17557

THE WORLD BANK ECONOMIC REVIEW

Volume 1

January 1987

Number 2

Small Manufacturing Enterprises in Developing Countries

I. M. D. Little

**A Valuation Model for Developing-Country Debt
with Endogenous Rescheduling**

Gerard Gennotte, Homi J. Kharas, and Sayeed Sadeq

The Political Economy of Economic Liberalization

Deepak Lal

**Sites and Services—and Subsidies: The Economics of Low-Cost
Housing in Developing Countries**

Stephen K. Mayo and David J. Gross

**Multimarket Analysis of Agricultural Price Policies in an Operational
Context: The Case of Cyprus**

Avishay Braverman, Jeffrey S. Hammer, and Anne Gron

Devaluation, Fiscal Deficits, and the Real Exchange Rate

Mohsin S. Khan and J. Saul Lizondo

THE WORLD BANK

ECONOMIC REVIEW

MANAGING EDITOR
Richard H. Snape

ASSISTANT EDITOR
Clara L. Else

STAFF ASSISTANT
Jane B. Sweeney

EDITORIAL BOARD
Assar Lindbeck, Institute for International Studies, Stockholm
T. N. Srinivasan, Yale University
Joseph Stiglitz, Princeton University

Deepak Lal
Sarath Rajapatirana

John Holsen
Marcelo Selowsky

The World Bank Economic Review is a professional journal for the dissemination of World Bank-sponsored research which informs policy analyses and choices. It is issued three times a year, with an occasional additional issue devoted to a special area of current research or policy interest.

The *Review* is directed to an international readership among economists and social scientists in government, business, and international agencies, as well as in universities and development research institutions. The *Review* emphasizes policy relevance and operational aspects of economics, rather than primarily theoretical and methodological issues. It is intended for readers familiar with economic theory and analysis but not necessarily proficient in advanced mathematical or econometric techniques. Articles will illustrate how professional research can shed light on policy choices without becoming essays in advocacy. Inconsistency with Bank policy will not be grounds for rejection of an article.

Articles will be drawn primarily from work conducted by World Bank staff and consultants. Before being accepted for publication by the Editorial Board, all articles are reviewed by two referees who are not members of the Bank's staff; articles must also be recommended by two out of the three external members of the Editorial Board.

The *Review* may on occasion publish articles on specified topics by non-Bank contributors. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages in length to the managing editor. Non-Bank contributors will receive a nominal fee for any article accepted for publication.

The views and interpretations in articles published are those of the authors and do not necessarily represent the views and policies of the World Bank or of its executive directors or the countries they represent. When maps are used, the designations employed are solely for the convenience of the reader and do not imply the expression of any opinion whatsoever on the part of the World Bank or its affiliates concerning the legal status of any country, territory, city, or area, or concerning the delimitations of its boundaries or national affiliation.

Comments or brief notes responding to *Review* articles are welcome and will be considered for publication to the extent that space permits. The Editorial Office welcomes inquiries and requests to translate or reprint material from the *Review*. Please direct all editorial correspondence to:

Managing Editor
The World Bank Economic Review
The World Bank
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Articles from the *Review* may be reproduced with appropriate acknowledgment of their source.

Subscriptions to *The World Bank Economic Review* for the first volumes will be made available without charge upon written request to the Publications Department. To remain on the mailing list, each year readers must reconfirm in writing their interest in renewing their subscription. See the last pages for subscription information and subscription request forms.

Inquiries concerning subscriptions should be addressed to:

World Bank Publications
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

© 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK
All rights reserved
Manufactured in the United States of America
ISSN 0258-6770

THE WORLD BANK ECONOMIC REVIEW

Volume 1

January 1987

Number 2

Small Manufacturing Enterprises in Developing Countries <i>I. M. D. Little</i>	203
A Valuation Model for Developing-Country Debt with Endogenous Rescheduling <i>Gerard Gennotte, Homi J. Kharas, and Sayeed Sadeq</i>	237
The Political Economy of Economic Liberalization <i>Deepak Lal</i>	273
Sites and Services—and Subsidies: The Economics of Low-Cost Housing in Developing Countries <i>Stephen K. Mayo and David J. Gross</i>	301
Multimarket Analysis of Agricultural Price Policies in an Operational Context: The Case of Cyprus <i>Avishay Braverman, Jeffrey S. Hammer, and Anne Gron</i>	337
Devaluation, Fiscal Deficits, and the Real Exchange Rate <i>Mohsin S. Khan and J. Saul Lizondo</i>	357

Small Manufacturing Enterprises in Developing Countries

I. M. D. Little

Emphasis is often placed on the promotion of small enterprises in developing countries, particularly as a means of improving the lot of unskilled workers. This focus raises questions about the relationship between establishment size and the pattern and efficiency of factor use, and about the nature and effects of price differentials in factor markets. This article goes some way toward answering these questions with data from surveys of small manufacturing enterprises in India and Colombia sponsored by the World Bank and relevant material from other countries. The article also examines India's long-standing policy, unusual among developing countries, of providing special support and protection for small enterprises. Analyses based on disaggregated data found that small firms are not reliably more labor-intensive than their larger counterparts; nor are they consistently more technically efficient in their use of resources. In light of these findings and an analysis of factor markets, this article discusses the general implications of the research results for industrial policy in developing countries.

This article is about the relationship between the size distribution of manufacturing industry in developing countries and the nature and efficiency of production processes. It does not purport to be a comprehensive survey of analyses of this relationship. Rather, it discusses the findings of World Bank surveys conducted in India, and assesses their broad implications with additional information provided by associated studies in other developing countries and relevant experience in developed countries.

Concern here is not with the concentration of economic power, and hence

The author, a consultant to the World Bank, is at the University of Oxford. This article is based largely on the book *Small Manufacturing Enterprises: A Comparative Study of India and Other Countries* by I. M. D. Little, Dipak Mazumdar, and John M. Page (forthcoming). The analysis and results reported on here are predominantly the work of Mazumdar and Page, although they are not responsible for the interpretations and conclusions reached. These are the author's; and in addition to the specific research results reported here, they draw on direct observations over many years of economic policies and development in developing countries. Other major outputs based on the same research are Anderson and Khambata (1981); Anderson (1982a, 1982b); Cortes, Berry, and Ishaq (forthcoming); Ho (1980); Mazumdar (1984a, 1984b); and Page (1979, 1984). References given in the present article are mainly to Little, Mazumdar, and Page (forthcoming) and occasionally to Cortes, Berry, and Ishaq (forthcoming). These works must be consulted for original sources.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

with the size distribution of *firms* and patterns of ownership, but with the size distribution of plants or establishments. However, the terms “enterprise” and “firm” are also often used; in the vast majority of cases small establishments in developing countries are independent firms.

The central issues discussed here are whether smaller enterprises use resources more efficiently than larger enterprises, and if so, for what size ranges this is true, and what might constitute an appropriate approach to industrial policy. This amounts to a concern for the best use of a country’s productive resources, including labor. Any inquiry into the efficiency of resource use is an exercise in applied welfare economics. It should not need to be stressed that this framework of analysis and the concept of efficiency of resource use are not in conflict with concern for the poor.

The best use of a country’s resources involves more than the efficient allocation of a given stock of capital and labor. Savings and the training of workers and entrepreneurs may influence and be influenced by the size distribution of industrial enterprises. These dynamic relations were not the central concern of the research, but where some light was thrown on them or more generally on the determinants of growth or efficiency, this is mentioned.

India has had by far the most varied and forceful set of industrial policies aimed at helping small-scale enterprises of any developing country. These policies are described, and their results assessed, below.

What do we mean by small? Size is most commonly measured by the number of workers. In industrial countries small-scale often means less than 200–300 workers. In developing countries, however, the average plant size is smaller, and “small” will generally be taken to mean 1–49 workers. Sometimes much finer distinctions are necessary, both up and down the scale, and will be introduced as required. Size may also be measured by capital employed, and for some purposes this may be a more appropriate measure. It is used for policy purposes in India, but even there size distributions by capital employed are rarely available. Both the availability of statistics and the fact that employment size is more easily envisaged make it the dominant measure.

I. INCREASING THE DEMAND FOR UNSKILLED LABOR

The central issue examined is whether the promotion of small-scale enterprises (SSEs) in developing countries is a good means of reducing poverty and inequality, insofar as it would result in an increased demand for unskilled labor. This was not a major concern of development policy until the late 1960s, when international institutions and some private development propagandists claimed to discover massive and rising unemployment in developing countries. The International Labour Organisation’s (ILO’s) World Employment Program was then launched. The concern for employment was soon linked to advocacy of SSEs, which were believed to offer much more employment than larger firms for any given investment.

As a part of this trend, the ILO's Kenya mission took up and popularized the concept of the "informal" sector. Discussion of how it should be defined, and of its role, has filled thousands of pages. But whatever exactly it was, it was accepted that it was labor-intensive, that it provided goods and services by the poor for the poor, and that it should be encouraged rather than being regarded as an unproductive slough that development would and should eliminate. However, I have not found the informal sector to be a concept that is of either analytic or operational value, and I do not use it.

Apart from the informal sector, there was much conceptual and analytic confusion as it came to be recognized that unemployment was not closely related to inequality or poverty, and that the evidence for any general increase in unemployment was weak and even false. One need not follow the changes in emphasis in international forums from unemployment to inequality to poverty. It suffices to say that a consensus has formed within the international development establishment in favor of a pattern of development that is more demanding of unskilled labor, and that the case for this does not depend on evidence of either unemployment or underemployment. Essentially the case made is that in developing countries a minority of people work with a lot of capital (or land) while a large majority work with very little indeed and can thus earn very little. If new investment called for more participating labor, more work would be done, real wages would rise, or both would occur. In any case, unskilled (or initially unskilled) workers would benefit. This is a consensus with which I am in full agreement.

Thus the main argument for encouraging small-scale enterprise development as a means of reducing poverty and inequality rests on the increased demand for unskilled labor that is supposed to follow.¹ It is clear, however, that output and, more specifically, the productivity of unskilled labor must also be considered. If small enterprises used both more capital and more labor per unit of output than larger enterprises, then investment in small enterprises would result in a smaller increase in output than investment in large enterprises and there would be no clear case for special investment in sses. Similarly, it is not sufficient to show that sses use more labor and no more capital per unit of output than do larger enterprises. This being so, one could as well employ more workers to do nothing in large factories (as is indeed often done) as employ them "productively" in sses. Labor almost always and everywhere has some opportunity cost—that is, if not employed wherever it is, it would somehow earn some money and produce something elsewhere. There is thus no escaping the fact that the case for promoting any particular type of enterprise is that it uses factors more efficiently, given their social costs. The so-called employment case *for* sses must rest on this. There will certainly be a strong case *against* sses if they are shown to use both more capital and more labor per unit of output: if such is the case, then one

1. It is also argued that sses produce goods that are more appropriate for the poor than do larger enterprises. The significance of this argument is recognized, but it is not discussed in depth in this article.

could only advocate the promotion of sses in order to increase the demand for labor if there were no way of increasing labor demand without reducing output—which is in general false.

Hence, to establish a case for special encouragement of sses, it must be claimed that (1) smaller enterprises are both relatively labor-intensive (low capital-labor ratio, K/L) and have relatively high capital productivity (output-capital ratio, O/K), low labor productivity (output-labor ratio, O/L) being implied; and (2) that these characteristics represent a more efficient use of capital and labor. If production functions are homothetic and profits are maximized,² then the first claim implies that small firms face a lower price of labor relative to capital than larger firms. Further, it implies that these resources are more efficiently employed if this lower relative price of labor is closer to its relative shadow price than is that faced by larger enterprises. This assumption implies distortions in capital or labor markets, or in both.

The above paragraph raises questions that are discussed in the next four sections. First, what are the facts concerning relative factor proportions and productivities? Second, if small and large enterprises face different wages should this be regarded as a distortion? Third, are capital costs similarly distorted? Fourth, are the facts about factor proportions and productivities consistent with differences in factor prices? Before turning to these questions I conclude the present section by pointing out that any case for reform to support sses will be strengthened if it can be shown that government policies, whether macroeconomic or industrial, have discriminated against small enterprises.

With the exception of India (the policies of which are examined in a later section), developing-country governments have shown little real (as opposed to cosmetic) concern for sses. Industrialization has been promoted in three main ways: first, by regulating trade; second, by legislating investment incentives; and third, by directly undertaking public sector investments, often with foreign aid or commercial loans. All of these policies have discriminated against sses. Foreign trade and payment controls have discriminated, since large firms are better placed to obtain import permits for capital equipment, components, and raw materials and are also better able to obtain the tariff rebates that are intended to alleviate some of the harmful effects of high protection. Investment incentive laws have sometimes actively discriminated against the small, by restricting their tax concessions to firms of some minimum, generally large, size. Where there has been no such overt discrimination, small firms have nevertheless often been ignorant of the concessions obtainable or unable to handle the paperwork needed to make an effective case to the public office concerned.

Selective credit controls combined with low, controlled real interest rates have the consequence that banks have not been able to compensate for the higher cost

2. In homothetic production functions, if relative factor prices are the same for large and small firms, then relative factor proportions will also be the same. Evidence presented in section II of this article suggests that production functions are homothetic for at least some of the industries surveyed. The possibility of differences in firms' technical efficiency should also be recognized (see section VI).

of small loans by charging more (higher collateral requirements not necessarily being a feasible alternative). The limited credit has therefore been allocated to large clients. Finally, when governments have directly promoted industrial investments, they have favored the large project even when it would be owned and operated by a private enterprise, whether foreign, local, or a joint venture.

Bias can also arise because the problems of dealing with the tax authorities and with governmental regulations weigh more heavily upon small firms, even in the absence of trade controls and industrial promotion. However, the hand of the law and the tax collector may often fail to reach the very small (say, less than ten workers) in most developing countries. When this is so, nonenforcement of taxes and of some regulations such as those governing minimum wages or working conditions will constitute a bias *in favor* of the very small (though this in turn may sometimes be outweighed by harassment of the small under zoning regulations in urban areas). Thus, taxes and regulations are likely to weigh more heavily on the small modern factory employing ten to fifty workers than on either the large factory or household/workshop activities.

Government concern for sses should be reflected in measures to offset the disadvantages imposed upon them, even if this does not involve major liberalizations of trade and industrialization regimes so as to remove these biases "at source." Indeed, in most countries, some such measures have been implemented. Some credit institution has usually been set up to lend to small and medium-size enterprises (the medium-size being favored in practice for the reasons already given). Industrial extension schemes, management training institutes, and industrial estates (the latter not always mainly for the small) have been started in many countries, often supported by the United Nations Development Programme, ILO, or some bilateral aid agency. Even where these have had some success, however, their reach has invariably been very limited, if only because of the difficulty of communicating their existence to all possible recipients. This applies also to the provision of institutional credit. Small firms (less than fifty employees) get very little of the credit reserved for the medium-size and small, and the very small (less than ten) almost never get any medium- or long-term loans. Except in the case of India, where some exceptional measures protect traditional industries, it is safe to conclude that the measures targeted to favor small enterprises have negligible positive impact relative to the negative impact of the general economic policies of governments.

However, the question remains open as to whether a bias against the small also amounts to a bias in favor of capital-intensive development. The contrasting experiences of India on the one hand, and the Republic of Korea and Taiwan on the other, suggest the opposite. India has done much to preserve and encourage small enterprises, and its industrial development has been very capital-intensive. Korea and Taiwan have done very little for the small, and tiny manufacturing enterprises have almost vanished. Medium-size and large enterprises accounted for the bulk of the extremely rapid rise in employment in these economies. For instance in Korea between 1963 and 1975 there was a rise in manufac-

turing employment of over 1 million persons, 86 percent of which was accounted for by firms with over 100 employees. The smallest census size group, 5–9 workers, showed no rise (Ho 1980). Moreover, large-scale industrialization was compatible with capital-output ratios that were among the lowest achieved by any economy.

To summarize, the *prima facie* case for policy interventions in favor of SSEs as a means of raising overall welfare in developing economies must rest on evidence that small units on average use factor inputs more productively than their larger counterparts, so that a shift of resources in favor of smaller units would yield a net increase in output, as well as an increase in the demand for unskilled labor. This is not to deny that income distribution is an important consideration; indeed, an increase in the demand for unskilled labor is probably one of the most effective means of improving distribution. However, this topic and the related claim that small enterprises produce goods more appropriate for the poor than do large enterprises are not considered in this article.³

II. FIRM SIZE, FACTOR PROPORTIONS, AND PRODUCTIVITY—SOME FINDINGS

I now address the first of the questions posed in the previous section: how do factor proportions and productivities actually vary with size? I first consider manufacturing in the aggregate and summarize evidence from industrial censuses for Japan and eleven developing countries.⁴ With only a few anomalies, capital intensity and labor productivity rise with size. But capital productivity quite often behaves unexpectedly. In Japan, it rises from the smallest size class (1–9 workers) to peak in the range of 20–49 workers (Ohkawa and Tajima 1976). In many developing countries, the figures behave erratically, but there is some evidence of a similar peak coming in the range 20–200 workers. For instance, in Korea capital productivity is virtually constant for 5–49 workers, rises dramatically in the range 50–199, and then falls again, though not to the level found in small enterprises (Ho 1980).

However, such aggregated figures, even if reliable, are of limited interest. For policy purposes, it is important to know whether small units survive and are labor-intensive because they produce the great bulk of a limited range of goods that are technologically unsuited to large-scale, capital-intensive production or whether a very wide range of goods is produced in both small, labor-intensive and large, capital-intensive units. In the former case, significantly more small-scale production could be engineered only by a change in demand in favor of the sorts of things that can be efficiently made by small units. In the latter case, a large relative shift to small units may be both possible and desirable without

3. They are dealt with, though not systematically, in Little, Mazumdar, and Page (forthcoming), especially in chap. 13.

4. For more detail and sources see Little, Mazumdar, and Page (forthcoming), chap. 7. The eleven developing countries are Brazil, Chile, Colombia, India, Malaysia, Mexico, Pakistan, Paraguay, the Philippines, Sri Lanka, and Venezuela.

such a change in demand. In narrowly defined industries, it is likely that the products of large and small firms are or could be highly competitive, though this is by no means always the case.

For this reason, more disaggregated figures obtained from censuses and surveys in India, Korea, the Philippines, and Taiwan were examined; industry classifications ranged from the two-digit to the five-digit standard industrial classification (sic) level. The greater the disaggregation, the less frequently were smaller enterprises found to be more labor-intensive, and to have lower labor and higher capital productivities (Little, Mazumdar, and Page, forthcoming, chap. 7).

In India, disaggregation to the level of 32 industries produced evidence that factories in the range 10–49 workers (10–99 if no power was used) were more labor-intensive than larger ones in a high proportion of cases (Mehta 1969). But analysis at the much greater level of disaggregation possible in Korea (373 industries in the 1975 survey) showed that the incidence of highest labor intensity is almost even throughout the range of size classes from 5–499 workers. Moreover, highest *capital* intensity was as often found in the range 5–50 workers as in the range 50 plus (Ho 1980).⁵ The moral is obvious. Even if one were constructing a policy purely to favor labor intensity, it would be essential to discriminate finely between industries. However, no such policy ignoring output could be sensible.

The crudest possible scale which might yet have claim to measure efficiency in a labor-abundant economy is capital productivity. The results here are not favorable to either the very small or the very large, even for manufacturing as a whole. In all the economies examined, capital productivity for manufacturing (taken in the aggregate) peaks in the range 20–200 workers. The highly disaggregated Korean figures produce the results that capital productivity peaks within the range 50–500 workers in two-thirds of the industries.⁶ Further analysis of the figures shows that capital intensity was lower and capital productivity higher in the range 100–99 workers as compared with the range 20–49 workers in about half of the 139 industries for which this comparison could be made. A Reserve Bank of India survey (1979) of small industries (defined by capital) suggested that enterprises employing over 20 workers (the “large-small”) have higher capital productivity than the “small-small.”

In Korea, in many industries, smaller size groups were “dominated” by their larger counterparts, that is, they used both more labor and more capital per unit of output. Where this is not the case, it is useful to shadow price the inputs. (Shadow pricing is also desirable because of variation in labor skills.) Using actual wages as shadows⁷ and a uniform 20 percent as the opportunity cost of

5. All the Korean figures derive from the census of 1968.

6. The relatively low capital productivity of large firms, especially the very large (say, over 2,000 workers), that is often observed may have little to do with size as such. The large size classes are likely to include a relatively high proportion of firms whose prices are controlled, and a high proportion of public sector firms.

7. It is widely agreed that the labor market in Korea is such that wages closely map opportunity costs.

capital, it was found that in only 32 of 138 cases was social cost lowest per unit of value added for small enterprises (less than 50 workers); only 7 of these cases were in the smallest size class (5–9 workers).⁸

The evidence thus far reviewed is very unfavorable for the SSE case.⁹ It is, however, to some extent misleading. The smaller size categories include some cases in which capital intensity is abnormally high and productivity low. Thus a large firm measured by normal employment will fall into a smaller size category in a bad year, and a sick or dying firm may have only a skeleton staff. The smaller sizes also include start-ups that have installed most of their equipment but have not yet reached planned employment levels. There is some evidence that, when size is measured by capital, the factor proportions and productivities behave in a more orderly manner, and that the small thus defined would show up better, but census data are not normally available on this basis.

Partly for the above reasons, and out of some general distrust of census figures, a number of narrowly defined industries were intensively surveyed in India under the World Bank research program. These yielded further and less biased evidence concerning the variability of factor proportions and productivities with size than that provided by censuses. The very intensive inquiries made of the enterprises should also have resulted in these survey figures being more reliable than those of censuses. The industries studied included shoes, soap, printing, machine tools, and metal casting. In general it was found that the intragroup variations were so large that mean differences between groups were rarely statistically significant.

Survey findings for establishments ranked by the number of employees are shown in table 1. Some comments can be made on the results for particular industries. In printing, the relatively low capital productivity of the two largest size categories, which is not offset by relatively high labor productivity, suggests that these establishments are technically inefficient relative to their more labor-intensive, smaller counterparts. Machine tools was the only industry in which a statistically significant difference was found between the factor productivities of very small firms (less than 10 workers) and those of larger firms. The soap sample was restricted to firms using no power, and the associated uniformity of technique accounts for the similarity of capital intensities and factor productivities across size categories. Similarly, in shoes technology is probably a better guide to capital intensity and factor productivities than is size, with the marked difference in findings for establishments with more than 100 workers reflecting the use of factory rather than traditional handicraft methods. In metal casting, the size-groups with over 50 workers had both higher capital and higher labor productivity, but the intragroup variation was too large for this to be statistically significant.

8. For this comparison the number of industries was reduced by excluding those with entries in less than four size classes, as well as those for which considerable heterogeneity of output was suspected.

9. This article has reported mainly Korean figures. Other broadly confirmatory evidence is given in Little, Mazumdar, and Page (forthcoming, chap. 7).

Table 1. *Factor-Output and Capital-Labor Ratios in Firms in Five Indian Industries, by Number of Employees*

Size (<i>N</i> = number of workers)	Printing			Machine tools			Soap			Shoes			Metal casting		
	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L
$N < 5$	8.92	4.86	1.84	—	—	—	10.80	2.47	4.37	5.23	14.28	0.37	3.20	10.67	0.30
$5 \leq N < 10$	12.15	3.51	3.46	4.16	1.11	3.75	9.33	2.20	4.24	5.61	29.87	0.19	13.91	4.97	2.80
$10 \leq N < 25$	14.18	3.26	4.35	6.73	2.66	2.53	9.78	2.04	4.79	5.67	16.20	0.35	12.32	5.77	2.14
$25 \leq N < 50$	14.56	3.51	4.15	7.84	3.28	2.39	11.19	1.98	5.65	6.30	13.13	0.48	14.44	7.63	1.89
$50 \leq N < 100$	11.59	1.24	9.35	7.44	2.57	2.90	10.92	2.93	3.73	—	—	—	19.40	29.96	0.65
$N \geq 100$	13.81	1.21	11.41	12.31	2.13	5.78	11.10	2.74	4.05	15.32	4.93	3.11	22.47	8.59	2.62

— Not available.

Note: Y/L and K/L are in thousands of rupees per worker. Y/K is a pure number.

Source: Little, Mazumdar, and Page (forthcoming, table 11-1).

Similarly, high intragroup variation was found in two surveys of small and medium-size firms in the metalworking and food processing industries conducted in Colombia (see Cortes, Berry, and Ishaq, forthcoming). In metalworking, capital intensity was lowest and capital productivity highest in the range 41–60 workers, and labor productivity was also higher than in the small size groups. In food processing, capital intensity was lowest and capital productivity highest in the size group 16–29 workers. Labor productivity was lower in the 16–29 worker group than in the 8–15 and 30 plus groups, but higher than in the smallest size class (1–7 workers). The authors of these surveys also estimated a social cost-benefit ratio, which was best in the largest size classes (61 plus workers in metalworking and 30 plus in food processing) (Cortes, Berry, and Ishaq, forthcoming, chap. 3).

When establishments are ranked by capital (measured as the undiscounted cost of fixed capital at 1979–80 prices), the factor proportions and productivities vary with size in a manner that is closer to expectations than when employment is used as the size discriminator. In the two Colombian surveys, the ratios behaved exactly as expected: capital intensity and labor productivity rose with size, and capital productivity fell. This confirms what has been found to be true in some other surveys or censuses, for example, in Japan (Kaneda 1980). However, in the Indian surveys (see table 2) substantial irregularities occurred, and the differences between size groups were frequently insignificant, intragroup variation remaining high. Shoes was the only industry except for the largest size class, in which there was a regular rise of capital intensity and labor productivity with capital size and a regular fall in capital productivity.

The surveys also distinguished between skilled and unskilled workers. In the cases of machine tools, shoes, and metal casting, there was a distinct tendency for the proportion of unskilled workers to rise as the size of firms rose above twenty-five workers. In soap there was little variation, while in printing the results were highly erratic (probably reflecting difficulties with the classification of workers).

As between the industries, printing was the most capital-intensive. It may be noted that for these five industries, intraindustry differences in capital intensity were of the same order of magnitude as interindustry differences. However, this similarity certainly does not carry over to the whole range of industries, as witnessed by Samuel Ho's examination of the Korean and Taiwanese censuses. In Korea, some industries were more than a hundred times as capital-intensive as others, while within industries the difference between size classes seldom exceeded three times.

This is an important observation. It suggests that, to achieve a more labor-intensive manufacturing industry, there is far more mileage to be obtained by changing the industry mix by altering the pattern of demand than by influencing the size distribution within industries. It should be noted, however, that this argument rests on observation of *average* intensities and productivities. *Marginal* values will clearly be important in determining optimal policies.

Table 2. *Factor-Output and Capital-Labor Ratios in Firms in Five Indian Industries, by Value of Fixed Assets*

Size (K = thousands of rupees)	Printing			Machine tools			Soap			Shoes			Metal casting		
	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L	Y/L	Y/K	K/L
$K < 1$	—	—	—	—	—	—	—	—	—	3.77	42.30	0.09	—	—	—
$1 \leq K < 5$	6.08	4.81	1.26	—	—	—	—	—	—	4.76	18.41	0.26	3.20	10.67	0.30
$5 \leq K < 10$	8.38	5.40	1.55	—	—	—	9.86	4.20	2.35	6.13	9.72	0.63	—	—	—
$10 \leq K < 20$	10.83	5.77	1.88	—	—	—	9.11	3.82	2.39	7.37	9.42	0.78	5.84	6.84	0.85
$20 \leq K < 50$	13.02	3.98	3.27	6.25	3.17	1.97	8.39	2.35	3.15	10.84	8.64	1.26	12.74	10.72	1.19
$50 \leq K < 100$	14.40	3.52	4.09	7.38	3.25	2.27	10.06	1.41	7.14	10.88	6.26	1.74	23.06	8.91	2.59
$100 \leq K < 200$	13.00	1.44	9.03	6.45	2.08	3.10	12.08	1.61	7.50	—	—	—	12.12	6.45	1.88
$200 \leq K < 500$	17.24	1.57	10.98	8.41	1.79	4.70	10.79	2.75	3.92	—	—	—	27.82	8.44	3.30
$500 \leq K < 1,000$	15.13	1.42	10.66	10.63	2.22	4.79	13.03	2.41	5.41	14.45	3.35	4.31	21.26	7.83	2.72
$K \geq 1,000$	12.71	1.09	11.66	15.91	1.80	8.84	—	—	—	15.98	6.11	2.62	25.96	2.72	9.54

— Not available.

Note: Y/L and K/L are in thousands of rupees per worker. Y/K is a pure number.

Source: Little, Mazumdar, and Page (forthcoming, table 11-2).

The Indian surveys provided data at the firm level that permitted analysis of variations in inputs and outputs going well beyond the simple relationships of size, factor proportions, and productivities that have thus far been described.¹⁰ The analysis was based on the fitting of three factor translog production functions of the average and frontier type. The three factors were capital and skilled and unskilled labor.¹¹ It is convenient to record some of the more interesting econometric findings at this point.

Technical efficiency was investigated by measuring the ratio of the factor inputs of a hypothetical best practice or "frontier" firm to those of the actual firm (both having the same factor proportions). There was no significant relationship between firm size and efficiency except in the case of machine tools, in which it was positive. This was true whether or not one controlled for various other potential determinants of technical efficiency (these other determinants are discussed in section VI).

It has often been argued that the relative prices of factors of production have little bearing on the methods of production employed, and hence that they are largely irrelevant so far as employment is concerned. Even if this were not so, however, the substitutability of factors would be an important issue. In all four industries for which estimates could be made,¹² the elasticity of substitution was highest for the ratio of skilled to unskilled labor (averaging 3.1, taken at the mean size), next highest for the ratio of skilled labor to capital (averaging 2.0), and lowest for the ratio of unskilled labor to capital (0.7). There was no significant variation of substitutability with size.

Tests were made for homotheticity and constant returns to scale. With a homothetic production function, the scale of operations will not affect the capital-labor ratio, provided relative factor prices remain the same. In the case of machine tools, metal casting, and soap, the hypotheses of homotheticity and constant returns to scale could not be rejected. In printing and shoes, size tended to favor capital intensity, while returns to scale were variable. In no case was there strong evidence of increasing returns.

In summary, enough evidence has been presented to show that size, especially when measured by the number of employees, is a very poor indicator of those characteristics of firms that may be of interest to policymakers. If one were searching solely for labor intensity, then it would be better to study techniques than size. Of course, traditional or handicraft techniques are highly labor-intensive, and units are generally (though not always) very small. But very small modern factories or workshops may be very capital-intensive. This was pointed out long ago by Dhar and Lydall (1961).

