

III 41

Bamber

IMPORT SUBSTITUTION
as an
INDUSTRIALIZATION STRATEGY*

ISEN 72452
PNABF 563

by

John H. Power
Williams College

*A paper to be presented at the World Conference of the Society for International Development, March 16-18, in New York.

The author is indebted to his colleagues at Williams College -- especially Hans W. Singer and Gordon C. Winston -- for helpful discussion of some of these issues.

Financial assistance from the Williams Research Project for the study of Import Substitution and Economic Policy in Economic Development, which is supported by the Agency for International Development, is gratefully acknowledged.

Williamstown, Massachusetts
January, 1966

A.I.D.
Reference Center
Room 1656 NS

1

Import Substitution as an Industrialization Strategy

John H. Power
Williams College

Introduction

An industrialization strategy biased toward import substitution is, I think, almost an inevitable phenomenon in less developed countries. Often it emerges in an apparently natural way -- perhaps even inadvertently -- following the imposition of import controls as a crude reflex response to a balance of payments problem. What is initially viewed as a curb on the consumption of less essential imports soon becomes rationalized as a protective device to encourage production of their substitutes. Since the market is already there for the taking, a sufficient degree of protection will promise quick easy gains in industrial output, primarily in finished consumption goods even though materials and parts, as well as capital equipment, must be imported.

Moreover a dual theoretical rationale for an import substitution orientation in development strategy easily emerges to justify its continuance.¹ The world demand for primary commodities is held to be growing too slowly for their export to play a leading role in economic growth. The export of manufactures in competition with the developed countries appears unpromising both because of the disadvantages of technological backwardness and diseconomies of small scale and because of protection in the wealthy countries where the principal markets are. This leaves production for the home market, balanced

¹The description of this rationale that follows is obviously based on the well-known theories of R. Nurkse and W. A. Lewis. It by no means does justice, however, to the sophistication and depth of their views. Popular versions of famous theories, which serve as the rationale for political decisions, are often unfortunately only caricatures of the originals.

in relation to home market demand, as the seemingly most promising avenue of growth.

The first aspect of the rationale for an import substitution bias is a defensive one, implying, in effect, that this is the only way out of a difficulty. The other has a more positive character, however, suggesting the possibility of an emerging self-sustaining growth mechanism. Behind import controls the domestic manufacturer can obtain high prices for goods even of inferior quality, thereby "earning" high profits that can be saved and reinvested. This means turning the terms of trade against agriculture (and other non-protected sectors) to create the saving for industrial growth. Eventually, it is hoped, agriculture too will benefit as industrial progress reduces costs and prices, and as industrial expansion offers higher productivity employment to rural labor.

This line of thinking has a natural appeal in newly developing countries. It promises less dependence on traditional exports -- hence, less fear that "export lag" or declining terms of trade (or both) will inhibit development. The problem of competition with the more advanced technology in developed countries is avoided by curbing imports and not encouraging new exports. The balance of payments problem -- the need to import growing quantities of capital goods -- is solved instead by saving foreign exchange through import substitution. The difficult problems associated with increasing agricultural productivity can be pushed into the background, since it is the twist of the terms of trade that initially wrests saving from agriculture. And, via this shift in income distribution to the "capitalist sector" and the latter's response, a cycle of profit-saving-reinvestment-increased productivity-higher profit-etc. can supposedly emerge to render growth self-sustaining.

While this is, no doubt, an over-simplified view of the origins of an import substitution strategy as well as of its rationale, it is suggestive, I think, of the experience of a number of developing countries that have had some success in the first stage of implementation of such a strategy -- the take-over of an existing market for consumption goods from the foreign supplier. Common also to their experience, however, seems to be the greater difficulty of meeting the challenges that lie beyond the first stage -- namely, extending production backward to intermediate goods, capital goods and raw materials, and breaking into the world market with exports of manufactures. Yet these are crucial to ultimate success in that, without one or both, the pace of industrial growth must falter and the emergence of a self-sustaining growth mechanism is frustrated.

This is seen most easily by focussing on the market limits for growing industrial output. The first stage of import substitution involves absorbing an existing market -- which is, of course, primarily for finished consumption goods. No market constraints need appear until this first stage is largely completed, but when this point is reached one or more of three developments must have occurred to prevent a market constraint from retarding industrial growth. The domestic market itself must have already begun to expand rapidly; industrial goods must have begun to penetrate the export market; or investment must have begun to move away from finished consumption goods to capital goods, intermediate goods and raw materials (both to substitute for imports and to add to the availability of investment goods).²

²If one assumes that investment requirements are given by growth targets and capital coefficients, any additional domestic production of investment goods within the limits of these requirements can be considered import substitution.

Moreover, the first possibility -- a more rapidly expanding home market -- itself requires some combination of expansion in investment goods production, import substitution, or export expansion to make possible the required rate of increase in investment. Simply expanding consumption goods production, unless it results in growing exports or import replacement, is not compatible with sustained growth.

