage goods which reduces overall employment and output for a given real wage. We need not be detained by the question of which of these effects dominates, since both are contractionary; it is worth noting however that if the demand side effect dominates then
disproportionality ceases to matter and the effect of trade liberalisation can be analysed in purely Keynesian terms: trade liberalisation then can be seen simply as precipitating a contraction even with no increase in the trade deficit and aggravating it if there is such an increase.

The third mechanism operates because of financial liberalisation which makes capital flows into and out of the country easy. This is done typically in the expectation of obtaining larger direct foreign investment. But while DFI inflows remain in general meagre (which does not of course preclude their moving in significant amounts to some countries for some time), it exposes the domestic economy to the vicissitudes of international speculation. Retaining the 'confidence' of international speculators so that 'hot money' does not flow out en masse becomes the obsession of economic policy, and the standard response is deflation. Deflation is always to the liking of finance capital: the inflation rate, which adversely affects rentier interests, is kept down, even at the cost of unemployment; high interest rates which accompany deflation ensure for them a high rate of return; public enterprises are "sold for a song" for keeping down the fiscal deficit which rentiers can buy up; the credit squeeze forces domestic small capitalists to sell out to international finance capital; and every episode of actual or threatened capital flight is followed by even larger concessions, for the acquisition of mines, of prime property, of lands etc. in order to entice finance capital to stay. In the name of improving the efficiency of financial markets, supposedly so necessary for drawing FDI, the country witnesses a loss of its economic sovereignty, an attenuation of its political independence through a restriction of its political choices (since the government must always be one that inspires 'confidence' among international financiers), and an attack on the economic conditions and political rights of its working masses (since strikes and wage increases frighten international speculators and can cause a capital flight). More germane to our argument however is the fact that deflation, a fall-out of the quest for efficiency (in the financial sector) leads to output contraction, and hence the very opposite result of promoting inefficiency. It is instructive that even South East Asian countries which appeared to have avoided this denouement and combined high growth with free capital movements (this was not the case with East Asia which had strict controls over capital movements), are now learning this painful lesson that free capital movements can, and ultimately do, precipitate deflation and stagnation.

\[V\]

Let me pull together the threads of the argument. While there would not be much disagreement over the view that a vector-wise larger output for given endowments would constitute an improvement in the efficiency of resource use, such an improvement does not follow from the replacement of an activity set with a lower marginal rate of transformation of one set of goods into another by another activity set with a higher rate. It would follow if the economy always had full employment of resources, but it does not; hence the two perceptions of efficiency are not synonymous. Neo-classical economists often forget this elementary fact with tragic consequences. Thus in Russia, GDP was made to fall by a third, enormous amounts of capital stock built up with the sacrifices of
generations were destroyed, millions thrown out of work, poverty which was banished during the Soviet era brought back with a vengeance, and the mortality rate made to climb to a level where the population itself started declining, all in the name of achieving efficiency of resource use!

When an activity involving a lower rate of transformation is substituted by an activity involving a higher rate, the latter releases some resources, i.e. some equipment and labour engaged in the latter become idle. The former however does not necessarily absorb them. If the former activity is foreign trade, then, unless the export sector's output was demand-constrained to start with, we find necessarily an actual contraction of resource employment and of overall output. But if the former activity is not external trade but one domestic process replacing another, e.g. machine-made goods replacing hand-made goods, the magnitude of resource employment necessarily falls while the overall output may or may not fall. If the replacement does not cause any deficiency of aggregate demand through a rise in profit-share, the overall output need not decline (as in the case of Hicks' (1973) 'classical traverse'). In the short-run therefore the quest for efficiency necessarily causes resource unemployment without raising overall output; indeed overall output is in many cases lower. The belief entertained is that this would be more than made up in the long-run through appropriate investment behaviour, so that the output profile would overtake what would have obtained otherwise. This belief however, which even Ricardo entertained in his discussion on machinery, does not reckon with the specificities of capitalists' investment behaviour. Since higher profits do not necessarily lead to higher investment, neither the resource employment profile nor the output profile necessarily rises above what would otherwise have obtained in the absence of activity replacement. Indeed in the case of trade liberalisation accompanied by restrictions on public investment both profiles would be lower than would have been the case.

This has important implications. It may be efficient for a country to carry on with what appear to be inefficient activities. A number of propositions which are argued on grounds of efficiency, e.g. loss-making units should be closed down, 'obsolete' technology should be weeded out, reservation of items for cottage industries should go, can in fact be criticised precisely on the grounds of efficiency, since all of them entail output losses. In short, recognition that ours is not a full-employment economy implies that policy measures have to be scrutinised for their short and long term effects on output and employment profiles and this may dictate, for the production of any good, the use of a combination of production techniques rather than one 'efficient' technique, a combination of activities rather than one 'efficient' activity. Planning alone provides the framework within which such assessments can be made. But the scope for meaningful planning has to be created first by halting and reversing the rush towards liberalisation that is currently occurring.
BIBLIOGRAPHY


Patnaik P. (1972) "Disproportionality Crisis and Cyclical Growth", EPW, Annual Number.


(1996a) "Should Domestic Prices be Equated to World Prices?", EPW, Special Number.

Patnaik U. (1996) "Export-Oriented Agriculture and Food Security in Developing
Countries and in India", EPW Special Number.