10. Surveys were also made for powerlooms and handlooms, but the data generated did not permit interesting size comparisons within these weaving methods because of insufficient variation in size of enterprise, insufficient variation in capital-labor ratios, or both.

11. The analysis is discussed in depth in Little, Mazumdar, and Page (forthcoming, chaps. 9–12).

12. Soap had to be excluded because the unrestricted translog function was insufficiently well-behaved to permit reliable estimates of the elasticity of substitution.

Small size certainly does not indicate high capital productivity. Indeed, the figures suggest that small firms rather often have both low capital and low labor productivity. Capital productivity and total factor productivity peak in the medium size range of 50–500 workers in most industries. However, census and survey returns based on annual figures tend to bias results against the small. Where firms can be ranked by capital employed, the results tend to be rather more as expected (capital intensity and labor productivity rising, and capital productivity falling monotonically, with size), but nevertheless usually with such variance within size groups as to make intergroup comparisons statistically insignificant. It should, moreover, be noted that when firms are grouped by capital size, there is a bias *in favor* of the small. This is because small groups, at least in censuses, will include firms with highly depreciated capital.

Our own enterprise surveys were almost free of the biases referred to. Sick firms with skeleton staffs were not observed. At least a year's operation was required, and there was no surprising preponderance of very young firms. The *undepreciated* value of capital was used in the Indian surveys, and commercial value in the Colombian surveys. These surveys did not provide evidence that small firms employ resources more efficiently (either technically or from a social point of view) than large firms, nor even that they are reliably more labor-intensive. If our research can be held to suggest anything about size and economic or social desirability, it is that beauty is to be found mostly in the middle of the size distribution.

III. LABOR MARKETS

One of the arguments in support of small-scale enterprises has been that they pay low wages that are closer to the social cost of labor than the higher wages paid by large enterprises. While doubt has already been cast on whether small enterprises will in fact be more labor-intensive—as theory suggests they should be if paying lower wages—this argument is worth examining in its own right.

Before proceeding it must be noted that the social cost of skilled workers is greater than that of the unskilled. There is no *prima facie* distortion unless labor of similar skills is paid differently for similar work in enterprises of different sizes. The crudest control is to restrict comparisons when possible to those classified as either skilled or unskilled. Further to this, allowance can sometimes be made for differences in age, sex, job experience, and education. In the following, differences in skills that have been controlled for when making comparisons will be noted.

For manufacturing *in the aggregate*, there is no doubt that small enterprises (say, with less than 50 workers) pay lower wages than large enterprises (say, with more than 500 workers), and that the differentials are too great to be plausibly attributed to differences in skill or capability. These differentials vary by country and are smaller the more developed the country. In the countries for which evidence was readily available, they were greatest in India and Indonesia (on the

order of 100 percent for similar workers); much lower in Korea, Malaysia, and Colombia (30–50 percent); and still less in the industrial countries (25–30 percent).¹³ There are “natural” (that is, noninstitutional) reasons for differentials which apply more strongly in developing countries and therefore explain country differences, at least in part. In general, differentials are not distortions to the extent that they are explained by these noninstitutional reasons.

The most important natural reason is that a stable labor force that can acquire firm-specific and industry-specific skills is relatively important for large firms, and for industries in which large firms predominate. One reason for this is that large firms are, in the aggregate, relatively capital-intensive.¹⁴ Machinery and equipment would be wasted and even damaged if associated with an unstable labor force, and there is evidence that labor turnover is higher in small firms. Another reason is that large firms usually produce higher-quality products, and that the scale of their marketing requires the goodwill associated with branded products of uniform quality. A relatively skilled and reliable labor force is needed if these desiderata are to be achieved.

Relatively high wages must be paid to get the reliable labor force which large firms want. This is particularly the case in the poorest countries. Mazumdar (1984b) has argued that large wage differentials existed in the Bombay cotton mills before there was any institutional reason for this, such as the strength of trade unions or governmental influence. The mills needed to attract permanent migrants, who had a higher supply price than transient workers. Small units, especially in handicrafts, services, and construction, had a comparative advantage in the use of cheap, transient labor; in countries such as India and Indonesia there is still a large supply of transient migrants, which keeps wages very low and turnover high in such activities. In countries like Colombia, Korea, and Malaysia, the urban labor force is more settled, and it is relevant that differentials have narrowed considerably in Korea with the very rapid growth of demand for urban labor (whereas they appear to be widening in India).

In the developed countries, labor forces are generally more skilled and homogeneous. But even where transient migration is a thing of the past, and the labor force is literate and manually adept, large firms demanding firm-specific skills may still need to pay relatively high wages to achieve the stability they require. Finally, it may be that even a stable labor force will increase its productivity if paid more, in which case the lowest wage that is required for stability may not be the least-cost wage (the so-called efficiency wage argument).

Indeed, there are so many natural economic reasons for wage differentials,

13. The general survey of this evidence, in Little, Mazumdar, and Page (forthcoming, chap. 4), is derived from Deshpande (1979), India (1979), *Indian Labor Journal* (1981), and Mazumdar (1984b), for India; Manning (1979) and World Bank (1979) for Indonesia; Park (1981) for Korea; Mazumdar (1981) for Malaysia; Cortes, Berry, and Ishaq (forthcoming); and Mohan (1981) for Colombia; and Lester (1967) and Wilkinson (1973) for developed countries.

14. It is hard to separate the influence of size from that of capital intensity. Small, capital-intensive firms are also likely to pay relatively high wages.

Table 3. *Average Annual Wage of Unskilled Workers in Firms in Five Indian Industries*
(rupees)

Size (N = number of workers)	Printing	Machine tools	Soap	Shoes	Powerlooms
$N \leq 10$	2,364	2,172	2,664	—	—
$10 < N \leq 25$	2,436	2,316	2,832	2,256	1,140
$25 < N \leq 50$	2,604	2,700	2,964	3,000	1,428
$N > 50$	2,292	2,544	3,240	—	2,244

Source: Little, Mazumdar, and Page (forthcoming, table 14-3).

including the relative disutility of jobs, that it is only if there is evidence that they are produced or widened by trade unions or by governmental pressure or legislation that there is a *prima facie* case for saying that higher-paid workers are, from a social point of view, paid relatively too much. In India there is some evidence that government intervention helped to create a labor aristocracy in the textile industry in the interwar period, although quite large wage differentials clearly existed long before. There is also evidence that labor legislation and the awards of industrial tribunals have widened differentials, especially for large firms, in the postwar period (Subramanian 1979). In Colombia, as in other Latin American countries, the proportion of fringe benefits rises sharply with size, and this is probably a distortion that results from legislation (Cortes, Berry, and Ishaq, forthcoming). In the other countries mentioned there seems to be little or no institutional influence operating. An exception may be that foreign firms usually pay relatively high wages, and it is difficult to attribute this to anything other than government influence or political prudence.

Although wages rise with size, they do not always do so in a smooth fashion. There is some evidence of a wage cliff at around 200 workers. For instance, in India the Annual Survey of Industry (1978–79) shows a wage jump at 200 workers, and Mazumdar (1984b) found some evidence of a cliff in his Bombay labor survey (after controlling for the quality of labor). In Colombia also, evidence was found of a cliff at 200 workers (Cortes, Berry, and Ishaq, forthcoming). Such a cliff could be an indication of institutional forces coming into play, a suggestion supported by the fact that the Korean data show a rather steady progression of wages with size (see Ho 1980, sec. V). However, the evidence is weak.

So far, I have been considering industry or manufacturing as a whole. The only disaggregated data available come from the research surveys. As can be seen from table 3, only in powerlooms was there a very large differential, with firms with over 50 workers paying about double those with 11–25 workers. Very large differentials seem to be a feature of the Indian textile industry.

In the two World Bank surveys of metalworking and food processing firms in Colombia, earnings of unskilled workers, including fringe benefits, were 30–50

percent higher in the largest than in the smallest enterprises (Cortes, Berry, and Ishaq, forthcoming, chap. 2).

These surveys were essentially of small and medium-size firms with up to 200 employees, thus excluding those beyond the wage cliff—if there is a wage cliff. The wage differentials are certainly far lower than for manufacturing as a whole when large firms are included, but it is not clear whether this is the result of disaggregation or of the exclusion of large, especially very large, firms. Finally, although the data above relate to unskilled workers, the arbitrary nature of this classification means that there may still be skill differences which account for part of the differences in earnings.

Nothing very firm can be concluded about distortions in the labor market from the evidence presented. A positive relationship has been found between wages and firm size. There is clear evidence of institutional factors at work in India, and probably also in Colombia, that make the wages paid in small and medium-size firms closer to an optimal social wage than those in large and very large firms in these countries. It seems likely that institutional factors are less significant for firms with up to about 200 workers, and this accords with the evidence of smaller differentials in this range. As has been pointed out, such differentials might be expected to arise quite naturally from differences in firms' demand functions for labor. It is arguably the case that in most industries and countries there may be little labor market distortion as between small and medium-size enterprises.

IV. CAPITAL MARKETS

Are capital markets distorted in a way which penalizes small enterprises? Although our surveys produced some information about the extent and sources of borrowing, they yielded no reliable information on costs. Consequently our response to this question is more speculative and less firmly based than one might wish.

To establish the existence of distortions, it is not sufficient to show that loans to small enterprises bear high rates of interest (except, perhaps, for loans from friends and relatives), or that they are unobtainable. The reason is that the costs of lending to small borrowers are high, often prohibitively high. Capital markets cannot come near to satisfying the conditions of perfect competition. Even when lenders are highly competitive, any loan involves unique elements. The lender has to assess the probability of repayment, which ideally requires intimate knowledge of the borrower and of the project for which the money will be used. Different lenders will assess the probabilities differently, as well as having different attitudes toward risk.

The riskiness of a portfolio of loans may be reduced by investigation of borrowers and projects. In principle, there is an optimum amount to be spent on such character and project analysis. This obviously approaches zero with very small loans, which is why such lending is largely confined to friends and rela-

tives. Collateral is another way of reducing the risk of nonrepayment and hence increasing the present value of a loan. This is less costly to arrange than investigation, and it is therefore not surprising that bankers usually demand collateral, but many would-be borrowers are unable or unwilling to provide it. Finally, higher interest charges will only raise the expected value of a loan up to a point, with higher interest rates reducing the ability, and perhaps the willingness, of a borrower to repay. Moreover, high interest rates may increase the portfolio proportion of dishonest or overoptimistic borrowers, and of risky projects (adverse selection). The upshot is that there is probably some maximum rate of interest beyond which rate increases would reduce the expected present value of the lender's portfolio. If such an optimum interest rate does not cover expected costs, then there will be no market at all.

Even in less extreme circumstances, some good projects will not be realized because their proposers are wrongly deemed not to be creditworthy, no chargeable rate of interest being judged capable of compensating for the risks. But such "imperfections" in intermediation are inevitable, since the future is uncertain and people are unreliable. The existence of imperfection does not imply that improvement is possible, though of course it may be.

In considering the case for policy, it becomes important to understand the reasons for imperfection. It has often in the past been attributed to a lack of formal lending institutions, to bankers' pessimism or excessive risk aversion, and to monopoly, rather than to the more fundamental reasons described above. As a result, governments have put ceilings on interest rates, and have tended to suppress informal lending institutions, while also often subsidizing banks and giving them directives to lend in particular directions. New-style institutions specializing in long-term industrial lending were intended to overcome the pessimism of traditional bankers. But high interest rates for some borrowers, as well as lack of access for others (even some who would have succeeded, and easily paid their debts) may just as well be signs of a healthy loan market as of a badly operating one.

How does all this relate to small enterprises? For any given expenditure on administration, appraisal, and surveillance, the cost varies in inverse proportion to the size of the loan. Not only this, but the amount of appraisal and surveillance needed to attain a given degree of risk is higher for small enterprises because their promoters and managers are generally less experienced than their counterparts in larger enterprises. If the loans are for new ventures or start-ups, the risks are still greater. The promoter has no track record to go by, and fund collateral cannot normally be offered. Furthermore, it is well known that a high proportion of new businesses fail within a few years. (Indeed, a high failure rate at this stage may be regarded as "healthy.")

The costs of reducing the risk of nonrepayment vary with the type of lending institution. The large, formal institution with little or no access to the local grapevine or personal information, and with no way of bringing moral influence to bear on debtors, is at one end of the spectrum. Friends and relatives lie at the

other end. In between is a wide variety of informal institutions, which have or are able to acquire local knowledge at little cost and are more or less capable of exerting pressure on debtors.

I now turn to a very brief overview of capital markets and loans to small enterprises in developing countries, based on numerous surveys of both borrowers and lending institutions. The picture that emerges accords well with the above *a priori* account of the nature of capital markets.

The vast majority of private sector firms start very small in terms of both capital and number of employees. The initial capital is predominantly provided by the promoter's own savings, with a leavening of contributions from friends and relatives.¹⁵ Bank loans are very rare, and recourse to moneylenders is also rare. Bank loans are available in modern industries for which the firm start-up size is not so very small, or if the entrepreneur is already established and running other businesses and hence has an observable reputation. (India is, to a limited extent, an exception, and I examine its case below.) As a firm grows in size, if it grows at all, the probability of its borrowing from banks increases quite rapidly.

If special institutions exist to lend to small and medium-size enterprises, and (as is usual) interest rates are controlled at very low levels, the loans go predominantly to enterprises at the upper end of the permitted size range. When loans have been made to very small and new enterprises by development banks, the default rate has been high—often catastrophically high (Levitsky 1983, Rangarajan 1980). (This result is common for similar institutions in developed countries.) This is not to say that it is impossible to lend profitably to the very small. There are some examples of success, mostly in the shape of either small, private, locally based and managed banks or clubs and cooperatives that can exert pressure on defaulters. Here loans are made on the basis of the character and experience of the recipient rather than any analysis of the business or project.

What can we say about the cost of capital? Most developing-country governments control interest rates and influence the allocation of credit. In general, this favors the large over the small. Credit is usually rationed in favor of large-scale modern industry and exporters. This either drives up interest rates to small enterprises, or, if lenders' rates are effectively controlled, enables fewer to obtain credit. The smaller the enterprise, the more likely it is to be denied credit. To be sure, in some countries there are special schemes for subsidized lending even to the very small, but the proportion of small enterprises reached by such schemes is invariably minute.¹⁶

However, the cost of loans is not the same thing as the cost of capital. As we have seen, very small enterprises rely predominantly on their own savings or those of friends or relatives. What is the private opportunity cost of such finance? This is hard to say. There may be opportunities for small savers to lend at

15. For some illustrative percentages, see Little, Mazumdar, and Page (forthcoming, chap. 15).

16. For an empirical examination of these issues, see Levitsky (1983).

high rates in informal markets. But this requires know-how and may be very risky. Not everyone is cut out to be a moneylender. The only reliable outlet is usually a savings deposit, and in many countries the interest rate on deposits is controlled at a level that results in a negative real yield. Thus controlled credit markets can result in an excessive stimulus to the formation of small business.¹⁷

The tentative conclusion seems to be that the controlled capital markets of most developing countries are likely to penalize the large-small or medium-size firms (covering about 20–100 workers) that aspire to rapid growth and therefore cannot rely on their own finance. Unless they are in a specially favored sector, either interest rates will be higher or access more difficult than with free capital markets.

V. RELATIVE FACTOR COSTS AND FACTOR PROPORTIONS

It has been shown that the cost of capital may be very low for very small firms, since they predominantly use their own savings for which other outlets may have very low yields. But it seems likely that the cost of capital varies greatly from one such enterprise to another, and this could help explain the variance of capital intensity (which is particularly high among the small size groups). Use of institutional finance rises quite rapidly as size increases. Even when there are low interest rate ceilings, banks often find other ways of charging more for loans. Unless government regulations can effectively prevent it, the cost of institutional finance almost certainly falls as size increases.

It was also shown for manufacturing as a whole wages tend to rise with size, but that within the range of up to 100–200 workers, the rise is relatively moderate, and that some surveys of narrowly defined industries show little or no rise in returns to unskilled labor through this range.

It can also be argued with some force that capital intensity causes variations in relative factor costs. There are independent reasons for variations in capital intensity, deriving from the quality and nature of the product even within narrowly defined industries, and it has been shown that capital-intensive firms are likely to want to pay high wages. This does not, however, preclude causation working in the opposite direction, and both directions of causation may operate simultaneously.

At a disaggregated level, the Indian surveys of particular industries found little significant variation of capital intensity with size (up to about 200 workers) while wages rose moderately or not at all (see tables 1 and 3). The exception was a large jump in capital intensity for shoemaking firms with over 100 workers, but unfortunately we have no wage data for these. Unfortunately again, the

17. It may not be a coincidence that the number of small manufacturing businesses ceased to grow in Korea at a time when deposit rates were permitted to rise to very high real levels. But good employment opportunities also reduce the rate of formation of new enterprises—the reverse of what is observable in Europe today.

surveys yielded no figures for the cost of capital, though in light of the controls on interest rates it may be surmised that any rise in capital costs would be moderate. Another exception was powerloom firms, in which wages rose rapidly with size, but there are no figures for these on capital intensity by size. It is clear that these Indian surveys did not produce data which could either confirm or refute the presumption that varying factor costs cause variations in capital intensity. As the surveys did not yield usable data on capital costs, it was not possible to generate correlations between factor proportions and factor costs. This would seem to be a fruitful area for future research.

The 1968 Korean census was remarkable in that among 213 industries, labor intensity peaked with about equal frequency over all size classes up to 500 workers (Ho 1980, p. 58). Unfortunately, the disaggregated earnings data from this census have not been analyzed, and there are no figures for the cost of capital, so again there are no figures for relative factor costs. However, it would be surprising if these relative costs did not favor a rise in capital intensity in many cases. The Korean result thus invites further investigation at the disaggregated level.

VI. EXPLAINING TECHNICAL EFFICIENCY, PROFITABILITY, AND GROWTH

The Indian surveys produced enterprise data for technical efficiency (the total productivity of factors), profitability, and the growth rates of employment and sales. These were correlated with the size and age of the firms, the age of the equipment, changes in capacity utilization, and various attributes of the entrepreneurs and the work force.¹⁸

There were wide variations in technical efficiency. As was mentioned in section II, the only industry for which size was significant was machine tools, in which large firms were more efficient. The best explanatory variable for firms' technical efficiency was the firm-specific experience of the labor force. (This corresponds with Stiglitz's basic theoretical model [1974] and with empirical evidence from other countries found by Chapman and Tan [1980] and Knight and Sabot [1982].) Its coefficient was positive in all of the four cases examined, and significant at the 1 percent level in three. In the latter three cases the elasticity was close to 0.2, that is, a 10 percent rise in the average length of service of the work force was associated with a 2 percent reduction in costs. This confirms the importance of a stable labor force, and the value to a firm of paying a relatively high wage (but there is no estimate of the wage elasticity of length of service). More surprisingly, managerial experience (industry-specific, not firm-specific) was significantly positive (at the 5 percent level) only in soap. This may have been because the manager was also usually a worker. The coefficient was actually negative in machine tools, though barely significant, which suggests

18. For a detailed description of this analysis, see Little, Mazumdar, and Page (forthcoming, chaps. 11, 12).

that in some industries older managers may get out of date. The literacy of the manager was positively significant, but any further education proved irrelevant.

The age of the enterprise, after controlling for the experience of the labor force, was a negative influence on efficiency in all cases. This is very probably because the age of the enterprise reflects the vintage of the equipment.¹⁹ In the two cases in which capital vintage could be measured directly, the coefficient was, not surprisingly, significantly negative.

I now turn to profitability. This was simply defined as value added less the wage bill, divided by the value of capital. Thus "profits" were gross of both interest and depreciation, and they might better be termed the "economic surplus." The Indian surveys show little relationship between profitability and size. In machine tools and metal casting, the smallest group (6–10 workers) showed significantly low profitability compared with the largest (51 plus workers); in printing, it was exactly the reverse. But there was some indication that capital-intensive firms were less profitable than labor-intensive ones in soap, metal casting, printing, and shoes. There was also a tendency for the variance of profitability to be greater in the smaller firms. I cannot from this evidence generate any strong dynamic argument to the effect that one size group yields more economic surplus than another. However, given the fact that failure rates are much higher among small firms, which is not allowed for in the analysis, it could be argued that small firms yield less economic surplus.

An attempt was made to explain profitability in terms of the same variables as technical efficiency (Little, Mazumdar, and Page, forthcoming, chap. 12). No significant variable was found in the case of metal casting. In machine tools, the age of the firm and the experience of workers were both significant (1 percent level). In soap, only high school education was significant (1 percent level); it explained 22 percent of the variance. In shoes, college education was significant (1 percent level); together with high school education (almost significant at the 10 percent level), it explained 9 percent of the variance. It is interesting that, although their explanatory value was low, intermediate and higher education were often significant for profitability (and growth, as will be shown), while only literacy was significant for technical efficiency. Presumably differing types of educational experience will be important for efficiency and profitability; for example, social contacts yielded by higher education may improve access to markets but may have little impact on technical efficiency. It is also notable that the experience of the work force—the most successful variable for technical efficiency—did not affect profitability.

I now turn to growth rates. These are of interest in light of the argument that the small firm sector is an important source of economic, and particularly employment, growth. It cannot be too strongly emphasized that only surviving enterprises were sampled. Consequently the average growth rate of any initial

19. It should be recalled that the measurement of capital was the undepreciated constant price cost of fixed assets, plus working capital.

Table 4. *Percentage of Sample Indian Firms in Very Small Size Group and Average Ages of Firms*

<i>Industry</i>	<i>At founding</i>	<i>At time of survey</i>	<i>Average age</i>
Shoes	91	65	12
Soap	83	63	19
Printing	83	34	16
Machine tools	74	6	6
Powerlooms	65	20	8
Metal casting	41	15	31

Note: Very small firms are those with 1–10 workers.

Source: Little, Mazumdar, and Page (forthcoming, table 12-11).

cohort is exaggerated and the smaller the size the more it is exaggerated. The mean employment growth rate of survivors over their lives was very high in powerlooms, shoes, machine tools, and printing (from 15 to 27 percent), lower in soap (7.1 percent), a slow in metal casting (0.7 percent). Most firms, except those in metal casting, started very small, with 10 or fewer workers.

Table 4 shows the percentages of surviving firms that started very small (1–10 workers), and the percentages that were very small at the time of the survey, together with average ages. There was a large fall in the proportion of the cohort in the very small size class in all industries except shoes and soap, and even in these two industries, in which small firms are utilizing primitive techniques, a substantial proportion have “graduated.”

The growth rates shown in table 5 reflect the fact that young survivors grow much faster than older firms, except in the very sluggish metal casting industry. The variance of growth rates is also higher for the small. If the failed firms could also have been included, the growth rate of employment attributable to the initial cohort of the small would undoubtedly have been very much smaller, and it could even have been less than that of larger firms.

As with technical efficiency and profitability, attempts were made to explain growth differentials by reference to the characteristics of firms. In the case of

Table 5. *Mean Employment Growth Rates by Age of Firms in Indian Survey*

<i>Age (years)</i>	<i>Machine tools</i>	<i>Powerlooms</i>	<i>Printing</i>	<i>Shoes</i>	<i>Soap</i>	<i>Metal casting</i>
1–5	41 (7)	47 (6)	71 (8)	48 (20)	19 (3)	4 (3)
6–10	24 (21)	9 (22)	17 (16)	13 (24)	11 (7)	6 (6)
11–15	16 (14)	7 (9)	10 (7)	4 (7)	7 (6)	5 (4)
16–20	9 (19)	n.a. (0)	9 (11)	6 (15)	8 (9)	7 (9)
21–25	10 (10)	n.a. (0)	1 (2)	3 (1)	4 (14)	2 (5)
26 plus	8 (4)	n.a. (0)	4 (10)	2 (8)	3 (8)	(9)

n.a. Not applicable.

Note: Figures are annual compound growth rates based on the firms' initial size. Figures in parentheses are number of firms.

Source: Little, Mazumdar, and Page (forthcoming, table 12-15).

growth, however, initial rather than final size was used to categorize firms. The dependent variable is the growth of sales.²⁰ For machine tools, no variable was remotely significant. For metal casting, only initial size was (negatively) significant (1 percent level). For printing, age was negatively significant (1 percent level), and graduate education positively so (5 percent level); the two together explained 20 percent of the variance. For soap, age and graduate education were similarly significant (5 and 10 percent levels), and so also was the experience of the entrepreneur, 30 percent of the variance being explained. For shoes, age and initial size figured negatively, and college education positively, 11 percent of the variance being explained (but all only at the 10 percent level). Thus, as with profitability, intermediate or higher education was clearly significant for three industries. It is somewhat surprising that education played no role in the two metalworking industries, either for profitability or for growth.

Except for the confirmation that the education of the entrepreneur is of some value in explaining profitability and growth, the regression analysis yielded little. That age or initial size (the two being closely related) were prominent in explaining the growth of surviving firms is only common sense. Initial size and the education of the entrepreneur were also closely related. Clearly the better educated either have more money, are better placed to tap sources of money, or both.

VII. INDIAN POLICIES

Indian small-scale enterprise policies are of particular interest because India is the only country with a really extensive set of policies supporting and protecting many kinds of small enterprise, while also being the country where very large enterprises (over 1,000 workers) dominate manufacturing *factory* employment to a greater extent than any other country in the world—including the United States.

Indian concern for handicrafts dates back at least to the early years of this century. Under the influence of M. K. Gandhi, *khadi*, the almost vanished art of hand spinning, was resuscitated and became a political and cultural symbol (Gandhi 1940). No other country has for three-quarters of a century revived and expanded primitive methods of making not only cloth, but also sugar, soap, and several other products. In the 1950s, however, India also became notable for its insistence on the importance of heavy industry, under the influence of J. Nehru and his close adviser, P. C. Mahalanobis. Although Nehru's views prevailed for the vast bulk of industrial investment, Gandhian thinking has continued to be influential. The Mahalanobis heavy industry model reserved a place for the Gandhians in that "small and household industries" were supposed to supply the

20. Growth of employment was also used as the dependent variable. The results were very similar to those reported.

increased demand for consumer goods with very little investment but greatly increased employment.

The Gandhians wanted an expansion of traditional, mainly rural, manufacture rather than the creation of modern small urban factories, and a series of boards was established in about 1950 to support traditional industries. It was not until rather later that the Small Scale Industries Board was established to encourage small factories producing modern products—"small" at the time excluding enterprises with over Rs500,000 of fixed assets (Rs100,000 without power) or over fifty employees. By the beginning of the second five-year plan (1956), the main administrative framework for government encouragement of both traditional and modern small enterprises was in place. Thus, a pattern for policy was established that has held ever since. These policies protect traditional manufacturers by taxing and banning the expansion of medium- and large-scale enterprises that compete with them, by directly subsidizing them or exempting them from taxes, and by giving them preference in government purchases.

However, from the second through the fifth five-year plans the amount spent on these policies was a small and declining proportion of the amount spent on "organized" industry. The traditional or village industries got the lion's share and, except in the case of textiles, probably also benefited most in the 1950s and 1960s from those protective measures that did not involve public money. But in the late 1970s there were moves to preclude larger enterprises from increasing their output of a greatly expanded number of modern products (over 800 at the peak), with the intention that growth of production would then be achieved by increases in the number of small enterprises. These products were predominantly in the light engineering and chemicals industries. Apart from this product reservation, powerful encouragement was given to SSES by exempting them from excise taxes, or by reducing the rates of taxes that depended on their volume of output.

Since the nationalization of the commercial banks in 1969 there has been increased pressure on them to lend to small enterprises, with an impressive increase in such loans. The number of recipients is estimated to have risen from 57,000 in 1969 to 301,000 in 1977 and 794,000 at the end of 1980, and the proportion of total bank credit accruing to the small has risen from 8 to 13 percent (Ojha 1982; Rangarajan 1980). These loans are made at preferentially low rates of interest. The normal result, that of reducing the amount of credit available to the favored group, has been avoided by forcing the banks to devote a certain proportion of their advances to priority sectors (including SSES) and by compensating the banks for the increased risk through a Reserve Bank of India guarantee scheme. However, despite these efforts, I would estimate that only a very small proportion, perhaps 3 percent, of very small manufacturing enterprises (less than 20 workers) got loans in 1977. Although reported default rates are low by comparison with much experience elsewhere, the expansion has been too recent to assess the real costs, let alone the benefits.

At this point it must be noted that over the years the official definition of small

has been changed frequently. The employment limit of fifty workers was removed early on. Now the limit is Rs3.5 million in fixed assets (or Rs4.5 million for “ancillary” units). Since the value of fixed assets is usually half or less of the value of total assets, it is clear that many enterprises are included that are medium-size in the sense that they employ more than fifty workers.

What have been the results of these policies? It was stated at the beginning of this section that Indian factory employment is heavily concentrated in very large establishments. This is, of course, consistent with a high proportion of total manufacturing employment being in households and workshops, as is indeed the case. The dearth is especially in small and medium-size factories. This was noticed long ago. In 1961 Dhar and Lydall wrote: “A peculiarity of the Indian distribution is that, while it has a high concentration of *establishments* in the lower size group, it has a high concentration of employees in the highest size group. Indian industry . . . is somewhat thin in the middle. . . . This is rather an important conclusion, for the medium sized firm (say 50–99 employees) is very often close to the optimum size, especially in the lighter manufacturing trades” (pp. 30–31).

In about 1960, the contrast between India (and Pakistan) and Korea and Taiwan (and even Japan) was enormous. Broadly speaking, Indian factory employment was concentrated in establishments of over 1,000 workers, and East Asian employment in factories of less than 50 workers.²¹ In the 1960s and 1970s, medium-size and medium-size-to-large factories with 100–1,000 workers grew relatively fast in India, while employment in the East Asian economies became much more concentrated in the range 100–500 workers. The contrasts are thus no longer quite so extreme, but nevertheless the extraordinary dominance of the large unit in India is still far from being eliminated. For instance, the range 50–500 workers accounts for well over half of factory employment in the three East Asian economies, while in India it accounts for less than a third.

It is doubtful whether the partial filling of the middle size range in India owes much to the small-enterprise policies described. First, in the 1960s the main beneficiaries were very small traditional units, the thrust in favor of small modern factories occurring only in the second half of the 1970s. Second, enterprises in the range 500–1,000 workers, and most in the range 100–500, will be outside the range of the “small.” Indeed the policies both present a barrier to organic growth above the size determined by the value of fixed assets that defines the small and seem certain to create an artificial trough in the size distribution, even if that trough occurs at a rather larger size than before.