The conclusion is, then, that beyond the first stage in an import substitution strategy -- the expansion, behind protection, of finished consumption goods production to the limits of the domestic markets -- lies the necessity of developing production of intermediate goods, capital goods, and raw materials; or expanding exports; or both. And here is where many developing countries have found the going more difficult, and the challenges to economic policy-making more demanding.

It is a simple matter to formulate and implement a policy of protection for the first stage. Often this happens almost inadvertently, as was suggested above. But the crude policies of protection that may serve adequately in the first stage, and the economic structure that they encourage, are likely, in my opinion, to become barriers to growth in subsequent stages.³ Why this is so, and what might be done to prevent it or correct it, is the subject of this paper.

³ Because of a rising import bill of materials, parts and equipment to sustain production in the protected industries, and because of a resistance on the part of unprotected sectors and income groups (e.g. agriculture and labor) to any deterioration of their terms of trade, barriers to growth taking the form of balance of payments difficulties and inflation may arise long before the first stage is completed.

Section I comprises a discussion of the emergence of barriers to growth under three headings: economic inefficiency (misallocation of resources); technical inefficiency (failure to minimize costs); and the saving gap (failure to achieve an adequate rise in domestic saving). While the distinctions may not always seem clear-cut, this scheme of presentation does serve to emphasize that a naive import substitution strategy can impede growth via an adverse influence on the marginal saving rate, as well as on the social product; and that its influence on the latter over time depends as much on inducements to efficiency and innovation as on resource allocation.

There follows in Section II a brief summary of the policy implications implicit in the critique of such an import substitution strategy.

I

Economic Inefficiency

An import substitution bias implies a balance of payments policy that favors import control or restriction (often via exchange control) over export encouragement. This, in turn, implies a lower value for foreign exchange than that appropriate to a policy of equal encouragement to exports and import substitution. If market prices could be taken to represent unit costs and utilities at the margin, the resulting resource allocation would require a greater value of resources at the margin to save a unit of foreign exchange through import substitution than to earn a unit of foreign exchange through export expansion.

Since this kind of welfare loss is generally well understood, the persistence of this direction of bias in balance of payments policies suggests either that considerations other than economic efficiency are considered to be more important, or that the assumptions

underlying this kind of welfare judgment are considered to be invalid. About all an economist can do with regard to the former is to point to the cost and, since this emerges anyway in a discussion of economic efficiency, I will focus on the latter.

Before turning to the validity of the assumptions on which welfare judgments against interference with free market results are based, however, we should note another kind of misallocation that appears to be both very likely and very substantial in the context of an import substitution strategy. That is the bias against production of intermediate goods, capital goods and raw materials. The reason is, of course, that these are inputs in the industries which develop in the first stage and, as such, are usually more liberally imported than are the finished consumption goods that compete with the emerging domestic industries. This means not only a bias against vertical balance in import substitution -- i.e., backward linkage is discouraged -- but also an inflated and irrationally differentiated structure of protection at the finishing stages of production.

This is so because the total rate of protection depends not only on the particular rate of protection that applies to the product of that industry, but also on the particular rates that apply at the preceding stage in the production process. The former acts as a subsidy while the latter act as taxes on value added in a particular industry. It may be useful to put these relationships more formally at this point.⁴

Let X_i represent the output of any industry and $\sum_j a_{ji} X_j$ its intermediate inputs, both valued at given world prices -- i.e., the prices that would prevail with free trade.

⁴The formal exposition is patterned closely after that of Harry G. Johnson in his "Tariffs and Economic Development," Journal of Development Studies, October, 1965, p. 20.

Then

$$V_i = X_i - \sum_j a_{ji} X_j \quad (1)$$

is value added at free trade prices, and

$$V_i(1+T_i) = X_i(1+t_i) - \sum_j a_{ji}(1+t_j)X_j \quad (2)$$

is actual value added under the system of protection. The t 's represent the proportions by which the system of protection permits the actual domestic prices of the outputs of various industries to exceed their free trade prices, while T_i is the total rate of protection of the i^{th} industry -- the proportion by which its value added can exceed what would be its free trade value. This can be written also as

$$V_i(1+T_i) = X_i + t_i X_i - \sum_j a_{ji} X_j - \sum_j a_{ji} t_j X_j$$

and by substituting (1) in the right-hand side

$$V_i(1+T_i) = V_i + t_i X_i - \sum_j a_{ji} t_j X_j$$

We can solve this for the total rate of protection

$$T_i = \frac{V_i + t_i X_i - \sum_j a_{ji} t_j X_j}{V_i} - 1 \quad \text{or}$$

$$T_i = \frac{t_i - \sum_j a_{ji} t_j}{V_i/X_i} \quad (3)$$

From (3) we can see that the total rate of protection of an industry will be greater the greater is its own particular rate of protection, the smaller are the particular rates of protection of its supplying industries, and the smaller is the proportion of its value added to the total value of its output. Now consider the distorted pattern of protection that can result from a policy of restricting most severely the import of consumption goods, while permitting inputs into these industries to be more liberally imported.

First, as was noted above, exports are penalized by the lower value of foreign exchange that is consistent with the bias toward import restriction. But the extent of the bias can be much greater than the particular degrees of protection would suggest. Suppose, for example, that the protective device employed were a 50 per cent duty on consumption goods while intermediate inputs could be imported at free trade prices. Then if value added in manufacturing (at free trade prices) were 25 per cent of total (free trade) value, equation (3) tells us that the total degree of protection would be 200 per cent! And this may not be an unrealistic example. I suspect that in circumstances like these there are many such industries in which the cost of saving a unit of foreign exchange is several times that of earning a unit through exports.⁵

A similar magnification of the distortion in degrees of protection occurs, of course, between industries producing consumption goods and those producing materials, parts and equipment when the latter are more liberally imported. Thus, the bias against backward-linkage import substitution is more pronounced than a simple comparison of particular rates of protection would suggest. Moreover, the resulting relative lack of domestic sources of supply for these inputs, together with the fact that the total degree of protection is inversely related to the (proportional) value added contribution of the industry, means that such a system of protection particularly encourages heavy users of foreign exchange. Finally, we should note that in protecting the balance of payments via import restriction, it is a very common practice to restrict most severely the least essential

⁵This assumes no terms of trade effect, but see below, p. 9. Also see below, p. 19, for some evidence on this point from Pakistan data.

imports. This tends to bias import substitution, albeit perhaps inadvertently, in favor of less essential industries.

The conclusion is that an import substitution bias in development strategy, when accompanied -- as is, I think, typical -- by relatively liberal import policies with respect to "essential" imports (both in the form of inputs for domestic industries and special categories of consumption goods), can create a rather extreme distortion of incentives away from the pattern that would result from free markets. Moreover, the direction of distortion appears to be unfortunate in that it particularly discourages export expansion and backward-linkage import substitution, one or both of which is crucial to sustained industrial growth, as noted above; while it gives the greatest encouragement to industries most heavily requiring foreign exchange to produce less essential products.

Despite this, protection is often defended as a means of correcting "market failures," and once we abandon the assumption that free market prices are necessarily the best welfare indicators we are obliged to consider several more ^{or} less respectable arguments for this view.

Johnson has argued that the only economic justification for tariffs is the terms of trade effect of trading more or less.⁶ For an open economy that can affect via trade the prices of the goods it buys and sells, full Paretoan optimality requires equality between the ratio of domestic prices of exports and imports and the marginal terms of trade, rather than the international price ratios. This means restricting trade until

⁶Op. cit., p. 8.

$$\frac{p_h}{q_h} = \frac{p_w}{q_w} \cdot \frac{1 + \frac{1}{n}}{1 - \frac{1}{e}} \quad (4)$$

where p represents the price of imports and q the price of exports, the subscripts h and w indicating home and world prices, while n and e are the world elasticities of supply of imports and demand for exports, respectively.⁷ This could be accomplished by establishing a dual exchange rate system, the price of foreign exchange for imports exceeding that for exports in the proportion

$$C = \frac{1 + \frac{1}{n}}{1 - \frac{1}{e}} \quad (\text{the "correction" for terms of trade effects}).$$

The more common method of favoring import substitution over exports, however, is by tariffs or exchange control. In this case imports should be restricted (and the price of foreign exchange reduced) until the condition described by equation (4) is met.

How does this description of optimality relate to the picture of misallocation which preceded it? First, the general degree of protection would have to be equal to $(C - 1)$ and it should apply uniformly at all stages of the production process. If all industries were effectively so protected the total rate of protection for each industry would be equal to its particular rate -- i.e., the uniform general rate $(C - 1)$. This can be seen by setting $t_i = t_j = t$ in equation (3)

$$T_i = \frac{t - \sum_j a_{ji} t}{V_i/X_i} \quad (3a)$$

⁷S. Alexander, "Devaluation versus Import Restriction as an Instrument for Improving the Trade Balance" IMF Staff Papers, April 1951, p. 379.

and since

$$V_i/X_i = \frac{X_i - \sum_j a_{ji} X_i}{X_i}$$

$$\tau_i = \frac{t(1 - \sum_j a_{ji})}{1 - \sum_j a_{ji}} = t$$

As Johnson has pointed out, however, for a trading country the export industries' rates of protection (in the absence of export taxes) must effectively be zero.⁸ If any of these exportables are inputs in other domestic industries, total rates of protection will differ among industries in accordance with their use of these inputs and their (proportional) value added contributions. This does not mean, however, that resources will be misallocated in this case, since the resulting pattern of protection is just what is needed to bring domestic rates of transformation between exports and imports in line with the international marginal terms of trade.

Put this way there seems to be a perfectly respectable argument for protection when the marginal terms of trade are below the average. The appropriate rate of protection is $(C - \bar{t})$ and to avoid inter-industry distortions in the pattern of protection the simple rule is a uniform rate for all industries.

⁸ Johnson, *op. cit.*, p. 22. This assumes that the export industries are competitive. If they are not they can act as discriminating monopolists behind protection. But if they (all) exploit their monopoly power in the world market there is no need for the government to do so via a policy of protection. Since, however, the most important cases of high values for C are likely to be associated with primary commodity exports, the assumption that the home prices of exports are equal to the world prices (or below them by the amount of export taxes where these play a role in the adjustment to the condition described by equation (4)) is probably a reasonable one.