The effect of policies on three industries, textiles, sugar, and light engineering, are now considered in more detail.²² In the 1950s, the government banned the installation of new looms by the textile mills, except for replacement or export

21. See Little, Mazumdar, and Page (forthcoming, chap. 6) for a presentation of the evidence and for the measure of concentration used.

22. The textile industry findings reported here are based on a study by Mazumdar (1984a); the remainder of this section is derived from Little, Mazumdar, and Page (forthcoming, chap. 4).

production. It later reserved synthetic cloth production for the powerloom and handloom sectors, which also paid no excise duty (unless, *in theory*, more than five powerlooms were operated).²³ The intention was mainly to promote handloom production, it being envisaged in the second five-year plan that only 10 percent of increased cloth production would come from powerlooms. The outcome was different. From 1956–81, mill production fell from 4.9 to 3.1 billion meters (billion is 1,000 million). Small-sector output rose by 3.3 billion meters, but three-quarters of this increase came from the powerlooms, which the surveys showed to be much more profitable. Powerlooms have a further advantage in that handlooms are not very suitable for weaving the synthetic cloth that is denied to the mills. It is not even certain that handlooms have, on balance, benefited at all from the government's policies, for their production, which rose at 2 percent per year, was also rising slowly in the first half of this century, being "protected," as now, by very low wages. Mill wages are now between two and two and a half times those paid by the master weavers who control both powerlooms and handlooms.

The central question in considering the social value of the government's textile policies is thus whether powerloom production has been *socially* more profitable than mill production. A crude cost-benefit analysis—crude because a good analysis of the relevant mill costs never seems to have been made—suggests the opposite. Ignoring the distributional consideration that millworkers are wealthy compared with powerloom operatives, a switch from powerlooms to mills appears to have a high social yield at any plausible shadow wage rate. If direct employment in weaving alone is considered, it is true, of course, that given the level of mill wages, the policies have led to an increase in employment, since powerlooms use about three times as much labor per unit of output (and pay less than half the wages). But the cost-benefit analysis indicates that this employment has been "bought" too dearly. Furthermore, taking the government's textile policies as a whole, it is certain that exports of both textiles and, more important, clothing, which is highly labor-intensive, have been seriously damaged. The impact of this in reducing employment could well have offset the increased employment generated in the powerloom sector. Finally, everywhere else in the world the poor person consumes cheap, durable synthetic cloth, while cotton has become a luxury. This suggests that the policies may not have been in the best interests of the poorest—or other—classes of consumers, an effect that must also be taken into account in assessing their overall welfare effect.

Sugar policies have both differences from and similarities to textile policies. For textiles, the intention was to promote the traditional sector, while the main beneficiary was actually a sector somewhere in the middle. For sugar, the promotion of the intermediate product (*khandsari*) was intentional, and the tradi-

23. Powerlooms are similar to the nonautomatic looms used in the mills. Often they are bought secondhand from the mills, although new looms are also produced for the "powerloom sector" that are cheaper and of lower quality than those normally bought by the mills. The powerloom sector thus consists of small units that operate, with power, a few relatively old or cheap looms.

tional labor-intensive product (*gur*) was neglected and suffered. Unlike textiles, the labor input is quite small for both the sugar mills and *khandsari* units. There was more official analysis of the alternatives (but not of *gur*) than in the case of textiles, but it was not objective. *Khandsari* won the day as the result of intermediate technology enthusiasm and probably the lobbying of the main beneficiaries, who fall among the wealthier sections of rural society. The main losers have been the exchequer and the *gur* industry. Subsequent intensive and careful cost-benefit analysis shows mill production to be superior to *khandsari* on any reasonable assumptions.²⁴

Small and medium-size light engineering enterprises have been the other most important beneficiaries of the government's product reservation policies. The effects on the production of diesel engines, bicycles and their components, and handtools have been scrutinized. The finding is that the policies have had quite dramatic effects on the structure of these industries and on the adoption of technology. SSE reservation has resulted in reductions in the linkages between small and large enterprises, including the development of specialized subcontracting firms, although specialization by small enterprises and links between them remain important. The adoption of new products and processes has been inhibited, with consequential effects on exports, especially in the case of diesel engines and bicycles.

What is there to offset these disadvantages? These industries produce mainly capital goods, which reduces the potential to claim good distributive effects on the consumer side. Also, enough has been learned of the unreliability of any purported relationship between size and labor intensity within narrowly defined industries to indicate that any claim that the employment of unskilled labor has been promoted is unlikely to be sustainable.²⁵

More generally, it seems clear that product reservation policies reduce competition both within the large- and small-scale sectors of the industry and between them. It prevents organic growth of establishments through the size barrier and artificially drives the large-scale sector to concentrate on higher priced goods, for which quality differentials are important.

VII. AFTERTHOUGHTS, SUMMARY, AND CONCLUSIONS

Household manufacturing, which is confined to a few industries and is most prevalent in rural areas, has declined, relatively if not absolutely, in all the economies examined (India, Korea, Malaysia, the Philippines, and Taiwan). Workshops (1–5 workers) have also declined, except possibly in India. Yet cottage shop manufacturing still accounts for over half of manufacturing employment in the poorer countries (India, Indonesia, the Philippines, and most of

24. This paragraph is based on the work of de Haan (n.d.), which is more extensively reported in Little, Mazumdar, and Page (forthcoming, chap. 4).

25. This paragraph is based on the work of Berry, reported in Little, Mazumdar, and Page (forthcoming, chap. 4).

Africa). There has also been a relative fall in employment in small-to-medium-size factories (5–100 workers) in the more rapidly industrializing economies (Colombia, Korea, Malaysia, Singapore, and Taiwan). By contrast, in India the number of enterprises in the range 10–100 workers has risen since 1960 relative to large ones, especially those with over 500 workers, which partly redresses a previous anomaly.

The relative decline of small-scale enterprises in most developing countries, including many not mentioned here, has been accelerated by the industrialization policies adopted in these countries. Protection, investment incentives, credit control, and the promotion of industry in the public sector have all discriminated against the small. But another factor was important for the highly export-oriented economies of East Asia. Big firms have economies of scale in export marketing, and it is generally true that most small units export little or nothing (although evidence from such countries as Korea suggests the potential of medium-size units as exporters). This partly explains the very rapid decline of small firms in these countries (other probable explanations are good job opportunities and high deposit rates of interest). This decline in sses was compatible with a pattern of development that was highly labor-demanding.

In most countries there have been some countervailing measures in support of sses. But these have only scratched the surface and have had very little impact relative to the general economic policies of governments. India is a special case. In the factory sector (ten or more employees), employment has been and still is concentrated in very large establishments to a greater extent than in any other country that I know of (including the United States). At the same time much small-scale “traditional” manufacturing, especially in households and workshops, has been protected and subsidized. In recent years, however, small- and medium-scale modern enterprises have also been enthusiastically and forcefully promoted over a wider front, especially in light engineering and chemicals.²⁶

In surveys of narrowly defined industries, the idea that small, especially very small, manufacturing enterprises are relatively efficient users of resources in labor-abundant economies has been found to have no general validity. If anything, medium-size firms come out best. But size as such, especially when measured by employment, is a poor indicator of any attribute of social importance. In any case, as shown in section II, such differences in labor intensity as there are between size groups within an industry are dwarfed by differences between industries. This implies that, for manufacturing as a whole, small and very small firms are labor-intensive because labor-intensive industries have a relatively high proportion of small enterprises.

Apart from efficiency of resource use, other claims have been made for sses, such as that they are more innovative than larger firms. The available evidence, coming from developed countries, tends to suggest the opposite.²⁷ The most

26. It was reported in the Indian press that, in 1985, there was a reappraisal, and that there may be some backtracking on sse promotion.

27. See Little, Mazumdar, and Page (forthcoming, chap. 2) for a brief survey of the evidence.

interesting of the dynamic arguments is that SSEs form a seedbed for entrepreneurs, again an argument mostly made by SSE proponents in developed countries.²⁸ But, while true, it says nothing about how large the seedbed has to be. In most developing countries, this argument would seem to lack all force, since there is still an enormous numerical preponderance of very small firms, and the rate of formation is very high (given the high failure rates). Even in Korea, when the proportion of employment in small manufacturing enterprises declined drastically, there was no fall in absolute numbers, and many new firms were created.²⁹ Potential entrepreneurs seem to abound, and more than enough somehow find the resources to make a start. Africa south of the Sahara may, however, be an exception to this generalization.

In theory, low wage costs and high capital costs per unit of output will tend to result in labor-intensive methods (though the within-industry elasticity of substitution between capital and unskilled labor may often be fairly low—less than unity). Do small firms have these attributes? There is no doubt that earnings are generally lower in small enterprises, and this is usually matched by lower labor productivity. To a considerable extent this is because the workers are less able, but even allowing for this, earnings are lower. The increase in earnings with size is not uniform, however, and for unskilled labor in some industries it is quite small and even nonexistent within the small-to-medium-size range. There is some evidence from Colombia and India of a discontinuity around the 200-worker size. This may result from the greater influence of institutional factors on larger firms.

Labor markets cannot be considered to be distorted just because larger firms pay higher wages. There are many good economic reasons for this. Where wages are pushed up by trade unions or government action, however, there is a distortion. There is no doubt that some manufacturing wages are too high from a social point of view (for reasons of both income distribution and employment) in many developing countries, but this observation applies more to large and medium-size-to-large firms than to firms within the small-to-medium-size range.

The common idea that the costs of capital is very high for small enterprises is overly simple. Loans in a free market would certainly cost a lot (or be unobtainable, that is, infinitely costly) for small enterprises. For the lucky few in countries with loan programs for the small firm, they are very cheap because they are heavily subsidized, but concomitantly, the coverage of such schemes is minute. Far more important, however, is the fact that capital and loans from institutions are not the same thing. The opportunity cost of personal savings (entrepreneurs' own, or those of relatives) is often very low because of low rates available on deposits—and very small firms rely almost entirely on such savings. It is only

28. A sufficiency of new starts is more likely to be a problem for developed countries since informal capital markets hardly exist and the initial investment required is higher relative to income levels.

29. A survey in 1973 found that out of a stock of 21,045 small manufacturing firms, 2,510 had been formed in the previous eighteen months. The average initial investment of the new starts was far higher than in the less advanced developing countries (see Ho 1980, pp. 71–77).

when an enterprise cannot grow as fast as it would like using only self-finance that the cost of loans will be the same as the marginal cost of capital. For such firms, the cost of capital falls as size increases (or capital becomes more easily obtainable if interest rates are controlled).

Taking it as agreed that development has been less labor-demanding than was desirable, it may be asked both to what extent government policies are responsible for this result and how they should be reformed. It has been shown that policies have discriminated against small-scale enterprises. But the implication of what has been said is that this has very little to do with the excessive capital intensity of manufacturing that prevails in many developing countries. What is important is first that the pattern of industrial output has been twisted in favor of capital-intensive industries, and second that encouragement has been given to the choice of capital-intensive techniques, whether by small or large units.

International trade delinks domestic demand from domestic supply. For a given pattern of domestic demand, the labor-abundant country should import capital-intensive products and export labor-intensive ones. There is no need to retell here the story of how most developing countries have abrogated this principle with excessive import substitution policies. Shifting the pattern of domestic demand is also important. The government's own purchases, whether on capital or current account, are a large part of total demand. Apart from this, measures to make income distribution more equal will probably shift production in a labor-intensive direction, for the things demanded by the poor—food, clothing, and shelter—are made in a relatively labor-intensive manner. These aspects are self-reinforcing, in that the use of labor-intensive production methods benefits poor, unskilled workers and their dependents. It is no accident that the countries that have had growth led by exports of labor-intensive manufacturers are among those with the most equal distribution of income.

Apart from their effect through the pattern of output, government policies may be argued to have promoted capital intensity directly. Capital has been made cheap in numerous ways. Real interest rates have been held very low and have often been negative. Exchange rates are widely overvalued, while imported machinery is exempted from import duties. Other tax concessions, such as accelerated depreciation, favor the use of capital. Government investments often ignore prices, and the most modern, capital-intensive methods are chosen for reasons of prestige or to satisfy the pride and prejudice of engineers. At the same time the predilection of large enterprises to pay high wages has been reinforced by governments. In many countries and cases the shift in the relative prices of capital and labor has been enormous.

Indian policies have had all these defects, and Indian capital-output ratios are among the highest in the world. But India is also exceptional in the extent and range of its policies that directly support SSES. These interventions share several features. They have been romantic, rather than economic. They were not based on detailed objective inquiries into costs and benefits, and in some cases no such analysis seems to have been conducted at all. Even the most direct consequences

of intervention have been badly misjudged (powerlooms versus handlooms), and indirect consequences have been ignored in every case examined. There is not even any assurance when such indirect effects are allowed for that overall employment was promoted, or that the policies thus offset the unquestionably adverse effect on employment of the general thrust of measures taken to promote industrialization.

One particular intervention, common to India and many other countries, deserves further mention: the endeavor to make formal credit institutions reach down to very small enterprises. It is very doubtful if this is sensible. It is certainly proper to encourage banks to make risky loans by permitting them to charge what they like. To go further than this implies subsidization. Why should banks be subsidized to compete with informal markets or with the friends and relatives of entrepreneurs, which are more efficient suppliers of small loans to the very small, in the sense that the costs that they face in assessing and policing loans are relatively low. There may sometimes be a case for bank support to, rather than competition with, informal credit institutions. But it should not be so large as to change their nature and so become a kiss of death.

In an uncertain world with would-be borrowers of unknown reliability and capacity, it is inevitable that some good projects will not be financed. There is nothing wrong with a situation in which both inexperienced entrepreneurs starting up in a small way and small established entrepreneurs are unable to get institutional credit. Easy access to such credit has in many cases resulted in a very high proportion of defaults. The *lack* of institutional credit can be seen as a filter that arguably does more to eliminate dishonest, incompetent, and sluggish would-be borrowers than it does to prevent potential climbers from setting foot on the mountain of success. There is no shortage of new starts in most developing countries, and of these only very few get institutional credit. And if there were any shortage, the first step to overcome it should surely be the removal of institutional obstacles, such as these blocking access to power, materials, and land. Institutional credit is better seen as a means of facilitating the expansion of firms that have passed the survival stage and have acquired at least the beginnings of a good track record.

Some will complain that the above amounts to a sermon on the bounty of the price mechanism. Economists are adept at finding faults with free markets that undoubtedly do exist. (Others often find faults that do not exist.) Interventions that probably would improve markets and the allocation of resources can sometimes, though not always, be suggested. Unfortunately, government interventions are very rarely designed to improve the working of markets (or create good markets where they do not exist). Rather, they usually override markets. Even when such interventions are well-intentioned, their poor construction means that the outcome is often unforeseen and the results unfortunate. I have referred to a few of many of the possible examples of this. Someone has rightly remarked that while there is more to development than getting the prices right, getting the prices wrong may be the end of development. Overall macroeconomic and

industrial policies frequently result in massive price distortions that favor capital intensity. Trying to offset this with targeted interventions that favor small firms is, at best, fiddling while Rome burns.

REFERENCES

- Anderson, Dennis. 1982a. *Small Industry in Developing Countries: Some Issues*. World Bank Staff Working Paper 518. Washington, D.C.
- . 1982b. "Small Industry in Developing Countries: A Discussion of Issues." *World Development* 10, no. 11: 913–48.
- Anderson, D., and F. Khambata. 1981. *Small Enterprises and Development Policy in the Philippines: A Case Study*. World Bank Staff Working Paper 468. Washington D.C.
- Chapman, B. J., and H. W. Tan. 1980. "Specific Training and Inter-Industry Wage Differentials in U.S. Manufacturing." *Review of Economics and Statistics* 62, no. 3 (August): 371–78.
- Cortes, M., A. Berry, and A. Ishaq. Forthcoming. *Success in Small and Medium-Scale Enterprises: The Evidence from Colombia*. Baltimore, Md.: Johns Hopkins University Press.
- de Haan, H. H. N.d. "The Viability of Small-Scale Industrialization, The Case of Sugar-cane Processing." New Delhi: Institute of Economic Growth. Processed.
- Deshpande, L. R. 1979. "The Bombay Labor Market." Bombay University. Unpublished report.
- Dhar, P. N., and H. F. Lydall. 1961. *The Role of Small Enterprises in Indian Economic Development*. Bombay: Asian Publishing House.
- Gandhi, M. K. 1940. *An Autobiography*, 2d ed. Ahmedabad: Navajivan Publishing House.
- Ho, S. 1980. *Small-Scale Enterprises in Korea and Taiwan*. World Bank Staff Working Paper 384. Washington, D.C.
- India Central Statistical Organization (Industrial Statistics Division). 1979. "Wages and Productivity in Organized Manufacturing Sector, 1960–61 to 1976–77." Bulletin 15D/5 (August). New Delhi: Department of Planning.
- Indian Labor Journal*. 1981. December.
- Kaneda, H. 1980. "Development of Small and Medium Enterprises and Policy Response in Japan: An Analytical Survey." Studies in Employment and Rural Development 32. Washington, D.C.: World Bank Development Economics Department. Processed.
- Knight, J., and R. H. Sabot. 1982. "From Migrants to Proletarians: Employment Experience, Mobility and Wages in Tanzania." *Oxford Bulletin of Economics and Statistics*, 44, no. 3 (August): 199–226.
- Lester, R. 1967. "Pay Differentials by Size of Establishments." *Industrial Relations* 71, no. 7 (October): 57–67.
- Levitsky, J. 1983. "Assessment of Bank Small Scale Enterprise Lending." Washington, D.C.: World Bank Industry Department. Processed.
- Little, I. M. D., Dipak Mazumdar, and John M. Page. Forthcoming. *Small Manufacturing Enterprises: A Comparative Study of India and Other Countries*. New York: Oxford University Press.

- Manning, Christopher S. 1979. "Wage Differentials and Labor Market Segmentation in Indonesian Manufacturing." Ph.D. thesis, Australian National University.
- Mazumdar, D. 1981. *The Urban Labor Market and Income Distribution: A Study of Malaysia*. New York: Oxford University Press.
- . 1984a. *The Issue of Small Versus Large in the Indian Textile Industry: An Analytical and Historical Survey*. World Bank Staff Working Paper 645. Washington, D.C.
- . 1984b. "The Rural-Urban Wage Gap, Migration and the Working of Urban Labor Markets: An Interpretation Based on a Study of the Workers of Bombay City." *Indian Economic Review* 18, no. 2: 169–98.
- Mehta, B. V. 1969. "Size and Capital Intensity in Indian Industry." *Oxford Bulletin of Economics and Statistics* 31, no. 3 (August): 189–204.
- Mohan, Rakesh. 1981. *The Determinants of Labour Earnings in Developing Metropolises: Estimates from Bogota and Cali, Colombia*. World Bank Staff Working Paper 498. Washington, D.C.
- Ohkawa, K. and M. Tajima. 1976. *Small-Medium Scale Manufacturing Industry: A Comparative Study of Japan and Developing Nations*. Tokyo: International Development Center of Japan.
- Ojha, P. D. 1982. "Finance for Small-Scale Enterprise in India." *Reserve Bank of India Bulletin*, November.
- Page, J. M. 1979. *Small Enterprises in African Development: A Survey*. World Bank Staff Working Paper 363. Washington, D.C.
- . 1984. "Firm Size and Technical Efficiency: Applications of Production Frontiers to Indian Survey Data." *Journal of Development Economics* 16 (September–October): 129–52.
- Park, Fun Koo. 1981. "An Analysis of Wage Differentials according to Firm Size." *Korea Development Review*, December.
- Rangarajan, C. 1980. "Innovations in Banking: The Indian Experience, Impact on Deposits and Credit." Domestic Finance Studies 63, Washington, D.C.: World Bank Development Economics Department. Processed.
- Reserve Bank of India. 1979. "Survey of Small Industrial Units 1979." Department of Statistics, *Statistical Report*, vols. 1 and 2. Bombay.
- Stiglitz, J. E. 1974. "Alternative Theories of Wage Determination and Unemployment in LDCs: The Labor Turnover Model." *Quarterly Journal of Economics* 88, no. 2 (May): 194–227.
- Subramanian, K. N. 1979. *Wages in India*. New Delhi: McGraw-Hill.
- Wilkinson, R. 1973. "Earnings and Size of Plant in the Iron and Steel Industry." Cambridge, England: Cambridge University Department of Applied Economics. Processed.
- World Bank. 1979. "Cottage and Small Industry in the National Economy: Indonesia." Report 2490-IND (November), East Asia and the Pacific Regional Office, Washington, D.C. Processed.

A Valuation Model for Developing-Country Debt with Endogenous Rescheduling

Gerard Gennotte, Homi J. Kharas, and Sayeed Sadeq

Creditors have little recourse if a sovereign state repudiates its external debt obligations. They can, however, threaten to impose penalties if such action occurs which results in deadweight losses to the system as a whole. A preferred alternative for both borrower and lender is to recontract debt obligations. Reschedulings are a device that creditors can use to structure the incentives faced by borrowers such that repudiation is never a rational action.

This article develops a numerical method of valuing the option to reschedule. The model shows why fees are preferred to higher interest spreads during a rescheduling exercise; why maturities get shorter prior to a debt crisis but are lengthened in the rescheduling; why tougher supervision by regulatory authorities could be damaging to renewed voluntary loans; and why little has been done to attempt to seize the assets of countries that have not repaid any interest or principal for extended periods of time.

The model shows that lenders are willing to commit greater amounts if reschedulings are possible than if they are not, and that precommitment to provide additional funds at rescheduling can raise the market value of existing debt and should not be construed as concessions by commercial lenders. Alternately, the model can be used to improve systems for ranking country creditworthiness, to assess the degree of adjustment required to spark a full resumption of spontaneous lending, or to estimate by how much interest rates would have to fall to restore a country's creditworthiness.

In a period of financial crisis for heavily indebted developing countries, creditor guidelines are needed for optimal choices between rescheduling outstanding debt and demanding full repayment. The dilemma for the creditor is that the latter may precipitate a repudiation of debt by the borrowing country, because of an inability or unwillingness to bear the full burden of debt servicing in the short term. These decisions and the terms of rescheduling depend on the estimated valuation of a country's foreign exchange earning capacity, which is a determinant of the market value of its debt. Debt value is also a function of the variance in foreign exchange earnings, front-end fees, the interbank "prime" rate and the interest rate spread, new lending available to the country, debt maturity, and the

Gerard Gennotte and Sayeed Sadeq, consultants to the World Bank, are at the University of California, Berkeley, and at Credit Suisse First Boston, respectively. Homi J. Kharas is at the World Bank.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

amount of assets which a creditor could seize upon default. Reaching agreement on repayment terms of developing-country debt for this range of variables is understandably complex. Recent international debt negotiations reflect the need for further comprehensive analysis of debt valuation.

A predominant share of the international commercial debt of developing countries is in the form of syndicated credits. These credits are held by a relatively small number of banks. Absence of a widespread secondary market ensures that developing-country debt remains in the hands of a few known agents. A critical feature of such an arrangement is that it permits a negotiated rescheduling of obligations. Indeed, fourteen developing countries rescheduled commercial debt in 1983, in the midst of a severe downturn in the global economy, and there were a further twenty reschedulings in 1984. These reschedulings did not involve write-downs of principal or concessional interest rates. By contrast, during the last wave of reschedulings in 1932–33, during the Great Depression, international creditors lost up to 90 percent of the face value of their developing-country debt. These loans were in the form of bonds which automatically fell into default once coupon payments were missed.

The likelihood of external shocks triggering debt repudiation has been emphasized by Eaton and Gersovitz (1981) and Sachs and Cohen (1982). These papers suggest that repudiations stem from strategic behavior by debtors. When global conditions change, so do the relative payoffs from continued debt service and repudiation, and the latter may become an attractive option for the debtor. Kulatilaka and Marcus (1985) discuss the optimal timing of the repudiation decision. The analysis illuminates important features of developing-country debt: the existence of endogenous credit rationing, the need for self-enforcing contracts, the role of uncertainty, and the advantages of precommitment by a borrower. It is, however, couched in the context of bond financing, which limits strategic behavior by creditors. In addition, it predicts that reschedulings would involve concessional terms or debt write-downs. This does not seem to be an accurate description of recent events.

In this article we provide a more realistic framework in which a creditor is able to avoid the occurrence of a strategic repudiation by agreeing to reschedule debt payments under conditions of financial distress. We assume that a repudiation or default results in a deadweight loss.¹ It is then in the best interests of all parties to design a flexible loan contract. A country would prefer to reschedule rather than repudiate its debt because it then retains the option to repudiate at a later stage. The lender prefers rescheduling because there is then some probability that favorable future developments will allow it to recover the full value of its debt.

This article describes a methodology for valuing debt claims on developing countries held by commercial banks. The value of the claim is given by the

1. Repudiation refers to a debtor's action to unilaterally abrogate a debt agreement. Default refers to a creditor action to terminate a debt contract when a borrower fails to meet the contract obligations.

discounted stream of expected future debt service payments (the contractual payments weighted by the probability that a payment will be made). A lender's behavior is motivated by the desire to maximize the value of its claim, which it does by rescheduling and by choosing the maturity, spread, and fees on its loans. The value of the debt claim (the "market" price of the debt) will differ from its nominal face value because of uncertainty about the evolution of the economy's productive assets, which generate the revenue to service the debt. This contingent claims aspect of sovereign debt is discussed in section I.

The outcome of the valuation process is an estimate of the market price of the debt, the price that would prevail if secondary markets existed.² This price is useful for several reasons. First, it provides a cardinal measure of creditworthiness. If the price of debt in one country is higher than in another, it can be called more creditworthy in the sense that less adjustment would be required to restore full creditworthiness. Second, the price provides information about when and whether to expect new money to be forthcoming. Spontaneous lending from new creditors will occur only if the price of a debt claim is greater than or equal to its nominal face value. Lending from existing creditors may occur in other circumstances, however, as it could raise the value of the existing stock of creditors' debt. Third, the approach allows the effects of exogenous variables, such as changes in the interest rate, to be quantified in a systematic way by analyzing their impact on the price of debt. Fourth, some light is shed on rescheduling policies and the role of alternative financial instruments, such as bonds. In this respect, a distinction is made between a rescheduling to overcome liquidity problems and one designed to alleviate insolvency.

Section II presents a numerical valuation model for developing-country debt, based on a Black and Scholes (1973) option pricing scheme in which the value of the debt varies with the value of the underlying assets. The novelty here is that sovereign debt never need be actually paid off as long as it is serviced. A standard option pricing formulation, however, only considers a finite maturity. Section III presents some simulation results for various parameters. We also illustrate the impact of optimal rescheduling on the value of debt. This supports the conjecture (Sachs 1983) that the development of syndicated Eurocredits was an important innovation that promoted the sizable growth of external liabilities of developing countries. In the same framework, we are able to explain certain characteristics of the developing-country loan market—the shift to short-term maturities as the stock of debt grows; the existence of credit rationing and the associated cutoff of new lending at certain times; the process of reschedulings involving high front-end fees but only minor changes in interest rates; and the absence of any attempts by creditors to push for tough legislation to increase their ability to seize country assets in the event of a default.

2. Fledgling markets do exist for developing-country debt, but information on prices and quantities traded is very scanty. The markets are more in the nature of bilateral swaps than of competitive trades.

I. CONTINGENT CLAIMS ASPECTS OF SOVEREIGN DEBT

International lending to developing countries takes place in an institutional framework from which collateral, seniority provisions, and the seizure of assets upon bankruptcy or default are absent.³ This distinguishes sovereign debt contracts from corporate debt. Freed from the obligation to perform or pay damage measures upon unilateral breach of a contract, parties will adhere to contract terms only if the gains from doing so exceed the costs. When a firm goes bankrupt, creditors receive ownership of the real assets; by contrast, creditors of sovereign countries are able to recover only a small fraction of a debtor country's tradable wealth. Creditors can, however, inflict deadweight losses on a country that repudiates its debt by restricting its future access to financial markets and by creating general impediments to its international trade. The existence of these penalties allows some international lending to take place until outstanding debt reaches a ceiling (Eaton and Gersovitz 1981; Sachs and Cohen 1982).

As with any financial claim, the value of debt to creditors is determined by the present value of future repayments. The occurrence of a repayment, however, depends not only on the country's *ability* to repay but also on its *willingness* to repay. If creditors were able to distinguish between these two behaviors, they could choose not to apply deadweight penalties when repudiation results from an inability to repay. It is reasonable to think, however, that creditors cannot discriminate among borrowers in this way. Furthermore, enforcement of penalties in a particular case would be important to maintain the credibility of the threat in other cases. Since neither party benefits from such punitive measures, it is in their best interests to reduce ex ante the likelihood of unilateral repudiation. This can be done by designing flexible contracts that allow some sharing of the risk of unforeseen events without precipitating a repudiation or default.

Suppose a country has borrowed up to its credit ceiling. Following a large adverse shock, it may decide that repudiation is preferable to the further resource outflow involved in meeting a debt repayment obligation. It is then in the best interest of the creditor to negotiate and reschedule part or all of the required payment. In the latter case, the country's incentive to repudiate is removed: it does not suffer the immediate penalty associated with repudiation, and it retains the option to repudiate in the future. Similarly, the bank is better off: an immediate payment would not have been made, regardless, but rescheduling increases the likelihood of future payments. Indeed, the recent debt crisis shows the default is generally avoided by a combination of short- and long-term reschedulings of both interest and principal repayments.

Rescheduling of interest obligations leads to a buildup of debt. Over time, if unfavorable states of nature persist,⁴ this creates a situation in which the country cannot repay. If, conversely, favorable shocks occur, the country would be able

3. See Sadeq (1985) for a full exposition of the concepts developed here.

4. A "state of nature" is the complete characterization of economic conditions.

and willing to repay the full amount of the debt obligation. In our analysis, uncertainty rests with the value of future foreign exchange earnings discounted to the present. For example, unexpected changes in the real interest rate would affect the value of such earnings considerably. By explicitly incorporating the possibility of rescheduling into sovereign debt contracts, our analysis can focus on the long-run ability of the country to repay its debt.