There are, however, two serious weaknesses in the argument. First is the assumption of a single elasticity of demand for all exports⁹ -- presumably a weighted average. For most developing countries, however, the primary reason for a value of C in excess of unity is the relatively low elasticity of demand for one or a few primary exports which weigh heavily in the total. Basing a system of protection on the weighted average elasticity would mean a strong bias against all of the other (actual and potential) export industries.

This strongly suggests that the few exports with low elasticity of demand be removed from the jurisdiction of general trade policies and treated as special cases requiring taxes, supply restrictions, or something of the sort. If this is done, however, the second weakness, alluded to above, of the terms of trade argument for protection becomes more apparent. For the argument assumes no retaliation -- an assumption that might be valid in the case of modest across-the-board protection, but which can be held with less assurance in the case of a much stronger price influence concentrated on one or a few commodities. It is, in other words, precisely where the potential terms of trade gain from trade restriction are greatest that the threat of retaliation is most likely.

A situation where a country can avoid serious misallocation of resources only by means that hurt others and invite retaliation calls for some kind of international agreement to resolve the inherent conflict. So what appears to remain as valid of the terms of trade argument for protection is that it should apply selectively to a relatively

⁹ A single elasticity of supply for all imports is also assumed, but this is normally more defensible.

few commodities, and that this should lead to an international agreement on prices. Short of achieving this kind of international cooperation it is probably in the interest of the developing countries to apply unilaterally supply restriction schemes for these commodities because of the paramount importance of freeing all other exports from the penalty of an undervaluation of foreign exchange.

What of the other common arguments for protection: infant industry, external economies, and factor price disequilibrium? Johnson has argued that since these do not involve a failure of international price ratios (or marginal terms of trade) to represent true opportunity costs in international trade, any policy designed to implement these arguments that simultaneously disturbs the relation between domestic and international price ratios will thereby create, as well as correct, distortion.¹⁰ The appropriate measures to bring true social costs and values into line in each of these cases would involve a system of taxes and subsidies, not a system of trade restriction. His argument is correct on the assumption that the government can in fact implement a fiscal policy that itself involves less distortion than, say, taxing imports. If, on the other hand, the fiscal measures available are quite limited for institutional reasons, the argument is weakened. Nevertheless, the advantages of fiscal remedies are so pronounced in each of these cases that the argument for trade restriction is, in my opinion, very dubious. Let me consider each of these cases, in turn.

The infant industry argument has two roots. One is the relation of efficiency to scale and the other is the relation of efficiency to time. Increasing returns to scale and a time-consuming learning process then serve as a valid basis of the case for protection

¹⁰ Johnson, *op. cit.*, pp. 8-9. It is assumed here that the relation has already been corrected for terms of trade effects.

of infant industries.

The logic of the argument calls for specific protection of certain industries, however, rather than general protection of the sort described above (in the introduction) as an import substitution strategy. The reason is two-fold. First, industries differ with respect to their scale-efficiency and time-efficiency relationships. Long-run comparative advantage depends in part, then, on their relative differences in the response of efficiency to scale and time. Second, the extent and pace of response is itself likely to be a function of the concentration of resources. That is, even if all industries had the same response functions it would normally pay to concentrate on fewer industries, at least up to a certain point in their development, rather than to disperse resources across a broad front.

This is obvious for the scale-efficiency relationship, but it may be true for the time-efficiency relationship, as well. That is, a concentration rather than a dispersion of investment, technical and organizational skills, and education and training may mean a more rapid average rate of progress in efficiency for the whole economy. Scitovsky emphasizes in this connection the relationship between the pace of growth of an industry and its rate of innovations, concluding that the rapid pace made possible for some industries (as opposed to the pedestrian rate for all in a balanced growth context) means concentrated growth would permit and encourage a more rapid overall rate of increase of productivity.¹¹

So the logic of the infant industry argument calls for concentrated industrial growth

¹¹T. Scitovsky, "Growth--Balanced or Unbalanced?", Papers on Value and Growth (Stanford University Press, 1964), pp. 107-109.

rather than growth balanced in relation to domestic demand, implying greater emphasis on exports and less on import substitution. But protection against imports penalizes exports via the lower value of foreign exchange. An optimal set of policies would include, therefore, subsidies to exports from the selected industries equal to their rates of protection, both set (somehow) equal to the (discounted) future relative advantage of these industries.

There remains another problem, however. We do not have in this case a uniform degree of protection across the economy, with the result that industries using as inputs the outputs of the protected industries would be penalized, the degree of penalty depending not only on the amount used of the various protected goods, but also on each industries (proportional) value added contribution. In order to avoid misallocation from this source, these differential penalties would have to be offset by matching subsidies. At this point Johnson's argument for avoiding all of this patching up by subsidizing directly the "infant industries" in the first place begins to look very sound.¹²

The presence of external economies is another reason sometimes given to defend protection. Broadly viewed external economies comprise all elements of interdependence among industries, both direct and through the market. It will be convenient for what follows to distinguish within these simply between interdependence in production and interdependence in consumption.

An emphasis on interdependence in consumption leads to "horizontally balanced" growth in line with market demand. While this yields external economies via

¹²Johnson also rightly stresses the loss in consumer surplus from distorting the relationship of domestic to international prices. Op. cit., p. 10.

complementarities in consumption, it means that external diseconomies prevail in interdependence in production via competition for scarce resources. In contrast, an emphasis on interdependence in production leads to "vertically balanced" growth in line with backward and forward linkages in production. This, of course, yields external economies on the supply side, but ignores complementarities in consumption, thus requiring an ability to sell in the world market to solve the demand problem.

There seems to be no obvious reason for giving greater emphasis in general to one kind of interdependence over another -- i.e., for generally favoring horizontal balance and import substitution over vertical balance and export expansion because of the existence of external economies. On other grounds a preference could be established. For example, if the terms of trade were the only criterion, horizontal balance should be preferred. Or, if saving and growth were the only criterion, vertical balance should be preferred.¹³ But if the world market is available to fill the gaps in both demand and supply there is no critical advantage in emphasizing either domestic supply balance or demand balance.

The essence of the external economies argument for protection, however, is the inability of private decisions based on market criteria to take account of the results of interdependence. While this may be a nearly universal phenomenon in the context of a dynamic growth process, there will be certain areas where the total gains from interrelated decisions can be judged to be particularly large in relation to what the market promises. These then should be treated in a manner similar to that suggested above for infant

¹³See below, p.

industries. And the same argument for subsidy rather than protection applies. In the absence of any reason for altering the relation between domestic and international prices, the only defense for protection in these cases would be that import duties were the least inefficient method of taxation available to the government.

I turn finally to the factor price disequilibrium argument for protection of manufacturing. Put in its simplest form it is that wage rates in manufacturing¹⁴ exceed the opportunity cost of labor from other sectors and this puts domestic manufactures at an unwarranted disadvantage with imports. Protection of domestic manufactures is then the suggested remedy.

Lary has argued correctly that, since the factor price disequilibrium applies to manufacture for export as well as for import substitution, the former should be equally encouraged. He has advocated a dual exchange rate -- a higher price of foreign exchange for both exporting and importing manufactures and a lower price for trading agricultural products.¹⁵ This is a step in the right direction in that it corrects an unwarranted bias against exports. But there remains a bias against the use of labor in the factor mix and a bias against the use of domestic manufactures as inputs. These can be eliminated along with the others, however, by a simple subsidy on the employment of labor where its market price is above its opportunity cost. Again it seems that restriction of imports is an inept and costly way to correct a market failure.

The conclusion that emerges from this analysis of the economic efficiency of

¹⁴This need not be restricted to the manufacturing sector.

¹⁵H. Lary, "Economic Development and the Capacity to Import -- National Policies," in Lectures on Economic Development (Istanbul University, 1958).

favoring import substitution via protection is somewhat depressing. First, the system as it develops in the first stage is likely to misallocate resources by means of a strong bias against exports, against backward-linkage import substitution, in favor of less essential industries, and in favor of heavy users of foreign exchange. Nor can any of the arguments for correction of market failures bolster very much the case for this kind of protection. Even the terms of trade argument appears dubious when a concern with low world demand elasticities for a few exports dictates an under-valuation of foreign exchange that heavily penalizes all other exports (actual and potential).

This is a qualitative judgment, of course, and the really important question is its quantitative significance. Some striking evidence bearing on this has been presented recently by Soligo and Stern for Pakistan.¹⁶ Using the Tims-Stern input-output model for 1963/64, they have calculated implicit rates of protection for forty-eight manufacturing industries. Their "implicit" rates correspond to my "total" rates, but with an important difference. In my notation theirs is

$$U_i = \frac{t_i - \sum_j a_{ji} t_j}{V_i/X_i + (t_i - \sum_j a_{ji} t_j)} \quad (5)$$

which can be compared with equation (3) above.

Put more simply the difference is this:

$$T_i = \frac{W_i - V_i}{V_i} \quad \text{and}$$

¹⁶R. Soligo and J. Stern, "Tariff Protection, Import Substitution and Investment Efficiency," Pakistan Development Review, Summer 1965, pp. 249-270.

$$U_i = \frac{W_i - V_i}{W_i} \quad (6)$$

where W_i is $V_i(1 + T_i)$, actual value added under the system of protection (see equation (2) above). It follows that their implicit rate of protection

$$U_i = T_i \cdot \frac{V_i}{W_i} = \frac{T_i}{T_i + 1}$$

The Soligo-Stern measure of the rate of protection has one very decisive advantage in that it can apply to cases where V_i is negative, which my T cannot. And this turns out to be of considerable importance in assessing the Pakistan data.

A negative V_i means that for an industry, the value of output at world prices is less than the value of intermediate inputs at world prices. Abandoning production and importing the finished product would save both foreign exchange and domestic resources. Or, as Soligo and Stern put it, "the average revenue product of capital and labor is negative."