In the following paragraphs, we will first define a measure of the assets underlying the claim on a sovereign country and then discuss the effect of alternative rescheduling policies on the debtor's behavior and consequently on the market value of debt. Finally, we will show that sovereign debt can be viewed as a contingent claim on the value of the country's current reserves and potential future trade surpluses, its "tradable value."

The Underlying Assets

In the case of a corporation, debt is essentially a prior claim on corporate assets. This stems from the legal rights of creditors to appropriate all assets in the event of bankruptcy. The "assets" of a country, however, are not transferable in any meaningful sense. For our purposes, country assets have three aspects. First, they represent the capacity of the economy to earn foreign exchange to service the debt. The total value of the assets is the amount that a foreign investor would pay in exchange for the rights to all the country's future net foreign earnings. It is given by the present discounted value of potential future trade surpluses. Second, they represent the base against which creditors, or their governments, can impose deadweight penalties in the event of repudiation. Following Sachs and Cohen (1982), it is convenient to think of the penalty as an embargo on future trade. Then, the higher the assets, the higher the penalty. Third, they include some appropriable elements, such as a country's holding of foreign exchange reserves abroad, which can be seized by a creditor upon default.

The asset value is not readily observable by banks but must be derived by applying standard valuation techniques to the stream of anticipated foreign exchange cash flows. The key determinants are the level, expected growth rate, and risk characteristics of the net foreign exchange flows. Precise derivation of the asset value requires an in-depth analysis of trade prospects and is a key variable in the creditor's decisionmaking. Note that the asset value also incorporates current reserves and therefore fully characterizes the foreign wealth of the debtor. A short-cut measure to determine this value is applied to twenty high-debt developing countries in the appendix.

The creditor's valuation is the market price of the assets, X_t , which is, in general, different from the value of the assets to the sovereign country, X_t' . The preferences of the country (such as risk aversion), the extent of credit rationing, and the correlation of earnings with domestic (nontradable) production will all clearly play a role in determining what value the country places on its assets.

This, in turn, will affect how and when the country chooses to repudiate its external debt. We will show below that it is reasonable to assume that the country will find it is in its best interest to accept the repayment condition chosen by the creditor if the latter adopts a flexible rescheduling policy. We are thus able to abstract from the important issues discussed in the preceding section and look only at the value of the claim to creditors.⁵

By construction, the asset value at any time reflects all the information then available to creditors. Consequently, the value changes only as a result of unexpected events, and it can be viewed as a random variable. For example, when exchange earnings are unexpectedly high or when the expectation of future earnings rises as a result of exogenous events (for example, a major oil discovery) the asset value rises. Note that, although the expected dollar return is always positive, changes in the asset value can be negative. That is, the asset value can fall from one period to the next, and imports can exceed exports (the trade balance can be negative). The total value of assets at any point in time, however, is always positive.⁶

We make two technical assumptions about the stochastic process which characterizes the asset value. First, the returns on the asset are intertemporally uncorrelated. Note that we do not require foreign exchange earnings to be intertemporally uncorrelated. Since expected earnings are a determinant of the asset value, we are only assuming that unexpected changes ("shocks") in the earnings stream are not forecastable from past expected changes. Second, the variance of the asset return is constant through time. This means that the variance of the change in value is proportional to the asset value: the larger the value, the larger the amplitude of the change.⁷

5. We do not consider issues of moral hazard, which may make the lender's assessment of the asset value a function of a bargaining equilibrium between borrower and lender (see Gennotte 1986).

6. A positive X is an essential requirement for a net debtor country, or else the intertemporal budget constraint will be violated.

7. These assumptions are summarized by viewing the stochastic process governing the asset price as a lognormal process. The asset value evolves as:

$$dX = X\mu dt + X\sigma dz$$

where

X = instantaneous asset value as perceived by the lender

μ = expected instantaneous rate of return

dt = time (instant)

σ = constant instant mean standard deviation of asset returns

dz = standard Wiener process

X is assumed to have a continuous sample path. The value of the asset at time $t + \tau$ is lognormally distributed,

$$X_{t+\tau} = \exp [(\mu - 1/2 \sigma^2) \tau + \int_t^{t+\tau} \sigma dz]$$

with expected value $X_t e^{\mu\tau}$:

$$E(X_{t+\tau}) = X_t e^{\mu\tau}.$$

Rescheduling Policies

The value of debt depends on the magnitude and risk characteristics of promised payments. Consequently, it depends on the rescheduling policy chosen by the lender. The optimal rescheduling policy, from the lender's point of view, is the one which maximizes the value of its claim. The lender's choice, however, is constrained because the sovereign country always has the option to repudiate unilaterally. In addition, the lender is constrained by external and internal regulations. If there is little likelihood of any payment being received because a country is unable to repay, these regulations force the lender to declare a default and to write off its loans. This forces creditors to set a limit to the country's debt relative to its asset value, above which default occurs.⁸

Debtor Behavior. The sovereign country owns its current reserves and a stream of future exchange revenues, which are valued at X_t , but has contracted a liability in the form of the promised debt payments. The strategy for the borrower is thus to maximize the value of its residual claim on the assets. Because of the risk of repudiation or default, there is a difference between the market value of the debt, C_t , and the face value, D .

At any time t , the value of the country's net external or tradable wealth is given by $X_t - C_t$, the asset value less the value of the debt. If the country repudiates its debt obligations, it retains the asset value X_t , but suffers a penalty, with present value P_t , imposed by lenders. The country would only have an incentive to repudiate if $P_t < C_t$.

If lenders will reschedule a country's debt only if the penalty they are able to impose is larger than the country's valuation of its debt, then the country will never voluntarily choose to repudiate. The choice to repudiate arises at each repayment date; there is clearly no benefit to repudiation between payments. On repudiation, the sovereign country does not make the current payment, repudiates future payments, and thus incurs the associated penalties. If the creditor chooses to reschedule all payments, however, the country is always better off because it keeps the option to repudiate later. The decision is thus not whether to repudiate but rather whether to repudiate today or in the future. By repudiating now, the debtor forgoes its options to repudiate or to make debt payments later.

Creditor Behavior. Commercial lenders to developing countries, with access to efficient worldwide capital markets, are assumed, in contrast with the sovereign country, to be able to borrow and lend risklessly at the same rate. Their objective is to maximize the market value of their debt, subject to the constraints

8. The limit can be thought of as imposed either by external regulators or by internal rules. For example, several banks have recently cleared up their Bolivian debt (at 10 cents on the U.S. dollar) to eliminate the administrative costs of continuing to hold the debt on these books.

imposed by the risk of repudiation and the risk that a country will be unable to repay its debt. A complete rescheduling policy is the choice, for any state of nature, between rescheduling, declaring a creditor to be in default, and obtaining complete payment.

In what follows, we assume that banks can impose a penalty on a country in the event of repudiation in proportion to the value of the country's gross tradable wealth, $P_t = \lambda X_t$. A portion of this penalty is a deadweight loss, however, so that the return to lenders is not equal to the size of this penalty, but only to the seizable portion of total assets, ϵX_t , where $\lambda > \epsilon$. Lenders may be better off, therefore, if they avoid repudiation through a rescheduling policy rather than permit repudiation to occur.

Consider the following example of a rescheduling policy designed to eliminate the risk of repudiation. The incentive to repudiate varies with the country's tradable wealth and the amount of the promised payment. Suppose that the yearly required payment is a given fraction of outstanding debt. When the country is heavily indebted relative to its wealth, the required payment is large relative to repudiation costs (a fraction of tradable wealth), and incentives to repudiate are strong. The bank would jeopardize future payments by requiring immediate payment and thus prefers to reschedule. If the tradable value of the country increases relative to its debt because of good fortune, the repudiation incentive diminishes, and the creditor can safely require repayment of the debt obligations. Consequently, if the bank adopts a rescheduling policy which postpones payments until the country's tradable wealth recovers from very low levels, strategic repudiation will be avoided.

This proposition does not imply that the country's ability to repudiate the debt contract unilaterally is irrelevant. In fact, if there were no penalty that lenders could impose in retaliation for repudiation, the value of debt to lenders would always be zero. We only claim that lenders are able to remove the incentive for *strategic* repudiation by adopting a flexible rescheduling policy. They are then able to take advantage of the wedge between the penalty suffered by the country on repudiation and the amount that they directly recover. Avoiding the deadweight costs of repudiation renders the debt contract self-enforcing for both the borrower and the lender.

A similar logic applies to a rescheduling policy designed to minimize the risk of bankruptcy. As the value of tradable assets relative to debt approaches zero, the value of debt to the lender tends to zero as well because the country will be unable to repay. The lender does not benefit from calling the country in default since it can appropriate only a small fraction of assets, an amount which may be less than the total debt outstanding. Thus, when a country is unable to service debt, lenders would prefer to reschedule rather than call a default. However, our analysis requires some assumption about the termination of the debt contract in case of adverse economic events. We model this in terms of the constraints on exposure faced by international banks because of internal and external regula-

tions.⁹ Such limitations prevent the bank from granting complete reschedulings perpetually. They have the same effect on the lenders' choice of rescheduling policy as the risk of voluntary repudiation by the borrower. In addition, this "automatic" bankruptcy serves the bank in the sense that it represents a credible threat to the country (especially if bankruptcy is accompanied by punitive measures) and provides an incentive for the country to comply with the creditors' rescheduling policy. Given the deadweight costs of bankruptcy, it is in the mutual interest of both borrowers and creditors to adopt a debt repayment strategy that minimizes the risk of bankruptcy actually occurring. Following this principle, the lender will agree to reschedule only part of the required payment (say, the principal portion) for lower debt-asset ratios. Finally, when the level of debt is very small relative to wealth, the lender is increasingly confident that demands for full payment can be met, and debt becomes virtually riskless.

The value of debt is thus determined as a function of the value of the underlying assets and the rescheduling policy. When the asset-debt ratio, x_t , is high, both the ability and willingness of the country to repay are high, and the debt value approaches the stream of promised payments discounted at the risk-free interest rate. At the other limit, when x_t is very low, both the ability and willingness of the country to repay are very low. Default is called to comply with regulatory authorities, and the value of debt is equal to the value of appropriable assets.

For intermediate levels of x_t , the repayment is determined by the rescheduling policy, which is designed to be self-enforcing. In essence, the strategic alternatives of the sovereign country are controlled by the lender's choice of rescheduling policy.

Debt as a Contingent Claim

In the preceding paragraphs, we have shown that sovereign debt can be viewed as a claim on the tradable wealth of the country, the payoffs of which are determined at each repayment date by the current debt level. This allows us to model sovereign debt as a contingent claim held by the lender on the borrowing country. In their seminal articles on contingent claims pricing, Black and Scholes (1973) and Merton (1973) noted that the equity and debt of a corporation can be viewed as claims, the payoffs of which are contingent on the value of the corporation. We summarize below some concepts of the contingent claims pricing literature which provide powerful tools to price sovereign debt in our framework.

Let us first think of straight corporate debt: the firm promises to repay the

9. For example, the U.S. Federal Reserve Board has categorized several country loans as "value-impaired" and has forced creditors to set aside substantial loan-loss reserves. By forcing banks to write down their assets in this manner, the regulators make it more likely that banks will try to regularize the situation by calling in their debt with the country.

amount borrowed at a specified maturity date, T . If, on that date, the firm is unable to meet its obligations (that is, the worth of the firm is less than the debt), the bondholders take control of the firm. In the absence of additional features to the contract (such as indenture provisions or interest repayments), the owners of the firm have exchanged the value of the firm against the amount borrowed and a claim which allows them to repurchase the firm on the specified maturity date at a price equal to the face value of the debt. Such a claim is equivalent to a call option on the stock of an unleveraged company.

Call options and a variety of contingent claims are actively traded on all major stock exchanges, and extensive research has been conducted on their pricing. Black and Scholes's and Merton's contributions have been to derive a theoretical model for contingent claims pricing which accurately fits the observed prices of most actively traded options. The key insight of their analysis was to establish that the short-term variations in the option price can be perfectly replicated by a portfolio of the underlying stock and riskless bonds. By continuously rebalancing the amounts invested in the two securities, investors can perfectly replicate the terminal payoffs of any option at maturity. The composition of the equivalent portfolio, and hence the contingent claim price, depend on the stock price, on its variability through time, on the interest rate, and on the promised payoffs. The promised payoffs are also called boundary conditions because the value of the claim at any date, t , is the solution of a differential equation which is equal to the promised payments at maturity date T .

For a given rescheduling policy, similar reasoning will lead to the pricing of sovereign debt developed below. The inherent difficulty of debt pricing when reschedulings are considered is that the maturity of debt is itself a stochastic variable. Indeed, rescheduling may be repeated indefinitely, and the claim potentially has an infinite time to maturity. This is in agreement with observation of actual international debt contracts, which have de facto become revolving credit agreements. As is extensively developed in the next section, we solve this problem by viewing debt pricing as a contract which is periodic in the sense that the same set of repayment decisions is faced in every future period. Once the value of debt to the lender is determined, contingent on a rescheduling policy, the derivation of the optimal rescheduling policy and of the associated value of debt reduces to optimizing the value of debt over possible rescheduling policies.

II. THE DEBT VALUATION MODEL

In this section, we first outline the assumptions of the analysis and discuss their economic meaning. Next, we present the model of international lending with endogenous reschedulings. This requires a specification of the boundary conditions and of a valuation process for an infinite maturity claim.¹⁰

10. The method for deriving a numerical solution for the value of the debt and the optimal rescheduling policy is outlined in the appendix.

Assumptions

A sovereign country has an initial stock of debt and an associated stream of repayment obligations comprised of interest and principal. Both are assumed to be a constant fraction of the stock of debt and are denoted I and A_d , respectively. Debt payments come from the foreign exchange earnings of the country. The value of these is summarized in terms of a hypothetical asset. Debt is a claim held by the lender on this asset. The debt contract specifies the rule that determines the sharing of the earnings from the asset between creditor and debtor.

In the initial situation, the country has just made a debt payment. Prior to the next payment, the asset value changes according to an exogenous stochastic process. The creditor then makes a decision whether to demand full repayment of the obligations, interest only,¹¹ or no repayment. The country then chooses whether or not to make the requested payment. As discussed above, because the lender is assumed to know the country's willingness and ability to repay, and thus sets the rescheduling conditions accordingly, accepting the rescheduling offer is always a dominant strategy for the country. Because both the incentive to repudiate and the ability to repay depend on the same argument, the asset-debt ratio, x , only one will be binding on the rescheduling policy. In the development of the model below, we assume that it is the latter. This sequence continues indefinitely; it is halted only if the lender is forced, by its regulatory authorities, to call the country in default. If the cash flow from the country's asset is insufficient to service the debt, we assume that the country is able to "sell off" its asset to meet the required payment; this can be conceived as running down foreign exchange reserves or as obtaining financing from official creditors (the International Monetary Fund or World Bank) with claims on the country's assets senior to those of the commercial lender.

A Model of International Debt with Reschedulings

In the following paragraphs, we describe our formulation of the valuation problem, which can only be solved numerically. We rely on important results of the contingent claims literature (Black and Scholes 1973; Merton 1973; Cox, Ross, and Rubinstein 1979). Essentially, the instantaneous variation in the value of the debt is a function of the variations in the value of the asset. A continuous-time differential equation specifies this relationship. Our assumption on the set of possible rescheduling policies determines the boundary conditions of the numerical problem. The equation is integrated numerically.

Exogenous Boundary Conditions. Let D be the face value of debt, the amount borrowed by the sovereign country, and C the market value. There are two cases in which C is perfectly determined: (i) when the country has such a

11. In theory, the amount rescheduled could be anything between zero and the full amount of the obligations due. For simplicity, we assume that the decision is limited to rescheduling both interest plus principal or principal alone.

high asset-debt ratio, x , that debt becomes riskless, and (ii) when it has such a low asset-debt ratio that the bank is forced to declare default.

As $x \rightarrow \infty$, debt becomes riskless. This will translate into an arbitrary numerical ($\bar{x}$) limit beyond which it is assumed that the country will always be required to (and be able to) pay off its debt in full. C then attains its maximum, which is equal to the value of all promised payments on D discounted at the riskless interest rate. $C^{\max}(D)$ is exogenously specified in this model. Since repayments are a linear function of D and the present value operator is linear,

$$(1) \quad C^{\max}(D) = \alpha D$$

where α depends on both the risk-free rate and the rate charged to the country (that is, the rate including the margin over the London Interbank Offer Rate).

The bank faces a regulatory lending constraint for exposure to any one country. This can be assumed to emanate from the central bank or the depository insurance agency or to be exogenously set by the bank itself. This level is denoted by $D^{\max}$. Statutory and internal bankruptcy limits are set on the basis of the value of export earning assets and are not allowed to reach a level greater than some constant, b , times the asset value: $D^{\max} = bX$. The minimum asset-debt ratio is thus set as: $x^{\min} = X/D^{\max} = X/bX = 1/b$. When D reaches the full value, bX , regulatory bankruptcy occurs, and the bank appropriates a fraction, ϵ , of the total asset value. Thus the market value of the debt when D reaches its maximum is the seizable fraction of total assets:

$$C(D) = \epsilon X, \text{ when } D \geq bX$$

Now consider that the model thus far specified includes two state variables: X and D . Such models are inherently very difficult to solve numerically, and non-linearity can be a serious problem. From an economically appealing point of view, the value of debt actually depends on only one state variable, x , the ratio of asset value to debt. If we double the value of the asset and the face value of the outstanding debt, then the market value of the debt should also be doubled. If we denote $c = C/D$, the ratio of the market value to the face value of the debt, the value of c can be specified as:

$$(2) \quad c = c(x, t); c_x > 0$$

where

$$(3) \quad c = \epsilon x \text{ for } x \leq 1/b$$

and

$$(4) \quad c = \alpha \text{ for } x \geq \bar{x}$$

Equation 2 states that as the asset-debt ratio, x , rises, c also rises. At very low levels of x (below the default threshold, $1/b$), the value of c is exogenously given by the amount of the asset that can be seized by lenders, ϵx (equation 3). This is shown as the line $\overline{OA}$ in figure 1. Conversely, for very high levels of x (above $\bar{x}$, which is our finite proxy for infinity), the value of c equals its asymptotic limit, α

Figure 1. *Exogenous and Endogenous Boundary Conditions*

(equation 4). The value of c for interim levels of x , $1/b < x < \bar{x}$, is driven by the probability of receiving debt-service payments or of x reaching $\bar{x}$ or $1/b$ over time.

To simplify the numerical analysis, we do not actually derive the cost of debt service as valued by the borrower, or the penalty, P_t . Given this, to ensure that repudiation is never chosen by the country, we require that the market value of debt be no greater than the penalty to the country, $C_t \leq P_t$, or equivalently, $c \leq \lambda x$. We show below, however, that the constraint can always be satisfied by an appropriate choice of b , the limit on indebtedness set by the regulatory authorities. In figure 1, the line $c = \lambda x$ provides an upper bound to the value of c , because for $c > \lambda x$, the country would choose to repudiate and lenders get only ϵx per dollar of debt outstanding.

Endogenously determined rescheduling boundaries. So far, we have determined the value of debt, c , for two limit values of the ratio x : when it is very large and when it is very small. In all intermediate cases we assume that the bank has three alternative courses of action:

- Ask for a full repayment of the interest, I_t , plus the principal due, A_t^d . In this case the debt is reduced to:

$$D_t = D_{t-1} - A_{t-1}^d$$

- Require the country to make only the interest payment on the debt, that is, R . The debt level remains unchanged. If the bank charges a fee, F_1 , for this rescheduling, it is added to the face value of the debt:

$$D_t = D_{t-1} + F_1$$

- Receive no payment at all. In this case the level of debt increases by the forgone interest plus a second type of rescheduling fee, F_2 :

$$D_t = D_{t-1} + I_{t-1} + F_2$$

If, in addition, the creditor decides to lend some additional funds, L , the debt becomes:

$$D_t = D_{t-1} + I_{t-1} + F_2 + L_t$$

In each case, we assume that principal (A), interest (I), fees (F), and additional loans (L), and thus the repayment due at any point in time, can be calculated as some percentage of the remaining face value of the debt. This implies that even if the country makes the full repayment due each year, the debt is never completely repaid but the remaining debt progressively tends toward zero.

The optimal policy for the bank is reduced to the choice of the two rescheduling boundaries, x_1 and x_2 , in figure 1, which differentiate the request for payment from the country: $x \geq x_1$ when the bank requests full payment; $x_1 > x \geq x_2$ when the bank requires only the interest payment; and $1/b < x < x_2$ when the bank asks for no payment at all and may even make new funds available to the country. An important point is that the rescheduling boundaries are func-

tions of x alone rather than of X and D independently. This follows from the fact that the two exogenous limit boundaries can be written as functions of x , and that we have constant returns to scale.

We have restricted the bank's choice to three alternatives to simplify the argument, but the model could accommodate any number of alternative repayments. At every repayment date, the creditor decides on which repayment to require from the borrower. If the creditor could appropriate the full value of the productive assets in case of repudiation or default, it would be in its best interest to require full payment in all cases. As the debtor is a sovereign country, however, only the fraction, ϵ , of the value of the asset is appropriable in the event of repudiation or default. Hence the creditor will select a rescheduling policy which reduces the likelihood of repudiation or default in the future. The trade-off is thus between increased expected costs associated with default and a higher repayment at the present time. As the asset value increases, the incentive for the creditor to reschedule is lessened.

The Valuation Process

The value of debt at any point in time is determined by the value of the asset, the face value of debt, the rescheduling policy at all times in the future, and a set of parameters (the interest rates faced by the country and by the bank, the fraction of debt repaid as principal, and so forth). Once the function $c = c(x)$ is determined, the "market" price of debt, c , can be simply located on the vertical axis in figure 1, given any value for x .

One difficulty that is faced is the potential infinite maturity of the debt claim given rescheduling possibilities. This is overcome by observing that the valuation problem can be made stationary, or, to be more accurate, periodic. The model assumes the distribution of the asset value to be stationary, and the repayment schedule as well as the exogenous boundary conditions to be independent of time. Consequently, for a given level of x , c is independent of time. The problem does not exhibit perfect stationarity because the relative value of debt between repayment dates depends on the time left to the next repayment; but it is periodic.

The periodicity property is essential to our valuation method in that it closes the system of equations of the problem. The classical differential equation of contingent claims pricing specifies the value of a claim as a function of its payoffs at maturity. This is done by valuing the claim just before maturity and then recurrently deriving its price at earlier points in time. As the maturity of the debt is infinite, this method does not apply. We can still derive the relationship between the value of the claim at the present repayment date and its value at the next repayment date, however, contingent on a given rescheduling policy. The value of the claim one year from now is only known for the exogenous boundaries, but for a given level of X/D it is the same as the current value. This gives us a well-specified differential equation, which becomes a system of equations in the numerical approximation. Once the value of the claim is determined for any given rescheduling policy, summarized by the choice of the boundary values x_1

and x_2 , the optimal rescheduling strategy is found by maximizing the claim value over all possible policies. This last step yields both the optimal rescheduling boundaries and the value of debt to a fully rational lender.

III. SIMULATION RESULTS: THE MARKET VALUE OF DEBT

Based on the debt valuation model described above, simulations were run which allow us to compare the numerical values and the bank's optimal rescheduling policies obtained when we vary any one of the major parameters. The changes in the upper and lower boundary values indicate the resulting alteration in the debt values at which full payment or full rescheduling are optimal. From these simulations, it appears that the bank benefits substantially from voluntary rescheduling in realistic ranges of the parameters, in contrast to the old proverb "a bird in the hand is worth two in the bush." The expected costs of default are high enough for the bank to prefer to forgo part of an immediate payment even if the sovereign country is able to repay.

We review here the determinants of debt value and examine the comparative statics of the model for the ten parameters. Debt values are derived for 100 different levels of the reduced variable X/D , which range from 0.10 to 50 and thus include most realistic cases.¹²

The determinants of debt value are listed in table 1, along with the levels we used for our base case. Figures 2 through 7 below present the simulation results for different values of the parameters. Each simulation differs from the base case only in the value of one parameter. Table 4 in the appendix summarizes our numerical results.

The Base Case

The market-to-face value of debt for the base parameter values given in table 1 is shown in figure 2 for a range of x values, the asset-debt ratios. This ratio measures the indebtedness of the country and, consequently, the likelihood that the promised payments will be made. The value of debt is a linear function of x at default and a concave function of $x > 1/b$ (see figure 1). In case of default, the market to face value of the debt, c , is small (1.6 percent of face value at the default boundary in the base case). It jumps upward, however, if default is not called. On the brink of default, c equals 10 percent. This demonstrates the value to the lender of keeping alive the game by rescheduling rather than calling default.

The maximum value of c is reached asymptotically: as the asset value increases, the likelihood of full repayment increases until debt becomes virtually riskless.¹³ The debt is worth its face value, and thus $c = 1$ when the asset value is

12. These values can be compared with those estimated for twenty developing countries in the appendix, which were calculated using a short-cut asset valuation technique.

13. At the limit of very large values of x , debt becomes a perpetuity. It is easy to show that its value is $(A + R)/(A + R^*)$ times face value, that is, a debt value-face value ratio of 1.1 in the base case.

Table 1. *Debt Valuation Parameters: Base Case Values*

Parameter	Representative value (percent)		Figure showing simulation results
	Base case	Sensitivity	
Endogenous			
x_1 Upper rescheduling boundary for X/D (full repayment required for $x \geq x_1$)	108		3
x_2 Lower rescheduling boundary for X/D (interest payment required for $x_1 > x \geq x_2$)	49		3
Exogenous			
R Interest rate charged to the country	13	11, 15	4
R^* Riskless rate paid by bank	10	12	6
A^d Fraction of principal repaid per year	20	25, 50, 100	5
σ^2 Yearly variance of asset value	16	50	—
$1/b$ Boundary value of X/D for default	35	55, 100	7
ϵ Appropriable fraction of assets	5	10	—
F Rescheduling fee (percent of face value)	0	1	—
L Funding by bank (percent of face value)	0	5	—

— Not applicable.

approximately double the face value and $x = 2$. This is the credit rationing boundary. At this level of debt, a new lender that voluntarily extended additional funds, driving down x , would obtain a claim that would be worth less than the amount lent despite the risk premium built into the interest rate charged. The difference between full value (the value of debt if it were riskless) and face value measures the riskiness of debt (10 percent at the credit rationing boundary).

Creditor Policies

The value of debt is sensitive to changes in the rescheduling policy chosen by the lender. The rescheduling policy specifies the type of repayment requested by the lender as a function of the tradable wealth of the borrowing country. It is summarized by the parameters x_1 and x_2 in figure 1. If the bank is more "impatient," that is, if it starts requesting full or partial payment for a lower level of x , it increases the risk of financial distress and the expected probability of default.

Our numerical analysis yields the optimal rescheduling policy and the value of debt for that optimal policy and for alternate (suboptimal) policies. The optimal rescheduling strategy in the base case is for the lender to request full repayment until the asset-debt ratio, x , falls below 61 percent. Below this level, continued regular debt service would be likely to quickly trigger bankruptcy, and future debt service payments would then be jeopardized. In this case, the lenders would only request payment of interest due. When x falls below 48 percent, the lender is willing to reschedule both principal and interest. These optimal rescheduling boundaries are clearly sensitive to the level of x at which default is called. For

Figure 2. *Debt Valuation: The Base Case*

example, when defaults occur at $x = 100$ percent, optimal rescheduling occurs at $x = 127$ percent. In this case, both interest and principal are always rescheduled together.

Figure 3 also demonstrates alternative rescheduling policies. Bond financing would not permit any recontracting. It is clear that this sharply reduces the value of debt. Just prior to default, debt in the form of bonds would only be worth 60 percent of debt in the form of a credit that could be rescheduled. As the country gets wealthier, this premium shrinks toward zero when debt is riskless. Similarly, the credit rationing boundary is restricted under bond financing. The value of x for which $c = 1.0$ is about 10 percent higher than in the base case for credits; that is, bond lenders would cut off credit at lower levels of debt. Again, the numerical values are sensitive to the assumption about the size of x when default is called. The higher this is, the greater is the advantage of credit financing over bond financing. Figure 3 also illustrates two ad hoc rescheduling policies: creditors may be too impatient, ask for full payment, and thereby increase the probability of default; or they may be too lax and reschedule debt when asset levels would allow full payment. In both cases, the value of debt declines relative to an optimal rescheduling policy.

The interest rate charged to the country (R), together with the principal (A), determines the size of the repayments requested by the bank at each year end. The difference between R and the riskless rate, R^* , at which the bank can borrow in financial markets, compensates the bank for the riskiness of the loan. The "fair" interest rate R is that for which $c = 1$; that is, the market value equals the face value of debt. This is clearly a function of x . Figure 4 shows that in a competitive capital market, the spread charged to the country will be an increasing function of the long-run debt-asset ratio. Our simulations indicate that a 1 percentage point increase in the spread would support an 8 percent increase in debt.

Although higher levels of debt can be sustained with a higher spread, figure 4 shows that the spread is not an effective instrument with which lenders can compensate for ex post excessive debt; regardless of the spread, default is called at the same level of x . A high spread has two offsetting features. On the one hand, it raises profits to the lender when debt is risk-free. On the other hand, it raises the probability of bankruptcy or default. With the higher spread, the rescheduling boundary is raised relative to the base case. When debt is sufficiently risky that a rescheduling of interest is optimal ($x = 0.67$ for a 5 percentage point spread), these two forces balance each other. The value of debt to creditors is actually raised by lowering the spread. This explains why actual reschedulings have rarely involved an increase in spreads. The additional risk of default outweighs the additional current income.

There are similar offsetting factors governing the maturity of debt. In figure 5, we show how c varies as the proportion of debt to be repaid each year is increased. When debt is riskless, the maturity does not really matter, and c is the same for both levels of repayment. Longer-term debt is preferred by creditors

Figure 3. *Rescheduling Policy: Optimal versus Suboptimal Strategies*

Figure 4. *Interest Rate Spreads*

because they lock in higher total interest payments. As x falls and debt becomes more risky, lenders are keen to shorten maturities and recover their principal quickly. The shorter-maturity debt ($A = 25$ percent) has a greater value than the longer-maturity debt over the asset range from $x = 0.48$ to $x = 2.6$. Below this level, however, attempts by creditors to demand full repayment would precipitate bankruptcy more quickly. There is an incentive to lengthen maturities. In our simulations, we find that this switching point coincides approximately with the full rescheduling boundary, x_2 in figure 1. Thus, the model explains why short-term debt dominates as debt starts to become risky but is then rescheduled into longer-term maturities when the situation further deteriorates.