This was found to be true for twenty-three industries in Pakistan, including food processing, beverages, cigarettes, textiles and wearing apparel, petroleum and coal products, furniture, cycles, and motor vehicles. Since V was negative for each of these industries, U was greater than unity. Among the twenty-five with values for U less than unity (implying no absolute waste of resources from the activity), the values varied widely from -0.27 for grain milling to 0.92 for matches (suggesting relative waste of resources). The pattern of differential protection corresponds to what one would expect from an import substitution bias.

In general i) consumer goods are much more heavily protected than either intermediate or investment and related goods, ii) within the consumer goods industries, non-essentials, such as beverages and cigarettes, are much

more heavily protected than essential industries such as grain and rice milling, salt and tea, iii) textiles are the most heavily protected group of industries, although the protection is approximately the same for all components of the group and iv) the least protected industries are those producing heavy machinery, both electrical and non-electrical, and transport equipment other than motor vehicles and cycles. Fertilizer is also among the least protected group.¹⁷

Technical Inefficiency

Soligo and Stern report only the values for U -- not for W and V , its components. In the case of $U > 1$, there is what we might call an absolute disadvantage, and there is no doubt about the uneconomic nature of the activity. In cases of $U < 1$, however, a relatively high value could mean either a relatively high W , or a relatively low V , or both. It is possible, in other words, for a relatively high implicit rate of protection for an industry to cover high factor incomes or relative inefficiency, or both.

I have no evidence to present on this point, but it seems to me that for several reasons we might expect relative inefficiency to be widespread among those industries with the highest implicit rates of protection. First, a system of protection of the kind under discussion will inevitably include under its umbrella all kinds of comparatively disadvantageous industries. Second, for others (including "infant industries"), the protection against foreign competition insures monopolistic or oligopolistic market positions that take the edge off the drive for efficiency and technical progress. Third, the dispersion of resources in horizontally-balanced industrial growth sacrifices potential gains from economies of scale and the stimulus to innovations and learning from faster concentrated growth.

¹⁷ Op. cit., p. 259.

It is possible, on the other hand, that some of these highly protected industries have a real comparative advantage and are reasonably efficient, so that the protection permits high factor incomes. The factor-price disequilibrium case fits here. The protection may permit the industry to pay the required excess above labor's opportunity cost that the market dictates. We have seen above, however, that this is no more than a third-best sort of argument for protection.

Finally, however, the high degree of protection may mean high profits; and high profits suggest the possibility of a saving-reinvestment growth mechanism. This brings us, then, to the effects on saving of an import-substitution bias in development strategy.

The Saving Gap¹⁸

I have argued above (page 4) that to carry an import substitution strategy successfully beyond the first stage requires either breaking into the export market or extending production backward to materials, intermediate goods, and equipment. Continuing expansion of finished consumption goods for the domestic market, while perfectly compatible with a non-accumulation economy (wherein the growth of income occurs exogenously), can permit capital accumulation (other than accumulation of stocks) only so long as it reduces consumption goods imports. When the first stage is completed, of course, this is no longer possible. But even during the first stage there is a very real possibility that a bias toward the production of consumption goods balanced in relation to domestic demand will tend to

¹⁸The following discussion owes much to analyses of the Pakistan experience. See A. R. Kahn, "Import Substitution, Consumption Liberalization and Export Expansion," and my "Industrialization in Pakistan: A Case of Frustrated Take-Off?," both in Pakistan Development Review, Summer 1963.

erode the constraints on consumption that are needed to permit accelerating growth.

To see how this might be so, consider first the identity

$$C_d + I_d + E_d \equiv C_m + C_d + S \quad (7)$$

where C_d , I_d , and E_d are value added in domestic production for consumption, investment, and exports. S is domestic saving and C_m is the imported component of consumption. The left-hand side represents the national product and the right-hand side, the disposal of national income.

A rise in any component of the left-hand side implies an equal rise in saving (and investment -- domestic or foreign) if consumption does not rise. Thus a case of pure import substitution (the rise in C_d being matched precisely by a fall in C_m) increases saving exactly as does a rise in the production of capital goods or exports when consumption is constant. The analysis can be extended to the more general case in which consumption rises by some proportion of the rise in national product, and the conclusion is the same. The change in saving associated with a rise in output depends on the change in consumption regardless of the kind of goods the output increase embodies.

The key question is, then, how the marginal consumption rate might be affected by alternative patterns of investment leading to different mixes of output increase. This is usually analyzed in terms of the associated sectoral income increases and saving propensities, but I propose to look at it briefly from the other side -- to consider how the supply mix itself can affect consumption and saving.

Consider the following simple model of a closed economy.

$$\Delta Y = kI \quad (i)$$

$$\Delta S = s\Delta Y \quad (ii)$$

$$\Delta I = \alpha I \quad (iii)$$

$$\Delta S = \Delta I \quad (iv)$$

Y is national product, I is investment, S is saving, s is the marginal propensity to save, k is the incremental output-capital ratio (identical for all sectors of the economy), and α is the proportion of investment allocated to the investment goods sector.

Given ΔY (the growth target) and the investment coefficient, k , these four equations determine I , ΔS , ΔI and either s or α if the other is given. If both are given the system is overdetermined. That is, consistency is required between the marginal saving rate and the proportion of investment allocated to the investment goods sector. This leaves open the question of how consistency is achieved, however. If saving propensities govern, α must adjust to s -- the allocation of investment must respond to the pattern of final demand. Alternatively, however, marginal saving could be constrained by the output mix of consumption and investment goods as determined by the investment allocation -- i.e., by α . It is this latter possibility that I want to explore in the context of an import substitution strategy.

To do this we must introduce international trade into the model. This can be done most simply by assuming that any increase in exports or substitution of domestic production for imports going into consumption will result automatically in investment via import of equipment with the foreign exchange earned or saved. Allocation of investment to sectors producing for export or import substitution will then raise the rate of capital accumulation exactly as will investment in the capital goods sector, and α can refer to the proportion of

investment going to these sectors taken together.

Marginal saving depends, then, on α -- the allocation of investment to capital goods production, to production of exports, and to import substitution. But the import substitution strategy described above is strongly biased via the system of protection against both exports and the production of capital goods. And within the category of import substitution it is biased against investment in the production of material and parts. A high α must depend mainly then on (1) the rapid expansion of capacity to add value at the finishing stages of consumption goods production, and (2) the use of this capacity to reduce the import bill rather than to supply an expanding home consumption.

At first these conditions may easily be met, as import restriction serves not only as a balance of payments control, but also as the principal constraint on consumption. As domestic capacity expands rapidly in response to high rates of protection, however, two things happen. First, a kind of automatic decontrol of consumption takes place as the proportion of consumption constrained by import controls declines. This is partly due to the increased availability of goods and disappearance of scarcity premiums, and partly due to the shift in income distribution from government (customs duties) and profits of importers to income recipients in the new industries.

At the same time the expansion of consumption goods industries creates a rapidly growing demand for imports of materials, parts, and equipment. These two developments shift the focus of control over consumption to taxes and imports of inputs for the new industries. If control over the latter is tightened there arises the phenomenon of excess capacity due to scarcity of imported supplies. While this should be attributed to the misallocation of investment resulting from biases in the system of protection -- too much

capacity installed to produce consumption goods and too little to produce materials, parts, and equipment, the pressures are inevitably on the side of permitting the necessary imports. For the availability of excess capacity always promises a cheap way to get an increase in production. This, of course, also precludes the imposition of new taxes to offset the steady erosion of control over consumption. The result is what Khan has called "consumption liberalization."¹⁹

Consumption liberalization occurs, in a static context, when the rise in domestic output of consumption goods is not fully matched by a decline in imports -- i.e., in equation (7) (above), when the rise in C_d exceeds the fall in C_m , with a corresponding diminished effect on saving. In a dynamic context we must expect consumption to grow with growing output and the question whether an increase in production serves to replace imports or liberalize consumption is a more complex one.

Khan's solution²⁰ was to calculate a "normal" increase in consumption of a good based on population growth, per capita income increase, the planned marginal saving rate, and an expenditure elasticity of demand. Any increase in supply from production plus imports that was not exported or absorbed by normal consumption was defined as consumption liberalization. He then attempted to measure this over the period 1951/52 to 1959/60 for four of Pakistan's important import substitution industries: cotton cloth, sugar, cigarettes, and paper. In each case he found that a very high proportion of the output increase resulted in consumption liberalization -- from almost 50 per cent in cotton cloth to over 100 per cent in paper.²¹

¹⁹ Op. cit., p. 209.

²⁰ Ibid., pp. 208-212.

²¹ In the case of paper, imports increased more than "normal" consumption.

These results are at least consistent with the hypothesis that a part of the explanation for Pakistan's low saving rate during this period (despite rapid industrialization) was the bias toward consumption goods production for the home market.²² On the other hand, because of shifts in income distribution and in the proportions of rural and urban populations, because of the existence of controls and other abnormal influences affecting consumption, and finally because of the general complexity of the relation between the consumption of particular goods and aggregate consumption, one cannot be sure how important this was.²³

Nevertheless, on theoretical grounds a strong case can be made against an import substitution bias in development strategy because of its likely effect on saving. First, the various aspects of economic and technical inefficiency discussed above mean lower incomes, and especially lower profits, with obvious implications for saving. Second, the bias toward producing goods that can be consumed and against goods that cannot (e.g., capital goods and some exports) is likely to make political control of consumption more difficult. Finally, at some point there is an absolute necessity to move into exports or to the earlier stages of production, or both; and the longer it is postponed and the more biased against it is the system of protection, the more likely is the economy to find itself in the kind of

²²A case could be made for liberalizing the consumption of certain goods (via price or other inducements) to take advantage of economies of scale or other advantages of concentration. This "consumption distortion" has merit particularly if the favored goods are essential mass consumption goods. To avoid a general consumption rise, however, taxes would have to be raised elsewhere.