Front-end fees charged at rescheduling do not have the same intertemporal effects as raising spreads or shortening maturities. Our simulations reveal that the unit value of debt is unchanged by adding a 1 percent fee. The total portfolio of the creditor is increased, therefore, because the fee is assumed to be capitalized into the stock of debt. Fees are extremely flexible instruments which allow

Figure 5. *Principal Repayment: Fraction of Face Value Repaid Each Year*

the creditor to extract the maximum possible from the debtor but do not force the debtor to repudiate and do not unduly increase the risk of bankruptcy.

Finally, the lender may also provide new money to a country in financial distress to help stave off the risk of default. The promise by a lender to provide an additional 5 percent of debt in new money as well as rescheduling all payments due is shown by our simulations to be able to increase the value of the lender's portfolio. This is not, however, a time-consistent policy. When a country actually has severe financial problems, new money can help raise the value of old debt, but it does not do so by enough to compensate for the greater volume of debt that is then at risk.

Global Conditions

Our final set of sensitivity runs deals with the impact of changing global conditions; the riskless interest rate, the variance of the asset, the level of the default boundary, and the proportion of assets seized by banks upon default.

For any given level of x , a higher riskless interest rate lowers the value of debt because it raises the probability of default without increasing the value of debt if full repayment occurs. Figure 6 shows a rise of 2 percentage points in the riskless interest rate and in the rate charged to borrowers. The credit rationing boundary at which the face value equals the market value of debt shrinks by nearly 10 percent for a given level of x . In fact, the impact on nominal debt will be substantially higher, because the value of the asset will be significantly affected by a change in the interest rate. If a country's discounted earnings were worth 200 at $R^* = 10$ percent, it could take on a debt of 100 ($x = 2$). As R^* rose to 12 percent, however, the discounted value of earnings would fall (to, say, 160), while the rationing ceiling would also fall ($x = 2.2$). Thus, the total nominal debt ceiling that would be imposed by lenders would fall from 100 to 73. The simulation confirms the extreme importance of the level of the interest rate in determining country creditworthiness.

As would be expected, an increase in the asset variance from the base case level of 0.16 to 0.5 also causes the value of debt to fall.¹⁴ In a highly uncertain world, country creditworthiness will decrease.

If the regulatory authorities tighten limits on lending relative to asset value by calling bank loans value-impaired at an earlier stage (a higher x), the impact on the value of debt is dramatic, even for cases in which the probability of default is low. Figure 7 shows the impact of raising the ceiling from $x^{\min} = 0.35$ ($b = 2.8$) to $x^{\min} = 0.5$ ($b = 2$) and $x^{\min} = 1.0$ ($b = 1$). The most obvious advantage of the lower ceiling is evident at low levels of x , at which it is in the interests of the creditor and the borrower to continue to reschedule for as long as possible. What is perhaps more surprising is the strength and persistence of the effect at lower debt levels. If default is called at $b = 1.0$, then the credit rationing

14. A value of 0.5 is typical for a U.S. corporation in the petroleum, chemicals, or metals and mining industries.

Figure 6. *Interbank Interest Rates*

Figure 7. *Default Boundary*

boundary (when $c = 1$) is shifted out from $x = 2$ to $x = 3$. That is, the maximum level of debt that would be voluntarily lent to a country, holding the spread constant, is lowered by one-third. It is clear why regulatory authorities have been reluctant to enforce stringent controls on international lending. It is also clear that a tougher regulatory environment would seriously limit the volume of debt that commercial lenders would be prepared to extend.

The final sensitivity analysis consists of raising the share of assets appropriated by lenders at default from 5 to 10 percent. This has almost no impact on debt value. The intuition is simple. At default, the asset value is very low, as is the amount that lenders recover. In order to raise the value of debt, it is more important to change the probability of reaching the default boundary than to change the payoff at that boundary.

IV. CONCLUSIONS AND INFERENCES

Creditors have little legal recourse if a sovereign state repudiates its external debt obligations. There are, however, deadweight losses to the system as a whole that would arise in the case of a default. These may be thought of as the loss of mutually profitable opportunities for future loans, or as the imposition of efficiency reducing sanctions on international trade. It is in the interest of both parties to recontract to avoid such losses. The potential for rescheduling payments due on syndicated credits permits lenders to structure the incentives faced by borrowers in such a way that repudiation is never a rational action. In our framework, we show that rescheduling of principal, and sometimes of interest due, is in the best interests of both creditors and debtors as they strive to avoid damaging the country's asset base, on which future debt servicing capability depends. That is to say, reschedulings can free resources to add to investment, making future debt service payments more likely.

We have developed a numerical method for valuing developing-country debt. Debt is viewed as a contingent claim on underlying assets that have a stochastic value, which implies that the net payoff to repudiation is stochastic. This is the major source of risk in international lending, and it cannot be completely controlled by credit rationing. As the value of the asset-debt ratio falls, the "market" value of debt falls below the face value at an increasing rate.

Our numerical results are broadly consistent with certain stylized characteristics of international credit markets. The point at which the market value of debt falls below the face value is critical, for example, as it indicates when voluntary lending will be completely shut off. Under the assumptions of our simulation analysis, this occurs at an asset-debt ratio of 2 for an average developing country with a traded sector incremental capital-output ratio of 4. The cutoff point for new lending is shown to be sensitive to international conditions. This helps explain why new lending was suddenly curtailed in a period of high real interest rates and low growth prospects for developing countries—factors likely to

sharply reduce the value of the assets underlying the debt. Conversely, the methodology could be used to estimate by how much interest rates would have to fall to restore creditworthiness to any particular country.

A rough-and-ready application of the methodology to estimate the market value of debt relative to the face value (the price) of debt for twenty developing countries at the end of 1983 is given in the appendix. In all cases, the estimated price is less than unity. This suggests that these countries had reached their commercial credit limits. Additional lending would be forthcoming only in amounts that would leave the asset-debt ratio unchanged. That is, the growth of asset value would determine the feasible growth of debt. Eight of the countries fall below the rescheduling boundary. These countries have "solvency" problems, in the sense that they may be unable to repay debt service if further adverse shocks occur. According to our analysis, creditors would willingly defer the receipt of payments from these countries in order to avoid damaging the asset base on which future debt payments depend. The other countries seemed to be faced with classical liquidity problems: no new lending will be made available because the debt price had fallen below unity, and yet they face the need to pay interest and principal on a large outstanding debt stock.

In the appendix, countries are also ranked according to their debt price; this gives a creditworthiness ordering that corresponds closely with rankings from other sources. A more detailed application of our methodology offers the potential for improving creditworthiness orderings, and for assessing the degree of adjustment required to regain full resumption of spontaneous lending.

There are several actions that lenders can take to raise the value of debt they hold. Foremost among these is the potential to reschedule debt service obligations. This gives a clear advantage to syndicated credits over other instruments such as bonds, for which reschedulings are more difficult because of the anonymous and dispersed nature of ownership. When the asset-to-debt value of a country is low, we find that credits have a greater value than bonds, particularly during periods when adverse shocks are common and debt becomes more risky. This higher valuation also permits a greater volume of credits to be extended for any given asset value. Second, it seems that shortening maturities is always in the interest of lenders when debt is slightly risky, but it adds to the risk of default as the degree of financial distress gets larger. Similarly, we see that raising the interest spread has a mixed effect on debt values. If debt is relatively safe, higher spreads add to profits. But during a rescheduling, when debt is risky, higher spreads increase the probability of a default. The net benefits to lenders are small or negative over a broad region. In fact, lenders may even gain by reducing spreads somewhat. There are, however, no such adverse effects with respect to front-end fees. These are almost pure profit for the banks. It is not surprising, therefore, that actual reschedulings have been characterized by a lengthening of maturities, relatively minor adjustments to spreads (and sometimes even a narrowing), and heavy fees. Finally, we show that the precommitment to provide additional funds during a rescheduling can raise the value of existing debt. These

should not be construed as "concessions" on the part of commercial lenders. Rather, they are safeguards to protect the value of existing debt.

Our results can also shed light on the trade-off involved in some current proposals for international actions to alleviate the debt crisis. The development of a secondary market in developing-country debt could provide lenders with extra liquidity, but at the cost of a reduction in debt value if the ability to reschedule was lessened as a result of such a market. Limiting the number of debt holders seems essential to offset the threat of repudiation. Thus, it is unclear whether lenders would gain from further development of secondary markets. Developing countries, however, would probably lose in the long run from such markets. They would then be faced with a financial market similar to the bond market, to which they have had relatively little access in the postwar period because of the particular risk characteristics of lending to them. Calls for tougher supervision by regulatory authorities could also be damaging to international lending. The longer lenders can stave off a repudiation by capitalizing interest into debt in dire times, the greater the value of the debt. Finally, tougher policies regarding the seizure of assets in the event of a repudiation, even if feasible, would not be in the interest of lenders. When defaults are called, the value of such assets is typically very small and hardly protects the lender from risk.

APPENDIX

Numerical Solution Method

Given the periodicity of the problem, we need only derive the relationship between the values of debt at two successive repayment dates, T and $T + 1$. Suppose that the relative value of debt at the next repayment date c_{T+1} , is known for each possible asset-to-debt value, x_{T+1} . To evaluate c_T , we will derive the value of debt at the instant just prior to $T + 1$ and recurrently move backward to solve it at all prior times, $T < t < (T + 1)$. This yields the value of debt $c_T(x_T)$ at the beginning of the period as a linear function of all possible values at the end of the period, $c_{T+1}(x_{T+1})$. The periodicity argument implies that the two functions c_T and c_{T+1} are equal for any given value of the ratio, x . This reasoning gives us a well-specified equation which we solve numerically.

A variety of numerical methods can be used. We adopted the binomial method of Cox, Ross, and Rubinstein (1979), which provides a clear, economic understanding of the process. The continuous-time, lognormal distribution of x_t can be approximated by a series of independent multiplicative binomial jumps over discrete intervals. Divide the one-year period between two repayments into N equal intervals with length $\tau = 1/N$. Over each interval, the value of x_t can jump to two different values: a higher one, $x_{t+\tau} = u \cdot x_t (u > 1)$, or a lower one, $x_{t-\tau} = d \cdot x_t (d < 1)$. This probability distribution is discrete in the sense that the set of possible values taken by the random variable, x_b , is a finite set. Furthermore, it is defined only at discrete points in time. This distribution converges to the

lognormal distribution (with annual variance σ^2) when the number N of discrete intervals tends toward infinity and u and d are given by:

$$u = \exp(\sigma/\sqrt{\tau}) \quad d = \exp(-\sigma/\sqrt{\tau})$$

Using this approximation, we need to determine the value of the claim at all "nodes" of the "grid" defined by the possible values of x at times $T + n \cdot \tau$ ($n = 0, 1, \dots, N$). In the simulations, we adopted the value $N = 100$, which yields a good approximation of the lognormal process.

In this discrete framework, we can use the arbitrage argument (due to Merton 1973) to derive the value of the claim at time t as a function of its value at time $t + \tau$. Consider a portfolio composed of the contingent claim and a share, h , of the asset, x_t . The value of such a portfolio at time $t + \tau$ is:

$$c_{t+\tau}(u \cdot x_t) + h \cdot (u \cdot x_t) \quad \text{or} \quad c_{t+\tau}(d \cdot x_t) + h \cdot (d \cdot x_t)$$

If we choose the value $h = [c_{t+\tau}(d \cdot x_t) - c_{t+\tau}(u \cdot x_t)]/[u - d]x_t$, inspection of that equation shows that the value of the portfolio is the same in both cases. The return on investment in this portfolio is risk-free. It should thus be equal to the riskless rate, r , available to the bank in order to preclude arbitrage opportunities. We thus have the relationship:

$$c_t(x_t) = [p c_{t+\tau}(u \cdot x_t) + (1 - p) c_{t+\tau}(d \cdot x_t)]/r$$

where

$$p = (r - d)/(u - d), \text{ and } 0 < p < 1$$

We interpret p as the "quasi-probability" of an upward move. The value of debt at time t then becomes the expected value (for the quasi-probability p) of debt at time $t + \tau$ discounted at the riskless rate of interest. The expression for $c_t(x_t)$ depends in a linear fashion on the two possible values of $c_{t+\tau}$. This derivation can be iterated backward to determine the (linear) expression of c_t as a function of all possible values of c at time $T + 1$. The variable c_T can thus be interpreted as the present value of c_{T+1} associated with a series of multiplicative independent binomial processes defined by u , d , and the quasi-probability, p .

As noted before, the recursive relationship is independent of the true probability of an upward move. This is because movements in the value of debt can be perfectly hedged by a position in assets, the size of which depends only on the amplitude of changes in the asset price.

Recall that the underlying asset (the foreign exchange value) of the country is not traded and its value must be derived. Our valuation procedure does not require a perfect duplication of changes in the asset price and changes in the value of traded assets which are effectively priced by the market. In a capital-asset pricing model framework, for example, only the nondiversifiable risk is priced. Thus the valuation method is very general. Furthermore, it allows the lender to offset the risk of its claim by selling traded assets, which are highly correlated with the country value.

Repayment Equations

So far, we determined the value of the claim at time T as a function of all possible values of the claim at the instant prior to the next repayment date, $T + 1$. When repayment occurs, the variable x decreases or increases as specified by the following equations.

Normal repayment ($x \geq x_1$):

$$x \rightarrow (x - A - R)/(1 - A)$$

Interest-only payment ($x_2 \leq x < x_1$):

$$x \rightarrow (x - R)/(1 + F_1)$$

No repayment ($1/b < x < x_1$):

$$x \rightarrow x/(1 + R + F_2)$$

These equations are linear; consequently, the value of the claim at date T (after repayment) is a linear function of the value of the claim at time $T + 1$ (after repayment). The equations can be formalized in matrix notation to an equality between a vector c_T and the product of a matrix Ω by a vector c_{T+1} , where the i th component of c_T is the value of debt for the i th discrete value of the state variable x . Since the two vectors c_T and c_{T+1} are equal, the numerical problem reduces to:

$$c_T = \Omega \cdot c_T$$

The elements of the matrix Ω are known functions of the fixed parameters and of the rescheduling policy characterized by x_1 and x_2 . The optimal rescheduling policy and the value of debt associated with it are then obtained by maximizing the solution to the matrix equation over all possible rescheduling policies.

An Empirical Application of the Valuation Model

We have calculated the implicit value of debt for twenty developing countries at the end of 1983. The process requires evaluation of the country's underlying assets. The procedure used is as follows. We first took the total replacement value of the capital stock of each country by summing investment over the preceding fifteen years and assuming a 10 percent depreciation rate. The initial capital stock in 1968 was taken to be four times gross domestic product (GDP) of that year (four represents a "normal" capital-output ratio for a developing country). This is given in column 1 of table 2. We then make two adjustments to this figure to give a measure of the actual value of capital in the traded sector. The first converts the replacement value into the actual value. Because countries with low incremental capital-output ratios (ICORs) have many profitable investment opportunities remaining, the actual value of their firm exceeds the replacement value. Accordingly, we multiply the replacement value by $4/\text{ICOR}$. For the average developing country with an ICOR of 4, the actual value coincides with the replacement value. The second adjustment to obtain the share of capital in the

Table 2. *Valuation of Tradable Assets, 1983*
(millions of U.S. dollars)

Country	Replacement capital ^a (1)	ICORs ^b (2)	Tradables share ^c (3)	Value of traded capital (4) = (1) • [(4)/(2)] • (3)	Total reserves minus gold ^d (5)	Total assets (6) = (4) + (5)
1. Egypt, Arab Rep.	72,629	2.7	0.249	26,792.0	771.0	27,563.0
2. Venezuela	205,076	5.9	0.388	53,945.4	7,643.0	61,588.4
3. Thailand	81,567	3.5	0.208	19,390.6	1,607.0	20,996.6
4. Korea, Rep.	150,089	3.1	0.283	54,806.5	2,346.7	57,153.2
5. Malaysia	60,288	2.8	0.264	22,737.4	3,784.0	26,521.4
6. Yugoslavia	130,306	5.5	0.303	28,715.0	976.0	29,691.0
7. Cyprus	5,335	5.7	0.193	722.5	519.1	1,241.6
8. Colombia	69,372	3.5	0.226	23,309.0	1,901.0	25,210.0
9. Mexico	343,861	3.5	0.278	109,249.7	3,913.0	113,163.0
10. Brazil	425,265	4.2	0.271	109,759.0	4,355.0	114,114.0
11. Uruguay	10,461	3.2	0.269	3,517.6	207.0	3,724.6
12. Ecuador	28,336	4.1	0.282	7,796.0	644.5	8,440.5
13. Chile	44,291	4.1	0.294	12,704.0	2,036.3	14,740.3
14. Philippines	84,050	4.5	0.270	20,172.0	786.0	20,958.0
15. Peru	37,370	7.5	0.374	7,454.1	1,365.1	8,819.2
16. Bolivia	18,533	6.0	0.235	2,903.4	136.6	3,040.0
17. Argentina	163,565	11.3	0.286	16,559.1	1,170.0	17,731.1
18. Turkey	102,353	7.8	0.224	11,757.4	1,270.0	13,027.4
19. Costa Rica	11,511	4.9	0.192	1,804.2	311.2	2,115.4
20. Nicaragua	7,324	-3.9	0.236	0.	184.3	184.3

Note: Countries are listed in order of the estimated price of a dollar of commercial debt as shown in table 3, column 6.

a. Sum of real investment (1983 prices) for 1968–83, discounted at 10 percent, plus four times real 1968 GDP discounted at the same rate.

b. Incremental capital-output ratios. These are for the period 1975–80, with the following exceptions: Chile (1974–80), Cyprus (1974–80), Korea (1975–79), Turkey (1975–81), and Venezuela (1975–79).

c. (Mining and manufacturing)/GDP, average for 1970–81, at current prices.

d. Line 1.1.d, IMF, *International Financial Statistics* (July 1985).

Sources: World Bank national accounts data; IMF, *International Financial Statistics* (July 1985).

traded sector is calculated by multiplying the total capital stock by the average share of mining plus manufacturing in GDP. These sectors are chosen as they are likely to be worst hit by an international embargo and thus may be considered by creditors to be effective "collateral" on the country's debt. Production of agricultural exports could be quickly converted into production of domestic consumption goods. The result of these adjustments gives the value of traded-sector capital, shown in column 4 of table 2. As the final outcome of this process, the total traded assets of the country (column 6) are obtained by adding total reserves minus gold to the value of traded capital.

We present two versions of the asset-debt ratio in table 3. In practice, countries have official debt, which is serviced prior to commercial obligations. For example, World Bank and IMF loans have never been rescheduled. Paris Club reschedulings on other official loans are conditional on similar treatment being accorded commercial lenders, and thus we account for these as two debt categories. Values based on the first definition, given in column 4 of table 3, shows total assets over total debt. In the second variant, column 5, we first subtract out official debt from total assets, to give a measure of the net assets on which commercial lenders have a claim, and divide this by commercial debt.

Given the appropriate asset-debt ratio for a selected country, the estimated price of a dollar of commercial debt can be obtained from table 4. The first column gives the range of possible asset-debt ratios. The relevant price of debt under the base case assumptions is then read from the corresponding row of the second column. The additional columns give the debt price for the various simulations in which, for each column, the value of one of the nine determinants is altered.

The estimated prices can be used in two ways. First, for some developing countries there exists information on the value of Eurocredits implicit in swap arrangements undertaken by commercial banks. For example, the *New York Times* (June 1985) gives the following quotes: Bolivia, 0.20; Peru, 0.50; Argentina, 0.70; Venezuela, 0.90. Other sources suggest Brazilian and Mexican debt trades for about 0.80–0.85 cents on the U.S. dollar. A comparison of these prices with our crude estimates shows a close correspondence. We take this as a priori evidence that a model such as the one described above could be made operational.

The second test of the relevance of the estimated price is to use it to rank the twenty countries by creditworthiness, as shown in tables 2 and 3. Thus, we would have placed Egypt as the most creditworthy country of our sample in December 1983, Venezuela second, Colombia eighth, and so forth. We compared this rank ordering with the 1984 Euromoney rankings (published February 1985). The Spearman correlation coefficient is 0.65. While this is a reasonably positive association, inspection of the orderings shows that a major difference of opinion exists with respect to Turkey. Excluding this, the correlation coefficient for the nineteen remaining countries rises to 0.79. We conclude

Table 3. *Asset-Debt Ratios for Selected Developing Countries, 1983*
(millions of U.S. dollars)

Country	Official debt ^a (1)	Commercial debt ^b (2)	Total assets ^c (3)	Assets / total debt (4) = (3)/[(1) + (2)]	Net assets / commercial debt (5) = [(3) - (1)]/(2)	Estimated price: commercial debt ^d (6)	Euromoney rank, 1984 ^e (7)
1. Egypt, Arab Rep.	12,304.6	7,534.2	27,563.0	1.39	2.02	0.99	6
2. Venezuela	237.0	32,711.4	61,588.4	1.87	1.88	0.98	9
3. Thailand	4,312.1	8,909.0	20,996.6	1.59	1.87	0.97	3
4. Korea, Rep.	9,313.2	27,224.5	57,153.2	1.56	1.76	0.96	2
5. Malaysia	2,318.4	13,891.8	26,521.4	1.64	1.74	0.96	1
6. Yugoslavia	4,252.7	14,935.4	29,691.0	1.55	1.70	0.95	11
7. Cyprus	258.7	591.0	1,241.6	1.46	1.66	0.94	8
8. Colombia	3,488.1	14,508.4	25,210.0	1.40	1.50	0.91	5
9. Mexico	6,716.8	84,934.8	113,163.0	1.23	1.25	0.83	7
10. Brazil	10,058.2	83,756.9	114,144.0	1.22	1.24	0.82	13
11. Uruguay	382.2	2,724.7	3,724.6	1.20	1.22	0.81	15
12. Ecuador	1,638.8	6,244.9	8,440.5	1.07	1.09	0.75	17
13. Chile	1,479.8	16,045.4	14,740.3	0.84	0.83	0.59	19
14. Philippines	4,824.2	19,835.2	20,958.0	0.85	0.81	0.57	10
15. Peru	3,449.4	8,019.1	8,819.2	0.77	0.67	0.45	12
16. Bolivia	2,085.6	1,493.3	3,040.0	0.85	0.64	0.40	18
17. Argentina	2,013.5	46,635.0	17,731.1	0.36	0.34	0.10	16
18. Turkey	10,720.8	7,951.5	13,027.4	0.70	0.29	0.10	4
19. Costa Rica	1,436.7	2,686.9	2,115.4	0.51	0.25	0.10	14
20. Nicaragua	2,335.3	1,399.7	184.3	0.05	-1.54	0.10	20

Note: Countries are listed in order of estimated price of their commercial debt.

a. Defined in World Bank (1985).

b. Publicly guaranteed debt plus private nonguaranteed and short-term minus official debt.

c. For the definition, see table 2.

d. Estimated minimum price prior to regulatory default, for countries numbered 17-20.

e. Based on debt swap arrangements; reproduced for comparison with our estimate for country debt price.

Sources: World Bank (1985); *Euromoney* (February 1985).

Table 4. *Sensitivity Simulations: Calculated Market Value of Debt / Face Value of Debt*

Asset value / debt face value	Base case	Rescheduling ^a			Interest rate spread (percent)		Principal due (percentage of debt)			Front-end Fee (1 percent of debt)	Loans (1 percent of debt)	Higher interest (up 2 percent)	Higher variance (25 percent)	Default boundaries ^b		Seizure rate: (10 percent of debt)
		Excess	Limited	None												
					1	5	25	50	100					b = 2	b = 1	
0.20	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
0.30	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.03
0.42	0.18	0.15	0.15	0.11	0.21	0.18	0.18	0.14	0.17	0.18	0.20	0.15	0.17	0.02	0.02	0.19
0.53	0.31	0.27	0.27	0.22	0.34	0.30	0.31	0.26	0.28	0.31	0.33	0.27	0.29	0.03	0.03	0.31
0.62	0.40	0.37	0.37	0.31	0.43	0.40	0.41	0.36	0.37	0.40	0.42	0.36	0.38	0.25	0.03	0.41
0.73	0.51	0.47	0.48	0.43	0.53	0.51	0.51	0.48	0.48	0.51	0.52	0.46	0.48	0.36	0.04	0.51
0.92	0.66	0.63	0.64	0.60	0.67	0.67	0.67	0.66	0.64	0.66	0.67	0.62	0.63	0.54	0.05	0.67
1.00	0.71	0.68	0.69	0.65	0.71	0.72	0.72	0.72	0.69	0.71	0.72	0.67	0.68	0.60	0.05	0.71
1.49	0.91	0.89	0.90	0.88	0.88	0.94	0.92	0.94	0.91	0.91	0.91	0.88	0.88	0.85	0.60	0.91
2.05	1.00	1.00	1.00	0.99	0.96	1.05	1.01	1.02	1.00	1.00	1.01	0.99	0.99	0.98	0.84	1.01
3.07	1.07	1.06	1.06	1.06	1.01	1.12	1.06	1.05	1.02	1.07	1.07	1.06	1.06	1.06	1.01	1.07
4.95	1.09	1.09	1.09	1.09	1.03	1.16	1.08	1.05	1.03	1.09	1.09	1.09	1.09	1.09	1.08	1.09
6.30	1.10	1.10	1.10	1.10	1.03	1.16	1.08	1.05	1.03	1.10	1.10	1.09	1.09	1.10	1.09	1.10
8.00	1.10	1.10	1.10	1.10	1.03	1.17	1.09	1.05	1.03	1.10	1.10	1.09	1.10	1.10	1.10	1.10
10.18	1.10	1.10	1.10	1.10	1.03	1.17	1.09	1.05	1.03	1.10	1.10	1.09	1.10	1.10	1.10	1.10
15.18	1.10	1.10	1.10	1.10	1.03	1.17	1.09	1.05	1.03	1.10	1.10	1.09	1.10	1.10	1.10	1.10
50.40 (limit)	1.10	1.10	1.10	1.10	1.03	1.17	1.09	1.05	1.03	1.10	1.10	1.09	1.10	1.10	1.10	1.10

Note: The values shown are representative examples of the 100 points at which c , the ratio of the market to the face value of the debt, is evaluated.

a. "Excess" corresponds to a higher than optimal value of x_2 ; "limited" to an x_1 value that is too low (see figure 1).

b. The variable b = the regulatory maximum of the debt-asset ratio.

that the ranking produced by our methodology corresponds closely to the ex post Euromoney ranking.

Individual country observations also seem reasonable: the relative positions of Korea, Malaysia, the Philippines, and Thailand clearly correspond to their actual economic and financial circumstances. Colombia and Venezuela emerge as the most creditworthy of Latin American countries. Brazil and Mexico have very similar positions, while Costa Rica and Nicaragua are in critical financial straits. These results suggest that this methodology could prove valuable in analyzing country creditworthiness.

REFERENCES

- Black, F., and M. Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81: 637-54.
- Cox, John C., Stephen A. Ross, and Mark Rubinstein. 1979. "Option Pricing: A Simplified Approach." *Journal of Financial Economics* 7: 3.
- Eaton, J., and M. Gersovitz. 1981. "Debt with Potential Repudiation: Theoretical and Empirical Analysis." *Review of Economic Studies* 48: 289-309.
- Gennotte, Gerard. 1986. "Sovereign Debt: A Study of Alternative Payment Schedules." Berkeley: University of California Economics Department. Processed.
- Kulatilaka, N., and A. Marcus. 1985. "A Model of Strategic Default of Sovereign Debt." Boston, Mass.: Boston University School of Management. Processed.
- Merton, Robert C. 1973. "The Theory of Rational Optional Pricing." *Bell Journal of Economic and Management Science* 5: 141-83.
- Sachs, J. 1983. "Theoretical Issues in International Borrowing." Cambridge, Mass.: Harvard University. Processed.
- Sachs, J., and D. Cohen. 1982. *LDC Borrowing with Default Risk*. National Bureau of Economic Research Working Paper 925. New York: National Bureau of Economic Research.
- Sadeq, S. 1985. "LDC Debt: Growth Opportunities, the Option to Repudiate and Debt Rescheduling." Master's degree dissertation, Sloan School of Management, Massachusetts Institute of Technology.
- World Bank. 1985. *World Debt Tables*. Washington, D.C.

The Political Economy of Economic Liberalization

Deepak Lal

Two of the major policy problems facing governments of developing countries in the 1980s have been unsustainable external and internal disequilibria, and implementation of politically feasible stabilization cum liberalization programs which become necessary to correct these imbalances. This article discusses these "crises" and subsequent policy reform. The analysis suggests that balance of payments and fiscal deficits are frequently the result of use of an incorrect accounting system in a fixed exchange rate economy, and of public sector expansion beyond its economically feasible size; that governments usually seek to liberalize their economies during a crisis to regain control when the growth of the "transfer State" has led to generalized tax resistance, avoidance, or evasion; that reduction of the government role will be required to alleviate these crises; that sharp departures from past policies rather than gradual reform may be politically necessary; and that, contrary to the current technocratic opinion on this matter, the sequencing of a consistent and credible package of reforms which will most effectively reduce the costs of adjustment is initial liberalization of domestic capital markets simultaneous with cuts in the fiscal deficit, followed by floating the exchange rate and then commodity market liberalization.

With the growing importance of Internal Monetary Fund (IMF) stabilization and World Bank structural adjustment programs, there has been concern about the proper sequencing of the standard stabilization cum liberalization measures contained in these packages. In particular, there is grave concern at the very mixed and in some cases disastrous effects of liberalization attempts on incomes and employment in the Southern Cone of Latin America in the late 1970s and early 1980s. Much of the existing discussion of the order of liberalization is conducted within the traditional technocratic framework,¹ which seeks to determine the welfare-cost-minimizing deployment of policies in the standard stabilization and

1. This assumes a benevolent and well informed government maximizing a social welfare function subject to resource and technological constraints.

Deepak Lal, on leave from the University of London, is at the World Bank. This article is a revised version of parts II and III of Lal (1984b), part I of which now forms a revised and independent companion piece hereinafter cited as Lal (1986a). The author is grateful to Peter Kenen, Assar Lindbeck, and Joseph Stiglitz for valuable comments on earlier drafts of this article.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

adjustment packages. Not surprisingly, no clear-cut answers are possible to the unavoidably second best welfare questions that arise.

In a companion paper (Lal 1986a) these issues are analyzed using a standard two-good, two-factor Australian model of "real" trade theory, supplemented by a monetary model.² The effects of alternative policies on real wages and rental and exchange rates in the short and long run were examined. Here also, no absolute and generalizable answers about the optimal set of policies was found. The comparative static effects depend upon the relative factor intensities of the two goods (traded and nontraded), the degree of wage flexibility, and the extent of sectoral capital mobility. Comparative static results derived within this framework are summarized in the appendix for the case where the traded good is capital intensive (compared with the nontraded) and receives a subsidy to the use of capital. The apparent variations in these effects are compounded on the dynamic path of the real variables, which, depending upon the relative speeds of adjustment in the relevant markets, can lead to over- or undershooting from their relevant "equilibrium" values.

Nevertheless, based on the recent experience of the Southern Cone, there seems to be an emerging consensus that in liberalizing the foreign exchange, domestic credit, and labor markets, the capital account of the foreign exchange market should be liberalized last. Capital inflows should actively be discouraged (or held at a low constant level) while the other markets are being liberalized. (Edwards 1984 provides a good survey; see also Krueger 1984.)

In this essay I take issue with this emerging consensus on the order of liberalization. I argue that it is based on an implausible model of the attributes and behavior of the State and its citizens, and I suggest that a different, positive "political economy" approach may be more useful in answering questions concerning the sequencing and timing of measures of economic liberalization.

The typical stabilization cum adjustment program is launched when a country is in a "crisis," usually an incipient or actual balance of payments crisis³ that necessitates a reduction in the level of current expenditures. If there is underutilization of domestic capacity, an expenditure switching policy such as a devaluation may partly (or even fully) offset the extent of the cut in domestic absorption that may be required. Also, if the proper policy measures are taken, the country may be able to obtain some foreign resource inflows which could allow a grad-

2. The real models within this framework are based on Jones (1971), Mussa (1974), Burgess (1980), Neary (1978, 1982, 1985), and Corden and Neary (1982); the monetary model on McKinnon (1981), McKinnon and Mathieson (1981), Dornbusch (1974, 1980), and Krueger (1974); and the integration of the real and monetary aspects on Corden (1977), Corden and Jones (1976), and Prachowny (1981, 1984). Lal (1986b) provides an application of the real model to the Philippines and Lal (1985) of the real cum monetary model to Sri Lanka.

3. For our purposes, it makes little difference whether the crisis is caused by an increase in domestic expenditure, or, say, a decrease in the terms of trade facing the country. It also makes little difference whether the crisis arises from what are sometimes referred to as "insolvency" problems or to "liquidity" problems (though the validity of the conceptual framework underlying this distinction has recently been criticized by Eaton, Gersowitz, and Stiglitz 1986).

ual rather than sudden reduction in real expenditures. With any rate of social time preference greater than the cost of the foreign borrowing, the gradual approach must dominate any sudden shock treatment, if the former is feasible.⁴ But it is this very feasibility which is usually in question, essentially on grounds of the likely behavior of particular governments.

It may be possible to alleviate the crisis by persuading foreign creditors to extend more credit. This in turn may require that the foreign creditors believe that the country will actually carry out the proposed adjustment measures. This credibility will in turn be based on the reputation derived from the government's past policies.⁵ Thus, in the 1970s and 1980s the newly industrialized Southeast Asian countries (unlike many Latin American countries) have found it relatively easy to obtain external credit to smooth consumption following external shocks (from changing interest rates, commodity prices, and world demand). This is because the adjustment programs announced by these countries are credible given their reputations. By contrast, the announcements made by many Latin American and African countries may not be credible. If in the past a government has reversed preannounced plans because the costs of reversal (say, increased inflation) seemed to be lower than the benefits (say, financing a public sector deficit), then an announced adjustment program which is reversible may be unsustainable. Even if the "new" government has in fact changed its character, before outside creditors are willing to provide capital for smoothing intertemporal consumption, the government may have to demonstrate its newfound resolution by undertaking more Draconian disabsorption measures than would have been required if its announcements were credible. To improve its credibility, a government might choose to precommit itself to the new regime through external commitments, such as the binding covenants signed by Austria and Hungary with the Council of the League of Nations in the 1920s. These covenants reordered their fiscal and monetary systems, raised the costs to the governments of reversing their preannounced policies, and thus made the stabilization plans credible, which succeeded in stopping their hyperinflations almost immediately (see Sargent 1985).

Moreover, the announced speed of adjustment may be an important factor in determining the credibility of the stabilization cum structural adjustment program. For while from a technocratic viewpoint a gradual cut in expenditures may be desirable, the government may find that gradualism allows time for

4. This assumes concave adjustment costs. If adjustment costs are convex, discrete adjustments may be optimal.

5. *Fellner (1976, 1979) stressed the importance of the credibility of macroeconomic policymakers. Kreps and Wilson (1982) have provided one formalization of the notion of reputation. The formalization closest in spirit to that presented in the text, however, is that of Shapiro (1983). The literature arising from the problems of time inconsistency of optimal programs (see Kydland and Prescott 1977; Calvo 1978; Barro and Gordon 1983a, 1983b; Fischer 1980) has been extended by Backus and Driffill (1985a, 1985b, 1985c) in a series of papers which model dynamic games in which the credibility of the government's announcements is a central issue, as private agents are uncertain about the government's preferences.*

those hurt by the cuts to combine and exert irresistible pressure for their reversal. Politically, a long drawn out cut in real expenditures may thus be more difficult for a government to implement than a single, quick cut. One would expect that certain patterns of government behavior would be discernible which could provide some guidance on the question of the credible speed of adjustment. While this issue is of importance, I do not discuss it in any great detail in this article. But there is a presumption, given the fragility of many governments in the developing world, that long drawn out expenditure cuts will not be feasible in many developing countries.

There is another variant of the "easing of the pain" argument for gradualism in reducing real expenditures and removing policy-induced microeconomic distortions in the economy. These distortions place the economy well within the potential production possibility frontier (PPF). The various policy reforms undertaken to improve the overall efficiency of the economy are likely to induce growth in future real income and output (from the ongoing processes of growth—shifts in the PPF—as well as those increments induced by the improved policies—movement from inside the PPF to the frontier). As aggregate supply rises to eliminate the excess aggregate demand, this growth may preclude real cuts in current or future expenditures.⁶ Again, theoretically this would be uncontroversial. Controversy arises because of differing judgments on the *credibility*, tenacity, and political will of governments to complete a phased program of policy reform when the pain experienced by various sectional groups hurt by the removal of various "distortions" from which they have benefited is more immediate and more vocally expressed than the promised future joys from greater efficiency. This is also an empirical question of positive political economy.

Underlying these questions is a prior one: why countries get into a "crisis," and why some then choose and even succeed in not merely stabilizing but also liberalizing their economies. Section I considers this issue. The directions for the sequencing and timing of economic liberalization measures suggested by this discussion are used in section II to reexamine the question of the order of liberalization prescribed by the usual technocratic framework,⁷ which seeks to determine the welfare-cost-minimizing adjustment cum stabilization package for countries in "crisis." A brief final section summarizes my conclusions.

I. WHY DO GOVERNMENTS UNDERTAKE ECONOMIC LIBERALIZATION?

Most of the existing literature on the order of liberalization is based on the economist's traditional picture of disembodied, altruistic policymakers maximizing some social utility function subject to the usual resource and technological

6. This seems to be the argument underlying many of the contributions in Killick (1982).

7. I have found a specific factor version of the standard trade theoretic Australian model, supplemented by a simple monetary model, a useful heuristic device in answering these technocratic questions, as in Lal (1986a).

constraints. This view of the State is highly misleading, particularly when one considers the heterogeneous group of countries that comprise the developing world.

By contrast, it is more useful to follow the "new political economy" and view the State as composed of a group of self-regarding individuals and groups interacting strategically with private agents (see Staniland 1985; Findlay 1986; Lal 1984a; Srinivasan 1985). The State is then seen as seeking to maximize its own utility (including incomes, perquisites, and power) and not necessarily the welfare of its citizens. Nevertheless, in the resulting game between the State and its citizens (which will at different times have both cooperative as well as noncooperative elements), the latter may undertake various strategies which lead the State to serve the interests of its citizens.⁸ Thus, there may be mutually beneficial outcomes in some, perhaps most, circumstances.

Without setting out a formal model, I show in this part how a particular interpretation of past development experience may help to explain why and when States in the developing world choose to undertake economic liberalization (particularly structural adjustment programs) and also suggest the essential elements which must underlie a sustainable stabilization cum liberalization program. This analysis also leads to some empirical judgments about the order of liberalization.

Some Stylized "Crises," and When is a "Crisis" a Crisis?

It is useful to begin by presenting a taxonomy based on three sources of the excess absorption which usually leads to "crises." First, consider the case of excess absorption generated by monetary expansion accompanying increased government expenditure during the political cycle.⁹ With flexible money wages, the increase in absorption will initially lead to a rise in real wages.¹⁰ But with the ensuing leakage of the domestic money supply through the resulting payments deficit (at a fixed exchange rate), or else a contraction of real money balances

8. See Barro and Gordon (1983b); and Backus and Driffill (1985b, 1985c), for game theoretic models which provide a formal analysis of the problems of credibility and sustainability of a new policy regime when the reputation of the government is uncertain. However, while providing some useful insights, the robustness of these models may be questioned. Backus and Driffill (1985b), in an interesting paper which attempts to provide an endogenous explanation of the "political cycle" (as outlined by Nordhaus 1975; Lindbeck 1976; and Frey 1978), develop a two-person, noncooperative model in which an inflation conscious private sector is uncertain if the government which cares about output and employment will stick to its preannounced plan to disinflate (or not to inflate) on coming to office. "Instead of having a government create a pre-election boom in order to increase its chance of reelection, our analysis generates a pre-election boom as the solution to a game with a wet lame-duck government. In fact, if there were a chance of reelection, the incentive to preserve reputation may actually restrain the spending spree" (p. 536). They show that inflating at the end of its term is the rational response of a "wet" government that cares about employment. I doubt whether this conclusion is at all a plausible reason for the usual political business cycle; even if particular governments are mortal, parties are not!

9. Lal (1983, 1986b) provides an illustration from the Philippines.

10. This assumes that wages are more flexible (upward) than prices.

through a devaluation, all real variables would eventually return to their original values. In the process, the initial rise in real wages would eventually be reversed, as absorption falls back to its original, sustainable level. During this process of temporary excess absorption, the country will lose reserves, which may present a problem for the government. If the excess absorption is not reversed, however, or if reserves (which for this purpose should include access to external borrowing) are insufficient, at some stage the country will face a crisis. Appropriate expenditure reducing and (if there is nominal wage or price rigidity) expenditure switching policies such as a devaluation will be part of the resolution of the crisis. All this is well known. What is of concern here is why some countries, mainly in Latin America, are prone to follow highly inflationary macroeconomic policies which lead to dramatic periodic crises of overabsorption.

Second, one can consider another stylized case exemplified by the group of countries, mainly in South Asia, which have by and large followed fairly conservative macroeconomic policies. They have introduced a host of policy induced distortions in the working of the domestic price mechanism, however, which have led to rigidities in employment and output. Periodic stabilization crises arise when one of these countries faces a shift in taste or technology or some exogenous shock such as a harvest failure or a terms of trade deterioration.¹¹ Any of these changes requires an alteration of domestic relative prices to smoothly switch expenditures and outputs to maintain internal and external balance. The country may then suffer from an unsustainable excess of absorption, at the old unchanged relative prices. As is well known, an appropriate combination of policies to switch and reduce expenditures may be able to resolve the crisis. Again, what is of concern here are the reasons why and when such microdistorted economies are likely to reduce the rigidities which lead to their intermittent stabilization crises.

Third, windfall losses and gains of foreign currency are another common cause of "crises"; these are dealt with in greater detail in this section. Here, it is useful to distinguish between those windfalls which accrue directly to the public as contrasted with the private sector. Two recently occurring types of windfall are the foreign currency rents from minerals accruing to the public sector (for example, oil exporters such as Indonesia and Nigeria)¹² and the foreign currency receipts of the private sector derived from remittances by their relatives working abroad (as in Pakistan, Philippines, Sri Lanka, and Turkey).¹³ In the following

11. It has been argued within the technocratic literature that *temporary* falls in income should be partly financed by external borrowing, while if the fall in income is expected to be *permanent*, the country should adjust (see Cooper and Sachs 1985). The problem in practice is that except possibly for harvest failures it will be very difficult to judge whether a particular adverse shock to national income is temporary or permanent.

12. See Gelb (1986) for a discussion of the uses and abuses of these rents by various oil producers in the 1970s and 1980s.

13. See Swamy (1981) for some estimates of these remittances in the 1970s.

subsections, I assume with the technocratic literature that the government is optimizing and benevolent but nevertheless gets into a crisis if it is myopic or badly informed.

Private Sector Windfalls. First, consider *remittances*. Assume the exchange rate is fixed and there are exchange controls. The private sector receives foreign exchange remittances, which it exchanges for domestic currency at the central bank. This rise in the bank's foreign currency assets is initially matched by an equivalent increase in the currency and demand deposit component of the domestic money supply. High powered money has risen, however, and the domestic money supply *pari pasu* will expand, unless the government takes some countervailing action, by either reducing its own demand for credit through a reduction of the fiscal deficit or reducing commercial credit by raising the reserve ratio.

Suppose it does not sterilize the effects of these inflows on the domestic money supply. All other things being equal, at the fixed exchange rate, the expansion in money supply will lead to a rise in the price of nontraded goods and a trade deficit. In the standard trade theoretic framework, the effect on the real wage will be ambiguous. But with the loss of foreign exchange reserves accompanying the trade deficit, the high powered money base and thus the domestic money supply will contract automatically until the initial equilibrium is restored. In this process the "real goods" counterpart of the foreign remittance transfer on the capital account will have been affected through the trade deficit.

Is there any problem or "crisis" for such a country? It will experience an initial boom with inflation and a balance of payments deficit. Though these are usually symptoms of a crisis, in this case such a diagnosis would clearly be mistaken.

Suppose, however, as is likely, that there is some time lag between the receipt of the remittances on capital account and their subsequent implicit spending through the trade account. In the interim, foreign exchange reserves would have risen. Suppose, as is common in many developing countries, the government does not use the notion of high powered money in its budgetary planning. The treasury is advised by various economists that it should use the rising foreign exchange reserves for development purposes by running a larger budget deficit than it would have otherwise. If, as is usual, the government has not sterilized the foreign exchange inflows, the increase in high powered money entailed by the larger budget deficit provides a further boost to the money supply. Without this increase in the government's fiscal deficit, both the domestic inflationary process and the trade deficit would have been self-correcting, but this further expansion will entail more inflation and a larger trade deficit. Their correction through the "money-specie" flow mechanism will entail a *loss of foreign exchange reserves greater than those that had been received through the remittances*. The government, which had not intended to run down its initial level of reserves, *de facto* finds itself with an incipient rundown. It *now* has a balance of

payments problem, the cure of which will require the disabsorption we have already discussed.¹⁴

But this crisis is due to a misperception of the correct accounting system in a fixed exchange rate economy, with a government which is unable or unwilling to sterilize foreign currency flows. In particular, it arises from a misunderstanding about the relationship between high power money, the fiscal deficit (properly defined), and the movements in foreign reserves. Unless these are understood and the proper accounting framework is adopted, even though the government may (through financing or adjustment) adequately manage this crisis, it is likely that such problems will continually recur. The failure to adopt the correct accounting framework for budgetary and monetary management may thus be as important a factor in the generation of crises as any shocks administered by nature or by the external world.¹⁵

Public Sector Windfalls. The second type of "crisis" arises from windfall foreign currency gains or losses accruing to the *public sector*. Consider two radically different ways in which the government could in principle spend these rents. First, suppose it decides to distribute them annually, based on some suitable criterion of equity.¹⁶ A seemingly equivalent policy would be for the government to expand its expenditures by the annual inflow of these rents or to reduce general taxation. If the social value of increased government expenditure is considered to be greater than that arising from giving a form of national dividend or tax cuts to its populace, then the direct public use of these rents would seem to be desirable. Moreover, as long as the increase in government expenditure is covered by sales of the foreign currency to the central bank, the fiscal deficit which needs to be covered by domestic borrowing need not rise. Nevertheless, high powered money will increase with these inflows of foreign currency, unless the government actually reduces domestic credit, that is, unless they are sterilized.

The consequent rise in domestic money expenditure will raise the price level as (at a fixed exchange rate) the money prices of nontraded goods rise; part of the excess money demand will spill over into a trade deficit, financed by running down the newly built up foreign currency reserves. Assuming downwardly flexible prices of nontraded goods, this will tend to reduce the domestic money supply and bring the prices of nontraded goods down to their original level, and with it the overall price level.

But at a fixed exchange rate, if the inflow of foreign currency rents continues, the size of the domestic tradables sector will decline as demand for tradables is met by imports, and the relative price of nontraded goods will remain higher. To

14. Fry (1980) provides an illustration from Turkey in the early 1970s.

15. Of course, from the "political economy" perspective which I discuss in the next section, governments may deliberately choose not to employ a fully informative accounting framework, as this may allow them to conceal the implicit or explicit politically determined entitlements they wish to generate.

16. The U.S. state of Alaska did this with its oil revenues.

induce and then maintain the relative expansion of nontradables, *the real exchange rate (defined as the relative price of nontraded to traded goods) will remain at an appreciated level* for some time, depending upon how long it is expected to take the foreign currency rents to run out. The requisite rise in the relative price of nontraded goods need not come about through this inflationary process, if instead the government chose to *appreciate* the nominal exchange rate by the appropriate amount.

A problem would arise if the increased level of government expenditure was unsustainable over the long run. This could happen if the government misjudged the size of the annual foreign exchange flows and committed itself to long maturing investments or unsustainable consumption support programs, which would need to be cut back if there were any falling off in expected foreign exchange rents. The public sector expenditure program may thus be unsustainable.

If, however, the rents had been transferred to the general populace, each individual would have had his or her current income increased and expectations of future rises in income improved. The resulting decentralized decisions concerning the privately optimal consumption/investment mix would, in an open economy, involve portfolio choices between different domestic and foreign financial assets. The associated patterns of deficits and surpluses on the current/capital accounts of the balance of payments would have no overall welfare significance.

Dirigisme, Fiscal Crises, and Government Control

The above examples illustrate how most so-called balance of payments crises are in large part crises for the public sector and reflect misjudgments about its appropriate and feasible size and composition.¹⁷ They also suggest the reasons why self-regarding governments may choose to reverse their previous dirigisme (inducing macro and/or micro distortions) and liberalize an economy they have hitherto repressed.

One of the paradoxical dynamic effects of the past dirigisme of many countries in the developing world (as, of course, in some developed countries) has been that attempts to exercise political control over ever-increasing areas of economic life have often led, after a certain stage, to a diminution of the government's effective areas of control as private agents find numerous ways of avoiding them. What concerns us is the empirical observation that there appears to be a sort of "Laffer curve" of government intervention, so that after a certain stage, increased government intervention, instead of increasing the area of government control, diminishes it. It is rare for liberalization to follow some intellectual conversion of policymakers who, having seen the errors of their ways, seek to find a second best welfare maximizing transition from a controlled to a market economy. Rather, various measures of economic liberalization and/or stabiliza-

17. This is the important insight contained in Corden (1977).

tion are most often sought in order to reassert government control over the economy. It is important to determine the *costs to the State* of *not* liberalizing the economy. For sustained liberalization to be undertaken, these must be greater than the apparent costs of liberalization, namely, the inevitable political pressures which will follow from the changes in distributional rewards and rents which are entailed in any measure of liberalization.¹⁸ This subsection elaborates on these points.

A major motivation for liberalization from the viewpoint of many States in the developing world lies in an attempt to regain control over an economy which seems to be less and less amenable to the usual means of government control. Usually the most important symptom of this malaise is a creeping but chronic fiscal crisis (also reflected sometimes but not always in a balance of payments crisis), which has in different forms beset most economies—including developed ones—in the last decade (see Lal and Wolf 1986 for a fuller discussion). Its origins lie in the creation of politically determined “entitlements” to current and future income streams for various groups in the economy (the deserving poor; industrial labor; regional interests; old age pensioners; infant, declining, or sick, industries—to name just a few). As these entitlements are implicit or explicit subsidies to particular groups, they have to be paid for by implicit or explicit taxation of other groups in the economy. However justifiable on grounds of social welfare, the gradual expansion of this “transfer State” leads to some surprising dynamic consequences.

The gradual expansion of politically determined entitlements creates specific “property rights.” The accompanying tax burden to finance them leads at some stage to generalized tax resistance, avoidance, and evasion and to the gradual but inevitable growth of the parallel or underground economy. This has been the case with both developed and developing countries in the past decade. Faced with inelastic or declining revenues but burgeoning expenditure commitments, incipient or actual fiscal deficits become chronic. These can only be financed by three means: domestic borrowing, external borrowing, or levying of the inflation tax.

Many countries, particularly those in Latin America, have tried all three—with dire consequences. Domestic borrowing to close the fiscal gap may crowd out private investment (see Blejer and Khan 1984) and diminish the future growth of income—and thus the future tax base. The fiscal deficit may be financed by foreign borrowing for a time, particularly as in the mid-1970s, when

18. One could, following the “State as pressure group” school of political economy, seek to explain the move to liberalization as resulting from a new pressure group equilibrium (see Becker 1983). But this model of political economy relies on political institutions corresponding to those in Western democracies. Its applicability to the varied authoritarian regimes in the developing world would seem to be limited. Hence my attempt to explain why a developing world government which is *relatively* (but not completely) immune to democratic pressure group activity would seek to liberalize its economy. Another way of making this point is that instead of considering the State to be a passive transmitter of pressure group activity, one looks upon it as an autonomous agent with its own goals (see De Jasy 1984).

real interest rates were low and even negative. But this form of financing is inherently unstable. The debt service ratio can become unviable if, as in the late 1970s, world interest rates rise and the ability of the economies to generate the requisite export and fiscal surpluses to service the higher interest costs of publicly guaranteed debt is limited. This is often due to policy induced distortions inhibiting exports—for example, the maintenance of overvalued exchange rates and high and differentiated effective rates of protection which are an indirect tax on exports—and the difficulty in generating fiscal surpluses to match the interest on the debt. Thereupon, foreign lending can abruptly cease, leading to the kind of “debt crisis” which has plagued Latin America in the 1980s. The third way of financing the deficit, through the use of the inflation tax, is also unviable over the medium run, for it promotes a further growth of the parallel economy and a substitution of some indirect or direct form of foreign currency based assets for domestic money as a store of value. The tax base for levying the inflation tax thus shrinks rapidly.

With taxes being evaded, with domestic and foreign credit virtually at an end, and with private agents having adjusted to inflation to evade the inflation tax, the government finds its fiscal control of the economy vanishing. The growth of entitlements, moreover, reduces the *discretionary* funds available to the government, and it is discretionary funds which give the government power. It may not even be able to garner enough resources to pay the functionaries required to perform the classical State functions of providing law and order, defense, and essential infrastructure. This dynamic process whereby the expansion of the transfer State leads to the unexpected and very unMarxian withering away of the State has rarely reached its full denouement, although in some Latin American countries it may be close.¹⁹

But well before things come to such a dire pass, attempts are usually made to regain government control. Two responses by the government are possible—an illiberal and a liberal one. The former (which is rarely observed) consists of a further tightening and more stringent enforcement of direct controls. Tanzania provides an example of this response. If this tightening is effective, and the private utility of after-tax income received from legal productive activity declines to the level at which untaxed subsistence activities are preferable, however, producers may seek to escape the controls by ceasing to produce the taxed commodities altogether. The tightening and enforcement of controls could lead to an implosion of the economy.²⁰ The government might then find that as producers return to untaxable subsistence activities, the very production base over which it seeks control has shrunk or disappeared.

19. For example, in Peru it is estimated (by Hernando de Soto in private communication) that over 70 percent of the labor force in Lima works in “illegal” activities, the government has no domestic or foreign credit, inflation is high and rising, and nearly 70 percent of the money supply is in dollar denominated deposits.

20. See Collier and others (1985) for such an interpretation of recent Tanzanian economic policy and its outcomes.

The more usual response is to regain a degree of fiscal control through some liberalization of controls on the economy. Typically, however, these liberalization attempts are half-hearted and include some tax reform, monetary contraction, and some measures of export promotion. Their aim is to raise the economy's growth rate as well as the yield from whatever taxes are still being paid and to improve the debt service ratio in the hope that this will lead to a resumption of voluntary foreign lending. But unless the underlying fiscal problem (which is largely that of unsustainable public expenditure commitments) has been tackled, these liberalization attempts have usually been aborted.²¹

Without a commitment to reducing unviable levels of entitlements, the liberalization attempts have tended to worsen the fiscal situation. With the lowering of tax rates and lags in supply response, revenues do not rise and may even fall initially. The necessary reductions in money supply to contain inflation reduce the limited seigniorage previously being extracted.²² Government unwillingness to allow either public or private enterprises to fail entails absorbing the deficits of public enterprises as well as any newly sick units taken over, as the liberalization exerts competitive pressures on unviable firms. Moreover, where liberalization has been accompanied by large public or private capital inflows (often to finance the public sector deficit), there has been an appreciation of the real exchange rate sometimes accompanied by inflationary pressures arising from inappropriate nominal exchange rate policies (as in Sri Lanka; see Lal 1985). This appreciation thwarts potential export growth, so that as capital inflows diminish, the incipient fiscal deficit is once again reflected in a chronic balance of payments problem which the government then seeks to control in the old unviable ways—and the liberalization process is reversed.

The above patterns have been observed in a large number of countries which have attempted to liberalize in the 1970s.²³ The major lesson to be drawn is that liberalization is often undertaken to gain fiscal control, but if nothing is done to rescind unsustainable public expenditure entitlements a stabilization cum balance of payments crisis eventually emerges which undermines the attempt to liberalize the economy. It would thus seem that a *sine qua non* of a sustainable liberalization attempt must be the prior establishment of fiscal control through a reduction of unsustainable public expenditure commitments.²⁴ The stabilization

21. Some examples are the Sri Lankan 1977 liberalization episode (see Lal and Rajapatirana, forthcoming); the 1978 Argentina liberalization (see Nogués 1981, Calvo 1986); and the latter part of the 1974 Uruguay reform episode (see Hanson and de Melo 1985).

22. At the end of a period of hyperinflation, however, the demand for money will increase, as will the seigniorage associated with any increase in money supply to meet this incremental demand for the newly stable money.

23. This experience is being analyzed in two sets of multicountry comparative studies undertaken by the Trade Policy Research Centre and the World Bank. My reading of their preliminary findings provides empirical support for the following remarks.

24. Unless, of course, there are sufficient underemployed domestic resources so that an expenditure switching policy such as a devaluation can rapidly increase aggregate supply and thence public revenues by the requisite amount.

of the economy no less than any prospective liberalization also entails a willingness to overcome the resistance of those whose entitlements will be rescinded.

The *political* problem governments most usually face when considering economic liberalization is that the pressures from the potential losers from the liberalization tend to antedate the support which will subsequently be provided by all those who gain. As can be readily shown, under many circumstances there will be losers in the short run from the changes in relative prices and/or disabsorption flowing from stabilization and structural adjustment programs, particularly those whose relative returns on sector specific human and physical capital fall as a result of these changes. If the liberalization is sustained, however, it is likely to yield higher and more efficient growth in income, which will benefit most groups in the economy. Given the government's own rate of discount (which may be much higher than that of society), even if the resulting purely technocratic economic welfare integral is positive, policymakers may still be reluctant to undertake the reforms if they feel uncertain about their ability to survive the political pressures during the transition.²⁵

II. MINIMIZING THE WELFARE COSTS OF ADJUSTMENT

Consider now the problems facing the government of a developing country which finds itself in a "crisis" but has the requisite political will to implement the policy recommendations commonly contained in stabilization and adjustment programs.²⁶ These are (a) measures to reduce absorption by cutting the fiscal deficit, limiting overall domestic credit and devaluing the currency (if expenditure switching is required), and (b) measures to improve the supply side of the economy by reducing foreign trade and wage-price controls and by removing interest rate ceilings, changing reserve requirements, and possibly eliminating exchange controls.

The first group of disabsorption measures is unavoidable in a crisis situation. Only their extent is an issue, which (as was argued in the beginning of this article) depends upon the level of foreign financial accommodation available. The desirable level of this accommodation is a matter of judgment, and I have little to add to the earlier discussion.

The remaining questions concern the sequencing of the supply-side policy reforms and whether there are accompanying policies which might be able to minimize the inevitable costs of adjustment during the process of economic liberalization.

25. The "optimal" subsidy route for cushioning the transition for "losers" does not make much sense in this context, for it again assumes an omnipotent and omniscient government. The *political* difficulties I am emphasizing arise precisely because the government does not possess these attributes. Moreover, even if the government were able to implement this gradual adjustment policy, such a policy may not be time consistent and hence feasible.

26. Discussion of these will be found in Guitian (1981) on the IMF's conditionality, and in Stern (1983) on the World Bank's structural adjustment programs.

Optimal Sequencing of Policy Reforms

This subsection addresses welfare theoretic questions posed in the usual technocratic framework, which assumes a benevolent government concerned with maximizing some social welfare function. It is obvious that for an economy in which adjustments in both capital and labor markets are rapid, each of the policy changes proposed in the package must necessarily improve efficiency and hence lead monotonically to higher level of real national income until the new undistorted equilibrium is reached.

Moreover, within the standard trade theoretic Australian model the effects on real wages, particularly in the long run, of all these possible policy reforms is positive if, as one would expect, the factor market distortions are in the capital-intensive traded good sector. At least for the fairly flexible textbook economy outcomes which correspond to the Marshallian long run, the standard adjustment package will not lead to any efficiency or equity losses.²⁷ In these circumstances, there would be little sense in examining the sequencing and timing of liberalization measures, as there should be a simultaneous and instantaneous liberalization of all the relevant markets.

Short run immobilities of capital and rigidities in wages, however, can lead to short run falls in real income or real wages. This is one reason why governments seem unwilling to liberalize all markets simultaneously (see Krueger 1984). Two questions then arise: first, whether there are feasible supplementary policies which could minimize these efficiency and/or equity losses during the transition, and second, whether there is a policy sequence within the adjustment package which will reduce the pain of the transition. Arguments are presented below in favor of *one* plausible sequencing of policies which runs counter to the currently favored sequence (namely, opening the capital account last). This is to underline the fragility of much of the current argument for particular orders of liberalization being derived in the technocratic literature.