²³The use of cross-section expenditure elasticities of demand to estimate "normal" consumption may also be open to criticism. In the case at hand, however, the change in per capita income was so slight that their influence on the results was negligible.

trap that leads to consumption liberalization.

II

The conclusion I reach from this critique of an industrialization strategy biased toward import substitution is that it does not promise an easy path around the difficulties facing less developed countries. This is not a happy conclusion for the difficulties are very great and the alternatives to an import substitution strategy are not very promising either.

In any case, for what they are worth the policy implications, as they pertain to a single country, have more or less emerged in the course of the critique itself. They are, in general, to avoid the kind of excessive and distorted protection that biases growth toward a horizontal balance of consumption goods production for the domestic market, penalizing both exports and backward linkage import substitution. The costs of such a policy go beyond simple resource misallocation to adverse effects on technical efficiency, innovations and saving. More emphasis on vertical balance would seem to be essential to success in industrial growth beyond the first stage of import substitution.

This does not mean that policies should be biased against import substitution. What is needed rather are rational choices, both between import substitution and export expansion and among various potential import substitution industries. Especially important in helping the economy (public or private) to make rational choices in this area is to find some means of correcting the undervaluation of foreign exchange. Despite its obvious advantages, however, this is the kind of advice that will be widely ignored. Let me

suggest two reasons why this is so, only one to which I have any kind of an answer.

First is the terms of trade disadvantage from devaluation when world demand elasticities are significantly below infinity for important categories of exports. Insofar as these are primary commodities a particular country will normally count only a few in this category and can easily isolate them from the effects of the devaluation if international price agreements are beyond reach. If, however, new manufactured exports also face relatively low demand elasticities because of "reactive protectionism" in established manufacturing countries, there is a case for maintaining "overvalued" currencies even though this further penalizes such exports. How real is this case is difficult to estimate. Pakistan has discovered that a de facto partial devaluation by means of an export bonus scheme has elicited a very strong response from non-traditional exports. Whether what is true for one would be true for all less developed countries it is not possible for me to judge, however.

In some Latin American countries, apparently, another inhibition against devaluation is an automatic anticipation of an ensuing inflation that hastens it and renders the devaluation almost immediately ineffective.²⁴ So far as I can judge, economics is not yet able to teach us how to deal very effectively with social-psychological behavior of this sort. It might be of interest in passing to note, however, that the export bonus scheme was not generally recognized in Pakistan as a form of devaluation until economists began explaining it in these terms.

²⁴John B. Sheahan, "Imports, Investment, and Growth: Colombian Experience Since 1950," (mimeographed).

Beyond the difficulties of implementing exchange rate policy, however, lie more fundamental issues around which doubts will certainly arise. For what the above critique may appear to do is to reverse the classic argument of Nurkse in his lectures on "Patterns of Trade and Development."²⁵ There it was the difficulties faced by both traditional and new exports that dictated the option for balanced growth in relation to domestic demand. If the latter has all of the disadvantages catalogued above, however, the last escape route from economic stagnation would seem to have been closed off.

It is only fair to remind ourselves that Nurkse's view of an import substitution strategy bore little resemblance to that pictured above. He emphasized especially the prime importance of rising agricultural productivity in balanced growth and considered the inherent difficulties in carrying through an agricultural revolution to be the reason that "industrialization for domestic markets appears as a much more formidable task."²⁶

In addition he argued that: "When industrialization for the home market has taken root, it becomes easier to increase exports of manufactured goods to the more advanced economies."²⁷ It follows, I think, that he would have opposed measures that unnecessarily penalize such exports. Nevertheless, he was not sanguine about their prospects and I confess that I somewhat share his view.

The reason is that the distorted pattern of protection, described above, that tends

²⁵ Equilibrium and Growth in the World Economy, ed. by Haberler (Harvard, 1961), pp. 282-324.

²⁶ ibid., p. 315.

²⁷ ibid., p. 320.

to magnify greatly the total rate of protection for industries adding the final values to products is not a phenomenon peculiar to less developed countries. It is rather the rule for most countries.²⁸ Thus there is a strong general bias in the world against trade in finished manufactures-- the end of the production line at which less developed countries typically start.

But Nurkse suggested another escape route -- one which has by now achieved a popularity in principle far beyond its realization in practice. He wrote: "Manufacturing for home markets in the less developed countries must include also production in these countries for export to each other's markets."²⁹ This is clearly one way of resolving the Nurkse dilemma. What would appear to each individual country as new exports would represent a more rational pattern of import substitution for the group of countries. More stress on vertical balance within each country would be combined with some horizontal balance for the group. Whether this is a first-best or second-best solution to the trade problems of developing countries is a question that I won't attempt to answer in this paper.

²⁸Harry G. Johnson, "The Theory of Tariff Structure, with special reference to World Trade and Development," Etudes et Travaux de l'Institut universitaire de hautes Etudes internationales de Genève, Vol. IV, (1965), pp. 17-18.

²⁹ibid., p. 318 (Italics are Nurkse's).