On the first question, consider the case of wage rigidity. It is well known from the literature on project evaluation that, given certain assumptions, any stickiness or rigidity of wages may make the shadow wage lower than the actual wage rate (see Lal 1974 for a summary of the literature on the determinants of the shadow wage rate). In theory, a wage subsidy financed by lump sum taxes would yield the best outcome during the transition. If this is not possible, then some suitable tax-subsidy combination which essentially subsidizes the output of the labor-intensive industry may be desirable. Because self-employment predominates in most developing countries, it will not be feasible to institute a general wage subsidy, and a tax-subsidy scheme on output may be the only feasible policy. The required general subsidy to the labor-intensive sectors can be provided by in effect taxing the output of the capital-intensive sector, that is, by appropriate changes in the real exchange rate. Bhagwati (1979) and Krueger

27. We assume that there is little weight to be placed socially on the losses suffered by those capitalists who lose from the removal of the capital market distortion.

(1978) detail various ways in which the removal of trade distortions and changes in the nominal exchange rate can be combined to yield the requisite change in the real exchange rate.

The devaluation and reduction in output distortions required to obtain the real exchange rate change is likely to have an ambiguous short run effect on real wages if capital is immobile in the short run (see section 4 of Lal 1986a). By contrast, removing the capital market distortion in the capital-intensive traded good sector (with a given distortion in the labor market, and unchanged trade distortions) will raise both real output and real wages in the short and long run. This suggests the following sequencing of the adjustment package. First, remove the capital market distortion, and as the real output and real wage gains begin to appear, begin a phased program to reduce distortions in the commodity markets (particularly for traded goods). The possible short run declines in real wages resulting from the second phase will be mitigated by the wage increases arising from the continuation of the first phase of policy reform.

Removing the capital market distortion, however, is a tricky problem for a financially repressed economy (see McKinnon 1981). It is clear that, as a first step, interest ceilings on deposits and loans should be removed. If the distributional impact of these changes is to be spread over a period of time, then any quantitative credit controls (which imply capital subsidies to favored borrowers) should be converted into explicit subsidies and shown as part of the government budget. Without an increase in government revenues, the consequent rise in the fiscal deficit will require the levying of a higher inflation tax and a suitable adjustment of the reserve ratio on interest bearing domestic money deposits (this need not necessarily imply an increase in that ratio; see McKinnon and Mathieson 1981).

But most important, it will be necessary to keep the economy at the new steady state inflation rate to alter the nominal exchange rate at the same pace as the steady state inflation rate. A fixed nominal exchange rate would not be sustainable. In time, to reduce the inefficiencies of the financial repression associated with high reserve ratios, it will be necessary to reduce the fiscal deficit and thence the inflation rate. Whether this reduction in the steady state inflation rate will be welfare-improving depends on the alternative net social costs of reducing government expenditure and raising taxes to reduce the fiscal deficit. It may then seem to be better to sequence the inflation control program *after* the removal of interest ceilings, the introduction of flexibility in exchange rates, and the start of a phased program of trade liberalization. As real national income rises in response to the policy changes, government revenue will also rise to some extent. Then, if the government's real expenditure remains unchanged, the real fiscal deficit and hence both the reserve ratio and inflation rate should fall.

Though in theory one can couch arguments in terms of steady state inflation rates, in practice it will be very difficult (even for an omniscient and benevolent government) to maintain this steadiness. It may, therefore, be best to combine the reform of the domestic capital market with the reduction of the fiscal deficit

and inflation. The last part of this initial package will be made easier by the likelihood that some deficit cuts will have formed part of the initial disabsorption required to deal with the crisis. So the first stage of the adjustment package should probably combine the reduction of domestic capital market distortions with the reduction of the fiscal deficit and inflation.

What exchange rate regime should be adopted by the liberalizing government during and after the transition? I have argued elsewhere (Lal 1980) that an exchange rate regime with an *automatic* balance of payments mechanism be adopted in a world of irreducible uncertainty, where discretionary government action (required in all forms of managed exchange rate regimes) cannot be expected to achieve convergence or reduce deviations from an unknowable equilibrium. The only two choices are a completely fixed or a fully floating exchange rate regime.

It may be worth noting that Chile—unlike the other two Southern Cone countries, Argentina and Uruguay—followed the conventional sequencing pattern: it first controlled its fiscal deficit, second, liberalized its trade regime, and third, liberalized its domestic capital market and opened the capital account of its balance of payments. The Chileans also implemented the advice of international monetarists on exchange rate policy to the letter: they progressed from a crawling peg through a preannounced exchange rate regime to a fixed exchange rate system.²⁸ Prima facie, the subsequent Chilean debacle presents a dramatic real life argument against this exchange rate choice.

In my judgment, a fully floating rate would have served the Chileans much better: it would have choked off some of the short run capital flows which supposedly destabilized the economy in the early 1980s.²⁹ Moreover, as Black (1976) and Lal (1980) have noted, the institutional requirements of a floating exchange rate system are probably exaggerated, as the recent floating of a number of African currencies demonstrates.

In summary, my suggested sequencing of liberalization would be as follows: first, reduction of the fiscal deficit, accompanied by removal of domestic capital market distortions; second, elimination of exchange controls and free floating of the exchange rate, accompanied by an announcement of a phased program for removing commodity market distortions; and third, the implementation of this preannounced phased program.

On Liberalizing the Capital Account

A number of objections can be raised against this sequencing, however, and in particular against the proposed liberalization of the capital account simultaneously with the initiation of a phased program of trade liberalization. These

28. See McKinnon (1982), who wrote: "the correct order of liberalization . . . approximates the successful Chilean experience after 1975" (p. 159).

29. For alternative views about the role of the pegging of the exchange rate in the Chilean debacle, see Dornbusch (1985) and Harberger (1985).

objections have supposedly gained greater credence as a result of the effects of the liberalization of the capital account in the 1970s in a number of countries in the Southern Cone of Latin America.³⁰

But *prima facie*, it is not obvious that this experience provides arguments against my preferred sequence. As I have noted above, Chile liberalized its trade account before its capital account, and yet it suffered a debacle, arguably because it maintained a fixed exchange rate with real wage indexation and lax regulation of domestic financial institutions. Argentina liberalized its capital account without undertaking either any substantial trade liberalization or (more importantly) fiscal stabilization and instituted the system of preannounced exchange rates (the *tablita*). This program was clearly inconsistent and hence not credible or sustainable (see Fernandez 1985; Calvo 1986). Uruguay started with the opening of its capital account and the deregulation of its domestic financial system before it undertook trade liberalization. As Hanson and de Melo (1985) note, "no major imbalances occurred before 1979, when the financial liberalization was completed and the real exchange rate did not appear to be in disequilibrium from the viewpoint of the allocation of real resources in the long run," and hence "its experience does not seem to support the common view that the current account should be liberalized before the capital account" (p. 934). The debacle occurred because in October 1978, Uruguay adopted the *tablita*, and in 1982 and 1983 its fiscal deficit increased "due to an increase in social security payments and government salaries and a reduction in labor taxes" (Hanson and de Melo 1985, p. 933). These policies were clearly inconsistent. Thus, rather than blame the opening of the capital account for the crises, it would seem that in all three cases the *tablita* and/or pegging of the exchange rate made the programs inconsistent.

The lesson I would draw from this experience is that an appropriate exchange rate regime during the liberalization process and the maintenance of fiscal stability are essential for a sustainable liberalization program. For the reasons given earlier, a floating exchange rate which requires both domestic financial liberalization and the opening of the capital account would seem to be preferable to a fixed or managed exchange rate regime when liberalization is undertaken. We must consider, however, various technocratic *a priori* arguments that are raised against opening the capital before the current account of the balance of payments.

The first objection concerns the foreign capital inflows which might be induced as a result of the liberalization of the capital account while there are still distortions in domestic commodity markets. Theoretically, it is possible that capital inflows to a domestically distorted economy may be immiserizing (see Brecher and Diaz-Alejandro 1977). In my proposed package, however, the liberalization of the capital account would be accompanied by an announcement of a

30. See the symposium issue of *Economic Development and Cultural Change*, April 1986, and that in *World Development*, August 1985, on the liberalization experience of the Southern Cone.

future (dated) phased program of reduction in commodity market distortions. If one assumes that expectations are rational, and the public announcements credible, it is unlikely that long term capital flows (based on investors' time horizon extending beyond the trade liberalization phase) would be immiserizing. Because some foreign capital with shorter maturities is likely to flow into activities where private and social rates of return diverge during the trade liberalization phase, there may be a case for temporary and preannounced taxes on such flows by maturity, which are gradually reduced to zero as the trade liberalization proceeds.

The second objection relates to the likelihood of an initial appreciation in the real exchange rate with the removal of exchange controls, before trade has been liberalized. The argument can be stated as follows. With the liberalization of the domestic capital market, domestic real rates of interest will rise. Given the relative scarcity of capital in developing countries, they are likely to be higher than foreign interest rates at the original exchange rate. Hence, the removal of exchange control and the institution of a free float will appreciate the nominal exchange rate to establish interest rate parity. If, in the ensuing process of balance of payments adjustment, foreign capital also flows into the country, the real exchange rate will rise as the country runs a balance of trade deficit to match the surplus on the capital account of the balance of payments.³¹ This will lead to resource movements opposite to those required during the process of trade liberalization when the real exchange rate is likely to fall. Hence, it is argued that the opening of the capital account before the trade liberalization is completed will lead to unnecessary shifts in resources, in effect away from traded to non-traded goods industries.

Before discussing this resource shifting argument, it is desirable to contrast the alternative sequencing of the opening of the trade and capital accounts. Suppose, instead of floating, a managed exchange rate system requiring capital controls is maintained during the trade liberalization process. The government will have to judge the precise extent of the *nominal* exchange rate change required to yield the appropriate *real* exchange rate at each point during the transition. Suppose it misjudges and at some point during the transition undertakes a nominal devaluation which is *less* than that required to yield the appropriate real exchange rate. This will lead to an incipient or actual balance of payments deficit which the government may be tempted to reduce through import controls—thus aborting the trade liberalization.

Furthermore, as can be readily shown (depending upon the dynamic structure of the economy, the relative factor intensities of traded and nontraded goods and the degree of short run mobility of capital), the trade liberalization process could lead to an initial real depreciation greater than that required in the final equilibrium. In this case also, the real exchange rate would first fall and then rise, with

31. The resulting short and long run resource and real wage movements can be readily derived in the framework provided in Lal (1986a).

unnecessary resource movements similar to those caused by the opening of the capital account. It is not apparent how one could determine the extent, and choose between these two alternative sets of unnecessary resource movements, in advocating one sequencing pattern over the other.

By contrast, one of the major advantages of instituting a free float before trade is liberalized is that the nominal rate changes which are required during the process of liberalization become automatic. Given the uncertainty in the dynamics associated with the process of adjustment, it is very difficult in practice to judge the precise extent—and in some cases even the direction—of the *nominal* exchange rate changes required in a managed exchange rate system as trade liberalization proceeds. Arguments for a crawling peg, to be followed by a fixed exchange rate regime once full liberalization of commodity markets is achieved (advanced for instance by McKinnon 1982), ignore these difficulties. The arguments for this set of policies are based on the erroneous view that the government, by pegging the *nominal* exchange rate for a set period of time, can also peg the *real* exchange rate (see Stockman 1982, p. 189). Because a major cause of the crises in Latin America's Southern Cone was mismanagement of the nominal exchange rate, a floating rate, by obviating the need for discretionary nominal exchange rate management, will relieve the government of concerns about the balance of payments consequences of other measures of economic liberalization.

The third objection to capital market liberalization in the initial stages of liberalization relates to the likely dynamic effects of foreign capital inflows. These are symmetrical with those discussed in connection with the "Dutch disease" (see Corden and Neary 1982). If there is a sustainable level of foreign capital inflows following the liberalization of domestic capital markets and the capital account of the balance of payments, then the resulting real exchange rate appreciation cannot be said to constitute a problem for the country. Real income will be higher, even though the resulting sectoral shifts in resources will entail distributional shifts in the returns to sector specific factors of production. But these shifts will also result from other policies required to liberalize a repressed economy. They cannot in themselves constitute an argument against liberalization, but they do provide an indication of the likely directions from which opposition to particular liberalization measures might be expected.

The real worry of the opponents of the liberalization of the capital account is that it could lead to an overshooting of capital inflows with an accompanying overshooting in the real exchange rate.³² It is certainly conceivable that such overshooting may occur and that it will entail greater reversals of resource movements than on the "equilibrium" path of the real exchange rate. But with short run immobility of domestic capital, trade liberalization too will entail reversals in resource movements. In fact, in any dynamic economy with changing relative prices, such reversals (or more correctly changes) in resources move-

32. See Edwards (1984) for a summary of this argument, and the references to the Southern Cone experience with liberalization on which this currently dominant view is based.

ments can be expected to be taking place continually. Moreover, *all* resource movements involve adjustment costs. Unless some divergence between the social and private values of the adjustment costs associated with reversals in resource movements can be discerned, their mere existence does not in itself tell one whether moving resources in and out of particular industries during a dynamic process constitutes a policy problem which should determine the order of liberalization.

Consider the case where it is known by private agents (whose expectations are rational) and the government's Platonic advisers that, with my preferred sequencing of capital and current account liberalization, the real exchange rate will first appreciate and then depreciate over time, so that the relative profitability of nontraded relative to traded goods industries will first rise and then decline. Corresponding to this expected real exchange rate movement, private agents can calculate the present value of investments as relative prices change at any date during the transition, including the cessation of current income streams from investments which fail to cover variable costs. This will yield a private profit maximizing allocation of resources in the two sets of industries over time, which may well involve an initial expansion and then contraction of the non-traded good industry.

Moreover, as this dynamic sequence of investments and sectoral outputs is ex hypothesi based on the same expected real exchange rate movements as those foreseen by the government's advisers, this is also the "optimal" sequence.³³ Even a temporary capital inflow, by improving domestic resource availabilities, must raise the current and the discounted present value of future national income as long as either the domestic rate of time preference (if the inflow finances consumption) or the domestic rate of return of investment financed by the inflows is higher than the effective real interest rate on the capital inflows. It is only if one assumes that the government knows the correct time path of the real exchange rate while myopic and/or ignorant private agents do not that an argument against this sequencing can be made based on a divergence between the private and social costs of adjustment in this dynamic process.

Frenkel (1982) does present an argument based on divergences between private and social costs on alternative orders of liberalization. He argues that "if the capital account is opened up first, portfolio decisions are likely to correspond to the long-run undistorted conditions, but real investment will still be carried out in a distorted environment as long as the trade account is not opened up. Due to the distortions, the social cost of the investment is likely to exceed the private cost. These real investments will have to be reversed once the trade account is liberalized. Due to the difference in private and social costs, it is likely that the first order should be preferred" (p. 200).

33. I assume, of course, in this example that there is no divergence between social and private rates of time preference. But I do not seek to imply that this path is necessarily Pareto efficient. The main point is that there is *no policy* which in the postulated circumstances can yield a welfare improvement.

The crucial assumption in this argument is the assumed asymmetry of information between farsighted foreign portfolio investors and ignorant or myopic domestic real investors. Exactly the opposite outcome is likely if both groups are equally well informed. For as Frenkel himself argues, it is much "easier to reverse wrong portfolio decisions than to reverse wrong real investment decisions." Thus, it will be in the self-interest of investors involved in production decisions to have much longer time horizons in considering their real physical investments than do financial investors!

Finally, if this overshooting argument against the removal of capital controls is correct, it should apply to the liberalization of the capital account *at any time*, irrespective of whether or not trade liberalization has taken place. It would imply that permanent capital controls should be maintained because of the presumed suboptimal resource movements which myopic agents will undertake in response to any overshooting accompanying the liberalization of the capital account. This is hardly a credible argument.

I would conclude, therefore, that some of the currently conventional arguments against opening the capital account at the same time as a phased program of trade liberalization is announced must ultimately be based on the erroneous (but also common) assumption of the limited foresight of private agents and the clairvoyance of governments and their advisers.

III. CONCLUSIONS

Most of the "crises" which require the stabilization cum structural adjustment medicine are primarily crises of the public sector. For many developing countries in the past three decades, a vast expansion of the government bureaucracy, of the public sector, and of controls on industry, prices, and foreign trade have created a new system of subinfeudation, in which politically created property rights to rents for various groups are financed by implicit or explicit taxation of the general populace. In many countries these groups include a significant portion of the bureaucracy, public sector functionaries, industrial labor, and those who have been granted the monopolistic protection of the use of their labor or capital by the State's prohibitions on entry and exit of economic enterprises.

A crisis arises as the existing means of funding these entitlements become unviable. Economic liberalization is then seen as a way to regain government control of the fiscal accounts and the economy. The process of gaining fiscal control by reducing public expenditure to a sustainable level and the subsequent liberalization program must inevitably entail either confronting these vested interests or buying them out. The latter approach will only be viable, however, if it is ensured that the economic and legal framework contains enough checks and balances to prevent the State from creating future artificial rents for other vested interests. This is the strongest political and practical case—as distinct from that based on the controversial grounds of economic efficiency—for the promotion of a market economy.

Therefore, if a government can demonstrate its political strength in establishing fiscal control, the question of the order of liberalization which minimizes the pain to be suffered by sectoral and sectional interests during the movement to a market economy becomes politically less important. Moreover, for the *credibility* of the process, speed in implementing the liberalization of the various repressed markets may be of the essence. This speed is necessary both because of the time inconsistency of technocratic programs (whereby for the government it is usually optimal to renege on its past commitments)³⁴ and to preempt the formation of coalitions by the losers from the liberalization. To assure its credibility, the government might wish to institute a speedy bonfire of *all* controls and Ulysses-like tie itself to the mast by signing a stabilization and structural adjustment program with an international organization. If backed by sufficient resource inflows, this *could* provide the restraints necessary to forestall any future temptation for itself or its successors to reverse the liberalization and thus could make its announced program of liberalization sustainable.

Looked at from this perspective, most of the technocratic literature on the order of liberalization does not seem too relevant, though I have provided some arguments against the currently fashionable view that the capital account of the foreign exchange market should be liberalized last. There is still the controversial technical question of the exchange rate regime the liberalizing government should adopt during the transition, and the regime it should aim for when the liberalization is completed (if different). It was the inconsistency of exchange rate policy which largely explains the Southern Cone debacles. I have provided some reasons why an automatically adjusting floating exchange rate system may be desirable. These are essentially to minimize the information required to set particular nominal values at their optimal levels, and hence to avoid the risk of mistakes in the desirable extent and even the direction of changes of the *nominal* exchange rate.

There is, however, one major unresolved issue which the recent Chilean experience in particular highlights. This concerns the desirable form of regulation of domestic financial institutions which a developing country seeking to liberalize its domestic capital market should adopt during the transition, and as part of its ultimate economic framework. Besides wage indexation and exchange rate policy, it is the mishandling of the financial liberalization which many observers see as the third crucial ingredient in the Chilean debacle with economic liberalization (Congdon 1985). But these policies raise issues which go beyond the scope of this article.

34. Kydland and Prescott (1977) is the seminal contribution on this issue; also see Elster (1979).

Appendix. *Stabilization and Liberalization Policies: Direction of Change from Initial Equilibrium*

Policy changes and assumptions	Flexible wages								Sticky wages			
	Short run, sector-specific capital				Long run capital mobility				Short run, sector-specific capital			
	Real wage	Interest: nontraded goods	Interest: traded goods	Real exchange rate	Real wage	Interest: nontraded goods	Interest: traded goods	Real exchange rate	Real wage	Interest: nontraded goods	Interest: traded goods	Real exchange rate
Capital markets: removing nontraded good tax or traded good subsidy	+	0	-	-	+	-	-	-	0	0	-	-
Product markets: remove distortion in traded goods markets (assume no devaluation; demand management keeps nontraded good price constant)	?	+	-	-	+	-	-	-	+	+	-	-
Exchange rate market: devaluation:												
(1) No other change	?	-	+	-	0	0	0	0	?	-	+	-
(2) Increased preference for traded goods (assume fixed or no distortions)	?	-	+	-	-	+	+	-	?	-	+	-
Monetary policy: monetary expansion:												
(1) Fixed exchange rates	?	+	-	+	+	-	-	+	-	+	-	+
(2) Flexible rates (assume fixed or no distortions)	0	0	0	0	0	0	0	0	0	0	0	0
Capital account: sustained constant addition to capital stock in each period	?	+	-	+	+	-	-	+	-	+	-	+

Note: All prices are real, rather than nominal values. This analysis assumes that the nontraded good is labor-intensive and that the traded good is capital-intensive.

Source: Lal (1986a).

REFERENCES

- Backus, D., and J. Driffill. 1985a. "Credibility and Commitment in Economic Policy." Centre for Economic Policy Research Discussion Paper 63. London. Processed.
- . 1985b. "Inflation and Reputation." *American Economic Review* 75, no. 3 (June): 530–38.
- . 1985c. "Rational Expectations and Policy Credibility Following a Change in Regime." *Review of Economic Studies* 52, no. 169 (April): 211–22.
- Barro, R., and D. Gordon. 1983a. "A Positive Theory of Monetary Policy in a Natural Rate Model." *Journal of Political Economy* 91, no. 4 (August): 589–610.
- . 1983b. "Rules, Discretion, and Reputation in a Model of Monetary Policy." *Journal of Monetary Economics* 12 (July): 101–21.
- Becker, G. 1983. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98, no. 3 (August): 371–400.
- Bhagwati, J. 1979. *Anatomy and Consequences of Trade Control Regimes*. New York: National Bureau of Economic Research.
- Black, S. 1976. *Exchange Rate Policies for Less Developed Countries in a World of Floating Rates*. Princeton University, Essays in International Finance 119. Princeton, N.J.
- Blejer, M., and M. Khan. 1984. "Government Policy and Private Investment in Developing Countries." *IMF Staff Papers* 31, no. 2 (June): 379–403.
- Brecher, R. A., and C. F. Diaz-Alejandro. 1977. "Tariffs, Foreign Capital, and Immiserising Growth." *Journal of International Economics* 7, no. 4 (November): 317–22.
- Burgess, D. F. 1980. "Protection, Real Wages, and the Neo-classical Ambiguity with Inter-industry Flows." *Journal of Political Economy* 88, no. 4 (August): 783–802.
- Calvo, G. A. 1978. "On the Time Consistency of Optimal Policy in a Monetary Economy." *Econometrica* 46, no. 6 (November): 1411–28.
- . 1986. "Fractured Liberalism: Argentina under Martinez de Hoz." *Economic Development and Cultural Change* 34, no. 3 (April): 511–34.
- Collier, P., D. Bevan, and J. Gunning. 1985. "The Macro Economics of Liberalisation—with an Application to East Africa." London: Trade Policy Research Centre. Processed.
- Congdon, T. 1985. *Economic Liberalism in the Cone of Latin America*. Thames Essays 40. London: Trade Policy Research Centre.
- Cooper, R., and J. Sachs. 1985. "Borrowing Abroad: The Debtors' Perspective." In G. W. Smith and J. T. Cuddington, eds., *International Debt and the Developing Countries*. Washington, D.C.: World Bank.
- Corbo, V., and J. de Melo. 1985. "Overview and Summary: Liberalisation and Stabilisation in the Southern Cone of Latin America." *World Development* 13, no. 8 (August): 863–66.
- Corden, W. M. 1977. *Inflation, Exchange Rates, and the World Economy*. London: Oxford University Press.
- Corden, W. M. and R. Jones. 1976. "Devaluation, Non-Flexible Prices, and the Trade Balance for a Small Country." *Canadian Journal of Economics* 9, no. 1 (February): 150–61.
- Corden, W. M., and P. Neary. 1982. "Booming Sector and De-Industrialisation in A Small Open Economy." *Economic Journal* 92, no. 368 (December): 825–48.

- De Jasy, A. 1984. *The State*. Oxford: Blackwell.
- Dornbusch, R. 1974. "Real and Monetary Aspects of the Effects of Exchange Rate Changes." In R. Z. Aliber, ed., *National Monetary Policies and the International Finance System*. Chicago: University of Chicago Press.
- . 1980. *Open Economy Macroeconomics*. New York: Basic Books.
- . 1985. "External Debt, Budget Deficits, and Disequilibrium Exchange Rates." In G. W. Smith and J. T. Cuddington, eds., *International Debt and the Developing Countries*. Washington, D.C.: World Bank.
- Eaton, J., M. Gersowitz, and J. Stiglitz. 1986. "The Pure Theory of Country Risk." *European Economic Review* 30 (June): 481–513.
- Economic Development and Cultural Change*. 1986. April.
- Edwards, S. 1984. *The Order of Liberalization of the External Sector in Developing Countries*. Essays in International Finance 156, Princeton, N.J.: Princeton University Press.
- Elster, J. 1979. *Ulysses and the Sirens—Studies on Rationality and Irrationality*. Cambridge: Cambridge University Press.
- Fellner, W. 1976. *Towards a Reconstruction of Macroeconomics: Problems of Theory and Policy*. Washington, D.C.: American Enterprise Institute.
- . 1979. "The Credibility Effect and Rational Expectations: Implications of the Gramlich Study." *Brookings Papers on Economic Activity* 1: 167–78.
- Fernandez, R. B. 1985. "The Expectations Management Approach to Stabilization in Argentina During 1976–82." *World Development* 13, no. 8 (August): 871–92.
- Findlay, R. 1986. "Trade, Development and the State." Economic Growth Center Twenty-fifth Anniversary Symposium on "The State of Development Economics: Progress and Perspectives." April, Yale University, New Haven, Conn. Processed.
- Fischer, S. 1980. "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government." *Journal of Economic Dynamics and Control* 2 (February): 93–107.
- Frenkel, J. 1982. "The Order of Economic Liberalization: Discussion." In K. Brunner and A. H. Meltzer, eds., *Economic Policy in a World of Change*. Amsterdam: North-Holland.
- Frey, Bruno D. 1978. *Modern Political Economy*. Oxford: Martin Robertson.
- Fry, M. J. 1980. "Money, Interest, Inflation and Growth in Turkey." *Journal of Monetary Economics* 6: 535–45.
- Gelb, A. 1986. "The Oil Syndrome: Adjustment to Windfall Gains in Oil-Exporting Countries." In D. Lal and M. Wolf, eds., *Stagflation, Savings, and the State: Perspectives on the Global Economy*. New York: Oxford University Press.
- Guitian, M. 1981. *Conditionality: Access to Fund Resources*. Washington, D.C.: International Monetary Fund.
- Hanson, J., and J. de Melo. 1985. "External Shocks, Financial Reforms, and Stabilisation Attempts in Uruguay During 1974–83." *World Development* 13, no. 8 (August): 917–39.
- Harberger, A. C., ed. 1984. *World Economic Growth*. San Francisco: Institute for Contemporary Studies.
- . 1985. "Lessons for Debtor-Country Managers and Policymakers." G. W. Smith and J. T. Cuddington, eds., *International Debt and the Developing Countries*. Washington, D.C.: World Bank.

- Jones, R. 1971. "A Three Factor Model in Theory, Trade and History." In J. Bhagwati, R. W. Jones, R. A. Mundell, and J. Vanek, *Trade, The Balance of Payments and Growth*. Amsterdam: North-Holland.
- Killick, T., ed. 1982. *Adjustment and Financing in the Developing World*. Washington, D.C.: International Monetary Fund.
- Kreps, D., and R. Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27, no. 2: 253-79.
- Krueger, A. O. 1974. "Home Goods and Money in Exchange Rate Adjustments." In W. Sellekaerts, ed., *International Trade and Finance*. New York: Macmillan.
- . 1978. *Liberalization Attempts and Consequences*. New York: National Bureau of Economic Research.
- . 1984. "Problems of Liberalization." In A. C. Harberger, ed., *World Economic Growth*. San Francisco: Institute for Contemporary Studies.
- Kydland, F. E., and E. C. Prescott. 1977. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85, no. 3 (June): 473-92.
- Lal, D. 1974. *Methods of Project Analysis: A Review*. World Bank Occasional Papers 16, Baltimore, Md.: Johns Hopkins University Press.
- . 1980. *A Liberal International Economic Order: The International Monetary System and Economic Development*. Princeton University Essays in International Finance 139. Princeton, N. J.
- . 1983. *Real Wages and Exchange Rates in the Philippines, 1956-78: An Application of the Stolper-Samuelson-Rybczynski Model of Trade*. World Bank Staff Working Paper 604. Washington, D.C.
- . 1984a. "The Political Economy of the Predatory State." World Bank Development Research Department Discussion Paper 105. Washington, D.C. Processed.
- . 1984b. *The Real Effects of Stabilization and Structural Adjustment Policies: An Extension of the Australian Adjustment Model*. World Bank Staff Working Paper 636. Washington, D. C.
- . 1985. "The Real Exchange Rate, Capital Inflows and Inflation: Sri Lanka 1970-1982." *Weltwirtschaftliches Archiv*, no. 4 (December): 682-702.
- . 1986a. "A Simple Framework for Analyzing Various Real Aspects of Stabilization and Structural Adjustment Policies." World Bank Economics and Research Staff Research Unit. Processed.
- . 1986b. "Stolper-Samuelson-Rybczynski in the Pacific: Real Wages and Exchange Rates in the Philippines, 1956-78." *Journal of Development Economics* (April): 181-204.
- Lal, D., and M. Wolf, eds. 1986. *Stagflation, Savings, and the State: Perspectives on the Global Economy*. New York: Oxford University Press.
- Lal, D., and S. Rajapatirana. Forthcoming. *Trade Liberalisation in Sri Lanka*. Thames Essay. London: Trade Policy Research Centre.
- Lindbeck, A. 1976. "Stabilization in Open Economies with Endogenous Politicians." *American Economic Review* 66, no. 2 (May): 1-19.
- McKinnon, R. 1981. "Financial Repression and the Liberalization Problem within Less Developed Countries." In S. Grassman and E. Lundberg, eds., *The World Economic Order: Past and Prospects*. New York: Macmillan.
- . 1982. "The Order of Economic Liberalisation: Lessons from Chile and Argentina." In K. Brunner and A. Meltzer, eds., *Economic Policy in a World of Change*. Amsterdam: North-Holland.

- McKinnon, R., and D. J. Mathieson. 1981. *How to Manage a Repressed Economy*. Princeton University Essays in International Finance 145. Princeton, N.J.
- Mussa, M. 1974. "Tariffs and the Distribution of Income." *Journal of Political Economy*, 82, no. 6 (November/December): 1191-1204.
- Neary, J. P. 1978. "Short-run Capital Specificity and the Pure Theory of International Trade." *Economic Journal* 88, no. 351: 488-510.
- . 1982. "Capital Mobility, Wage Stickiness and Adjustment Assistance." In J. Bhagwati, ed., *Import Competition and Response*. New York: National Bureau of Economic Research.
- . 1985. *Theory and Policy Adjustment in an Open Economy*. Centre for Economic Policy Research Discussion Paper 61. London.
- Nogués, J. 1981. "Política Comercial y Cambiara: Una Evaluación Cuantitativa de la Política Argentina Durante 1961-1981." Buenos Aires: Banco Central de la República Argentina. Processed.
- Nordhaus, W. 1975. "The Political Business Cycle." *Review of Economic Studies* 42, no. 130 (April): 169-90.
- Prachowny, M. 1981. "Sectoral Conflict over Stabilisation Policies in Small Open Economies." *Economic Journal* 91, no. 363 (September): 671-84.
- . 1984. *Macroeconomic Analysis for Small Open Economies*. Oxford: Clarendon Press.
- Sargent, T. J. 1985. "The Ends of Four Big Inflation." In T. J. Sargent, *Rational Expectations and Inflation*. New York: Harper and Row.
- Shapiro, Carl. 1983. "Premium for High Quality Products as Returns to Reputation." *Quarterly Journal of Economics* 98, no. 4: 659-80.
- Srinivasan, T. N. 1985. "Neo-Classical Political Economy, the State and Economic Development." *Asian Development Review* 3, no. 2: 38-58.
- Staniland, M. 1985. *What Is Political Economy?* New Haven, Conn.: Yale University Press.
- Stern, E. 1983. "World Bank Financing of Structural Adjustment." In J. Williamson, ed., *IMF Conditionality*. Washington, D.C.: Institute for International Economics.
- Stockman, A. C. 1982. "The Order of Economic Liberalisation: Lessons from Chile and Argentina—a Comment." In K. Brunner and A. H. Meltzer, eds., *Economic Policy in a World of Change*. Amsterdam: North-Holland.
- Swamy, G. 1981. *International Migrant Workers' Remittances: Issues and Prospects*. World Bank Staff Working Paper 481. Washington, D.C.
- Williamson, J., ed. 1983. *IMF Conditionality*. Washington, D.C.: Institute for International Economics.
- World Development*. 1985. August.

Sites and Services—and Subsidies: The Economics of Low-Cost Housing in Developing Countries

Stephen K. Mayo and David J. Gross

Sites and services projects represent a major innovation in shelter policy in developing countries and have been sponsored by international aid agencies for somewhat more than a decade. Such government projects deliver a package of shelter-related services, the standards of which depend on the ability and willingness to pay of intended beneficiaries. Typically, such projects represent a sharp break with preexisting government shelter policies in that they attempt, in principle, to focus directly on lower-income groups and to deliver shelter and services with small or no subsidies. This article describes the background of the sites and services concept; reviews recent evaluations of sites and services projects; presents an analytical model of the sites and services paradigm (which is used to examine how major project outcomes are influenced by project design); summarizes recent research on housing demand in developing countries (which is relevant to designing appropriate sites and services projects); reviews planning assumptions used in World Bank sites and services projects; compares these assumptions with empirical evidence on willingness to pay for housing; and examines project experience in light of contrasts between actual planning assumptions and empirical research on demand for shelter. The article concludes with suggestions for ways to improve the project design process and reform housing sector policies to increase the efficacy of the sites and services paradigm.

This is an article about a major innovation in delivering shelter and related services to the poor in developing countries: “sites and services” projects. Sites and services projects are government-sponsored packages of shelter related services, which range from a minimal level of “surveyed plot” to an intermediate level of “serviced sites” to an upper level of “core housing” complete with utilities and access to community-based services. The level of services depends on the ability and willingness of beneficiary populations to afford them. Typically, such projects represent a sharp break with preexisting government shelter policies in

Stephen K. Mayo is at the World Bank. David J. Gross, a consultant to the World Bank, is at Abt Associates. Comments and suggestions by Alain Bertaud, Anthony Churchill, Michael Cohen, Emmanuel Jimenez, Stephen Malpezzi, Barry Pinsky, Bertrand Renaud, Evan Rotner, Paul Strassman, Richard Stren, Richard Westin, and James Wright regarding various drafts are gratefully acknowledged.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

that they attempt, in principle, to focus directly on lower-income groups and to deliver shelter and services with small or no subsidies.

The earliest sites and services schemes (in Chile, Kenya, and the Union of South Africa) were undertaken in the 1940s and 1950s, largely without external assistance. It was not until the late 1960s and early 1970s that international agency assistance began in earnest. Large numbers of sites and services and slum upgrading projects have since been implemented. The World Bank alone, for example, initiated sixty-eight projects through 1984, each benefiting more than 25,000 households on average. After a decade of experience with such projects, it is useful to look closely at how they have worked and to determine what can be learned from their successes and failures.

After we describe the background of the sites and services concept and review some recent evaluations and critiques of these projects, we will present the results of an analytical model of how the provision of sites and services is likely to influence household behavior. Relevant areas of impact are housing consumption and investment, consumption of nonhousing goods and services, incentives to sublet or to sell out to higher-income groups, and incentives to default on mortgages. In addition, the analytical framework provides a basis for estimating the economic benefits that accrue to project participants, the subsidies that are required to achieve selected project outcomes (such as housing that meets minimum project design standards), and the welfare losses to society associated with subsidy elements.

In a period in which most developing country governments face severe budget constraints, their ability to subsidize housing on a wide scale is limited. If subsidies are fiscally possible, and are desired to serve equity, efficiency, or project impact criteria, they should be explicitly set and rationalized to allow for full budgetary control. We approach the issue of pricing sites and services projects with an understanding that for many governments such projects will only be feasible if they involve minimal or no subsidies.

The analytical framework indicates that having a firm grasp of housing demand parameters in developing countries is critical for understanding project effects. After discussing the analytical model, we review recent evidence on housing demand in developing countries based on ongoing research projects at the World Bank. Implications of that research for sites and services projects are then described. This discussion is followed by a review of several major project outcomes of sixty-eight World Bank sponsored sites and services projects. That review suggests that failure to incorporate correct information on housing demand behavior into project design has resulted in problems which seriously threaten some of the fundamental goals of sites and services projects. The most serious of these are conditions that lead to either large subsidies, budgetary constraints, and hence poor prospects for widespread replication or to smaller subsidies with benefits going to relatively higher-income groups. The final section of the article presents conclusions and some recommendations for improv-

ing both the design of sites and services projects and the choice of housing policy instruments in developing countries.

I. SITES AND SERVICES PROJECTS: BACKGROUND AND RECENT EVALUATIONS

The growth of developing country populations and their increasing concentration in urban areas has put enormous pressure on governments to mobilize resources to meet the basic needs of their people. Shelter demands have been particularly acute. The initial response of many developing country governments to the press of urban population growth was to adopt the shelter "solutions" of developed countries: heavily subsidized blocks of public housing flats with high standards of construction and infrastructure, zoning and building code regulations discouraging production of lower standard housing, and in many cases, destruction of slum areas and squatter settlements in the name of either "law and order" or "urban renewal." These policies did not work.

Public housing did not reach most of the population; Grimes (1976) reported that from one-third to two-thirds of urban populations could not afford the least expensive public housing unit in the six developing country cities he studied. Subsidy levels were high enough to create serious budgetary pressures that virtually guaranteed that public housing could not be made available to most of the population. Zoning and building regulations were widely flouted as "informal housing" and squatter settlements proliferated. In Cairo, for example, where zoning laws forbid conversion of agricultural land to residential use and where building codes require approved architectural or engineering plans for legal housing construction, it was estimated that approximately 84 percent of housing built since 1970 is "informal"—built in violation of either zoning or building codes or both (Abt Associates 1982, which is available from Mayo). When squatter settlements were removed, hydralike they returned, sometimes nearby but often in the same place.

By the 1970s, Turner (1972), Mangin (1970), and others espoused an approach more in harmony with the natural processes of shelter acquisition and development of the poor themselves. It was proposed that public programs capitalize on the untapped energies and resources of the poor through "progressive development" schemes which simply serviced housing sites or provided housing that was affordable by low-to-moderate-income households and which could be progressively upgraded over time. Complementing such sites and services schemes, there were to be slum upgrading or squatter upgrading schemes that focused on improving existing residential areas of the poor rather than on developing undeveloped land.

The key to making such projects work was to bring down the cost of shelter and infrastructure from the high and unaffordable levels prescribed by most governments, as was noted in a World Bank policy paper (1975, p. 5): "This can be achieved, in the first instance and most rapidly, by reducing standards; per-

mitting and encouraging the use of low-cost (frequently indigenous) building materials and a lower quality of finish; providing communal rather than private plumbing and sanitary facilities; encouraging or providing higher density construction, with less land per dwelling unit; and providing less living space per dwelling unit."

It was also seen to be important to deal with problems which were assumed to restrict the level of housing investment among low-income households. These included shortages of urban infrastructure, uncertain and insecure land tenure, lack of formal housing finance, shortages of low-cost building materials, and difficulty in assembling land for development. Removal of such supply-side barriers to investment was seen as being essential to mobilize household resources for upgrading shelter and infrastructure beyond initial affordable but minimal levels.

It was also expected that public institutions responsible for implementing sites and services schemes and upgrading projects would price sites and shelter at a level that would not only fully recover costs but would also generate a modest surplus that could be recycled to replicate projects on a larger scale. While subsidies, particularly for the lowest-income participants, were not altogether ruled out of the sites and services paradigm, it was assumed that they would be modest, often generated internally by allocating profits earned on sales of higher-income household and commercial sites to reduce prices of sites for lower-income households. Thus, the important goal of "replicability" of projects was seen to depend critically on both "appropriate standards and sound pricing policies" (World Bank 1983, p. 9). This implied that prices should be sufficient to recover costs and entail small or no subsidies, and that the housing and infrastructure provided were both affordable and were those for which poor households were willing to pay.

These principles were reflected in the 1970s and 1980s in a growing number of sites and services and upgrading projects sponsored by developing country governments and funded in part by outside agencies such as the World Bank, the U.S. Agency for International Development (USAID), regional development banks (for example, the Asian Development Bank and Inter-American Development Bank), and other bilateral aid agencies. The World Bank alone made loans for 36 sites and services or upgrading projects between 1972 and 1981, which involved total Bank lending of more than \$1 billion and were estimated to have benefited nearly 2 million people (World Bank 1983, p. 46). Under its Housing Guarantee Program, USAID sponsored preparation of a number of similar schemes, many of them in Latin America (World Bank 1975).

Evaluations of such projects are beginning to emerge. Citing examples of policy changes consistent with the sites and services approach in a number of countries, Ayres (1983, pp. 176-77) concluded: "The most obvious manifestation of the [policy] changes is held to be that publicly constructed housing, the model of the 1950s and 1960s, has given way to private investment through self-help, thereby reducing the role of the public sector."

In addition to the effect of the sites and services approach on policy, Ayres noted that the impact "on project design, planning and investment programming is more measurable. Changes in design standards have brought shelter costs way down" (p. 177). In Zambia, for example, complete houses in sites and services projects were estimated to cost less than one-fifth as much as the least expensive government-subsidized housing; in El Salvador, the better quality sites and services project houses cost less than half as much as the cheapest conventional house (Keare and Parris 1982, p. xiv).

In a recent evaluation study of four early World Bank projects jointly sponsored by the Bank and the International Development Research Center (IDRC), Keare and Parris (1982, p. v), stated: "The experiment embodied in the first generation of Bank-supported urban shelter projects [has] been remarkably successful." As evidence, they cited (1) increased production of housing and infrastructure, (2) construction of higher quality housing than had been expected, (3) continued investments by beneficiaries in both housing and community facilities, (4) allocation of plots as low as the twentieth income percentile, (5) residential turnover no greater than among control group households, (6) "affordability" among target groups, and (7) generation of income and employment among project beneficiaries and producers of sites and services.

These accomplishments were achieved only at the expense of "hard and sometimes bitter arguments . . . to persuade borrowing nations to reduce their standards and costs, [and] to increase the amounts that were charged to project participants." The effect was that available public resources were spread far more broadly among the population than had been the case before the advent of sites and services and upgrading projects (Tym 1984, p. 217).

Nevertheless, a number of reservations and criticisms have been expressed about sites and services, including claims that (1) shelter provided is not affordable by the poor, (2) benefits accrue disproportionately to better-off households, (3) cost recovery is poor, (4) subsidies persist at unsustainable levels, and, because of the above, (5) it is not budgetarily possible to reproduce projects, as they are now designed and implemented, on a large scale.

Cost recovery has been noted as a problem by several reviewers, although evidence suggests that the experience has been mixed. While Keare and Parris (p. xii) cite cost recovery as a problem in three out of the four early World Bank projects they review, it appears that cost recovery in more recent projects is a less serious problem.

Evidence that projects have not proved affordable, either for the target residents, or, in the longer term, for the government, is less ambiguous. For example, Ayres asserts that affordability assumptions on which project designs were based were "over optimistic" (p. 193); Payer (1982, pp. 336-37), that they were "highly unrealistic" and resulted in slow housing consolidation, inability of some households to pay, and other problems such as the leakage of benefits to better-off households and poor cost recovery. Incidences of benefit leakage have been noted by Chana (1984), Cohen (1983), and Keare and Parris (1982). Chana, for

example, in reviewing the experience of a sites and services project at Dandora in Nairobi, Kenya, notes that about half of all occupied plots were rented out fully to non-allottees and some units had been sold, presumably to members of higher-income groups (p. 52). Keare and Parris found that sites and services participants "tend to be more representative of median income groups than of the poorest urban households" (p. 12).

In some cases, however, affordability seems to have been purchased with subsidies that are considerably higher than the minimal levels envisioned by the original architects of the sites and services paradigm. Land price writedowns and interest rate subsidies are common in sites and services projects and can constitute from 50 to 90 percent of the true resource cost of the shelter or land provided in a project. In a project in the Arab Republic of Egypt cofinanced by the British Overseas Development Agency, for example, land was priced at LE2.25 per square meter while the market price was LE10–LE15, and a 5 percent interest rate was charged when the prevailing rate was about 11 percent (Davidson 1984, p. 142). And despite the fact that many subsidy elements are implicit off-budget transfers, they imperil the long-term replicability of projects just as surely as do direct government payments to project beneficiaries.

Lowering standards and reducing subsidies has often proven to be extremely difficult for designers of sites and services projects. Little thought appears to be given in project planning to subsidies: their form, magnitude, distribution among different sorts of households, or their impact on either the behavior of beneficiaries (on housing consumption, investment, or tenure choice) or the viability of subsidy-generating institutions. In most developed countries with significant social housing policies, such questions are at the heart of policy and technical debates concerning the choice of housing programs and policies; in most developing countries they are ignored.

To meet this challenge means, among other things, having a clearer analytical understanding of how sites and services projects affect the behavior of potential beneficiaries and how one can modify project features in order to produce projects that are both affordable and replicable.

II. AN ANALYTICAL MODEL OF SITES AND SERVICES PROJECTS

This section briefly summarizes the major conclusions derived from an analytical model of sites and services projects that is presented in detail in Mayo and Gross (1985). That model is based on the standard neoclassical economic paradigm and explores the ways in which household behavior is likely to be influenced when households are given opportunities to participate in typical sites and services projects.

The model has the following stylized features:

1. The project designer picks a target income group whose income is y_0 .
2. An assumption is made concerning the fraction of income households are

- willing and able to pay for shelter. This fraction is a , the affordability ratio; it is sometimes assumed to vary for different target groups.
3. A target housing expenditure level (rent) is set; $R_s = ay_0$.
 4. Project standards such as lot size, location, infrastructure type and quality, and shelter size and quality are established such that units are built with a cost, C , set just equal to R_s .

Such units are then offered to target-income-group households at a price, R_s , which fully recovers costs. The amount of housing offered is equal to H_s , which is assumed to be greater than the amount originally consumed by the household (H_0), but it is offered at a lower price per unit (p_s) than the market price, p_b . The project price may be lower than the market price for a number of reasons, many having to do with removing supply-side constraints. For example, the project may offer long-term housing finance at a rate of interest below that of the informal money market (yet still at a level sufficient to cover the borrowing costs of the lending institution); water may be supplied at a price less than the cost from private vendors; secure tenure can be offered at less than the differential cost of securing legal housing in the marketplace.¹

The economic benefit to households of participating in the project depends on the amount of shelter and services received and its price. Once having entered the project, incentives to households to sell out, sublet, or default are determined by households' preferences, income fluctuations, market prices of units provided in the project, and rules and restrictions governing the ability of households to profit by selling out or subletting.

The model analyzes the complex incentives and disincentives that are set in motion when the standard sites and services paradigm is implemented. Among the more important findings are the following.

1. Economic benefits result from lowering the effective market price of shelter and services, relaxing market restrictions, or both. The magnitude of potential benefits depends on the household preferences for housing vis-à-vis other goods, on housing market features (especially the existence and magnitude of market imperfections and expectations concerning inflation in prices and income), and on project design features.
2. Potential benefits are determined by the standards of shelter and services, prices charged, the type and level of subsidies, and whether subletting and resale are permitted.
3. These factors affect consumption (and investment) in shelter and other goods, and incentives to invest further in project-provided units, to resell or sublet, and to default on project-related loans.
4. "Affordability" criteria for project design and pricing are of particular importance. Our earlier paper (Mayo and Gross 1985) indicates that there is

1. For a model of the "cost" of tenure security to urban households, see Jimenez (1984) and Friedman, Jimenez, and Mayo (1985).

some arbitrariness in establishing the percentage of income households are willing to spend on housing and hence in setting design standards so that incentives for household participation vary among households and projects. Some potential beneficiaries, particularly those with strong preferences for housing, will not only be induced to participate in the absence of subsidies but will also have incentives to upgrade project units. For many potential beneficiaries, however, participation requires subsidies, and benefits will be lower than is true for others. Consequently such households may have significant incentives to sell out to higher-income groups, to sublet part of their space, and to default. If standards are too high because effective demand has been overestimated, project problems such as high default rates, high turnover in favor of higher-income households, and slow project upgrading will occur. However, large subsidies and strong self-selection by beneficiaries (the latter particularly a problem when project scale is small) can disguise the existence of project problems. Thus, entirely "successful" small-scale projects with hidden subsidies can be implemented, but they would generate budgetary problems if replicated on a large scale.

III. AFFORDABILITY CRITERIA: HOUSING DEMAND, SUBSIDY, AND DESIGN

The previous section suggests the importance of getting good estimates of the effective demand for housing if sites and services projects are to be designed appropriately. Here we review recent research on household demand for housing in developing countries and present some of the major implications of that research for both the design and pricing of and the target groups to be reached by sites and services projects.

Housing demand

Until recently, few comparative studies of housing demand in developing countries existed. Most such studies were based on specialized data bases, which were not usually collected to estimate housing demand relationships.² Different analyses used different variable definitions, functional forms, and stratification variables, which made the comparison of results exceedingly tenuous.

In 1981, a major comparative study of housing demand in developing countries was initiated at the World Bank. In that analysis, high-quality data were collected for sixteen cities in eight countries (Colombia, Egypt, El Salvador, Ghana, India, Jamaica, the Republic of Korea, and the Philippines) and were used to estimate housing demand relationships using relatively comparable definitions of variables and identical functional forms and stratifying variables. For comparative purposes, identical models were estimated for two U.S. cities, Pittsburgh and Phoenix. A description of much of that research is contained in Malpezzi, Mayo, and Gross (1985).

2. For a review of the pre-1984 literature, see Malpezzi, Mayo, and Gross (1985).

In that study, a simple log-linear model of housing expenditures in each of the sixteen cities is estimated:

$$(1) \quad \ln R = b + \epsilon_y (\ln y) + cH + dH^2 + u$$

where R is rent; ϵ_y is the estimated income elasticity of demand; y is income; H is household size; b , c , and d are regression coefficients; and u is an estimated disturbance term. The model is stratified for renters and owners. For renters, rent is defined as net rent, exclusive of utility payments. For owners, rent is defined variously, and in order of availability, as (1) owner imputations of net rent, (2) estimates applying renter-based hedonic price equations to owners' housing characteristics,³ or (3) the product applying a fixed amortization ratio (from 1 to 1.5 percent per month depending on the country) to owners' estimates of housing value. While other functional forms were tried, and many other demographic variables were included in alternative estimating equations, results from the simple log-linear model were found to provide adequate fits and robust findings regarding major demand parameters.

Table 1 presents the estimated parameters of housing expenditure functions for renters and owners.⁴ In general the results are remarkably consistent with results from developed countries (see Mayo 1981). The regression fits are typical for this type of equation: typical R -squared statistics are in the 0.1–0.3 range (the minimum is 0.06; the maximum, 0.57). Fits are similar for owners and renters.

The median of all renters' income elasticities is 0.49; developing country elasticities range from 0.31 (Busan, Korea) to 0.88 (Davao, the Philippines). Most cluster between 0.4 and 0.6, with estimated U.S. elasticities, however, lower than developing country estimates. The median of all point estimates of owner income elasticities is 0.46, with extremes of 0.17 in Cairo and 1.11 in Santa Ana, El Salvador. Half of the point estimates lie between 0.4 and 0.7. For nine of the fourteen cases for which comparison is possible, estimated developing country owner income elasticities are greater than those of renters; this parallels findings in the literature for developed countries (Mayo 1981).

Comparing expenditure equations across countries reveals practically no systematic variation of income elasticities (ϵ_y) with country or city income level or population size but considerable variation in dollar-adjusted intercepts (b in equation 1) which are positively related to average city income. Rent-income ratios therefore decline systematically as income rises *within* cities but increase with average income *across* cities. It should also be noted that the relative variation in rent-income ratios is higher at low-income levels than at high-

3. The hedonic price equation statistically relates rent to housing characteristics. The coefficients of a simple hedonic price equation, $R = f(Z)$, (where R is rent and Z is housing features) would provide estimates of the rental "price" of each of the characteristics studied.

4. Malpezzi, Mayo, and Gross (1985) discuss findings concerning household size; here we focus only on income elasticity and intercept estimates.

Table 1. *Estimated Parameters of Housing Expenditure Functions*

Country/city	Renters					Owners				
	Constant	Log income ^a	Household size ^a		R Squared / N ^a	Constant	Log income ^a	Household size ^a		R Squared / N ^a
			H	H ²				H	H ²	
<i>Colombia</i>										
Bogota (1978)	1.11	0.66	0.09	-0.006	0.04	0.77	0.75	0.00	-0.003	0.49
		0.03	0.03	0.003	1,016		0.03	0.04	0.003	821
Cali (1978)	2.81	0.44	0.13	-0.006	0.27	1.25	0.69	-0.05	0.000	0.38
		0.06	0.07	0.007	257		0.06	0.07	0.005	256
<i>Egypt</i>										
Cairo (1981)	0.25	0.46	-0.17	0.010	0.16	0.89	0.17	0.12	-0.009	0.06
		0.06	0.09	0.008	303		0.12	0.21	0.019	76
Beni Suef (1980)	-1.20	0.51	0.38	-0.047	0.25	-0.09	0.42	0.14	-0.003	0.23
		0.14	0.28	0.029	63		0.13	0.14	0.010	63
<i>El Salvador</i>										
Santa Ana (1980)	0.37	0.48	0.13	-0.014	0.16	-2.50	1.11	-0.06	-0.004	0.37
		0.11	0.08	0.007	131		0.11	0.12	0.009	169
Sonsonate (1980)	0.79	0.50	-0.10	0.007	0.16	0.39	0.79	-0.13	0.001	0.57
		0.12	0.09	0.007	83		0.15	0.17	0.012	27
<i>Ghana</i>										
Kumasi (1980)	0.82	0.33	0.02	0.000	0.11	—	—	—	—	—
		0.04	0.03	0.002	814		—	—	—	—
<i>India</i>										
Bangalore (1975)	0.66	0.58	-0.08	0.003	0.18	2.84	0.43	-0.17	0.007	0.15
		0.04	0.04	0.002	1,041		0.08	0.06	0.004	205

<i>Jamaica</i>										
Kingston (1975)	-0.12	0.70	0.16	-0.012	0.30	—	—	—	—	—
		0.08	0.07	0.007	223		—	—	—	—
<i>Korea</i>										
Seoul (1979)	5.04	0.45	0.07	-0.004	0.15	6.06	0.44	-0.04	0.002	0.12
		0.03	0.04	0.005	952		0.04	0.04	0.003	952
Busan (1979)	6.26	0.31	0.05	-0.001	0.08	5.93	0.45	-0.05	0.002	0.10
		0.07	0.06	0.006	508		0.08	0.10	0.011	296
Taegu (1979)	4.95	0.44	0.03	-0.003	0.23	6.32	0.47	-0.19	0.011	0.18
		0.07	0.07	0.008	292		0.08	0.08	0.006	152
Kwangju (1979)	2.70	0.62	0.09	-0.002	0.32	7.53	0.41	-0.27	0.018	0.14
		0.09	0.13	0.014	134		0.11	0.18	0.016	84
Other (1979)	3.33	0.54	0.04	0.002	0.17	2.16	0.79	-0.12	0.003	0.26
		0.05	0.05	0.007	1,000		0.05	0.05	0.005	779
<i>Philippines</i>										
Davao (1979)	-1.60	0.88	0.00	-0.002	0.42	-3.20	0.99	0.04	-0.004	0.28
		0.03	0.05	0.002	1,376		0.04	0.04	0.003	1,968
Manila (1983)	1.27	0.56	0.01	-0.002	0.22	2.46	0.57	-0.02	0.000	0.31
		0.04	0.04	0.003	605		0.04	0.05	0.003	390
<i>United States</i>										
Pittsburgh (1975)	3.07	0.26	-0.02	-0.002	0.15	3.50	0.18	0.08	-0.005	0.21
		0.02	0.04	0.005	946		0.01	0.02	0.002	2,378
Phoenix (1975)	3.68	0.18	0.12	-0.015	0.13	3.62	0.18	0.13	-0.011	0.24
		0.02	0.03	0.005	918		0.01	0.01	0.002	2,284

— Not available.

a. The first of each pair of numbers is the coefficient; the second, the standard error.

income levels, a result that has implications for the degree to which self-selection problems such as those noted in section II will occur within sites and services projects with high standards (see Malpezzi 1984 for evidence). These relationships are shown graphically in figure 1 for renters in four representative cities. Relationships for owners are similar, although average rent-income ratios are invariably higher; owners' housing consumption is roughly 80 percent greater than that of renters at given incomes within particular markets.

It is important to recognize factors which suggest that these income elasticities are lower-bound estimates. In developing countries in which poorer families would prefer to purchase only the most rudimentary shelter goods and services, the absence of such packages in commercial markets means that measures of demand for conventional housing will not include this element of housing demand. Sites and services projects, however, may provide such an optimal package to meet the effective demand of the poor. Similarly, lack of access to credit for housing purchases may reduce demand for housing as income grows to some relatively high threshold level of income. Typically, sites and services projects include just such small increments of housing credit. Further, data currently available do not allow us to capture the transfer of families from renter to owner status as household income rises. Yet there is sometimes a major increase in housing consumption with income because of tenure shifts—another element of income elasticity which is not incorporated into our model. All these factors suggest that our present measure of the benefits of sites and services projects must be regarded as to some degree understating the actual rewards to participants.

The relationships portrayed in figure 1 are very similar to the consumption patterns within and across countries documented by Kuznets (see Kuznets 1961 and other works cited therein). Qualitatively, housing consumption is remarkably similar to total consumption; that is, within-country differences in average rent-income ratios at different income levels within a country are markedly smaller than are between-country differences at various average income levels. Malpezzi, Mayo, and Gross (1985) explore alternative theoretical explanations for these results and then test a series of cross-country housing expenditure models, which are designed to analyze presumably long-run cross-country housing demand relationships. They also estimate models for a pooled sample of U.S. and developing country cities. The simplest cross-country model they test is similar to the log-linear within-country model described above, but with the addition of a price term, the relative price of housing, which is constructed using data from Kravis, Heston, and Summers (1982).

Following this approach, and defining p_o as the relative price of housing, we estimate models for renters and owners in developing countries.

For renters:

$$(2) \quad \ln R = -5.39 + 1.60 \ln y + 0.15 \ln p_o$$

(0.18) (0.29)

Figure 1. *Rent-Income Ratios by Income for Renters*

Key: A average for each city at its average income B Bogota C Cairo M Manila S Seoul

$$R^2 = 0.90$$

$$\text{d.f.} = 13$$

For owners:

$$(3) \quad \ln R = 3.57 + 1.38 \ln y + 0.65 \ln p_o$$

(0.35)
(0.50)

$$R^2 = 0.76$$

$$\text{d.f.} = 11$$

In both models, the dependent variable, rent, and income are city means and are converted to 1981 U.S. dollars.⁵

The implications of these models, which are confirmed with alternative specifications, are straightforward: in the very long run, housing consumption is income-elastic, a result that conforms to recent findings of Kravis, Heston, and Summers (1982). Price elasticities are smaller in absolute value than income elasticities, although confidence intervals are quite wide for the former. Long-run income elasticities are estimated to be higher for renters than owners. This means that as cities' economies develop over the very long run, owner and renter consumption patterns increase at a similar pace, all other things being equal. However, because relative housing prices rise with income (at least in our sample) and because renter price elasticities are estimated to be higher than owner elasticities, the net effect of both incomes and prices rising with economic development is that owners' consumption increases faster than renters' consumption over most of the range of the data.

These findings have important implications for the design of housing programs and, in particular, of sites and services projects. An obvious general rule is that the fraction of income allocated by households for housing is highly variable, depending on income, the level of an area's economic development, the relative price of housing, and tenure (whether residents own or rent). Thus it is inappropriate to use any single affordability ratio (housing expenditure to total income) as the basis for establishing design standards in sites and services projects. If a single value is used, particularly if it is higher than normal spending patterns would indicate, then subsidies might be required to induce target groups to participate and higher-income groups might find their way into projects, either initially or by purchasing from initial allottees.

In order to get a sense of how serious these problems might be, it is useful to examine (1) the minimum subsidy (or benefit) necessary to induce typical target groups to participate, and (2) the likely income levels of participants in the absence of subsidies, based on the empirical findings on housing demand presented above and assuming alternative affordability rules of thumb.

5. Note that in a log-linear expenditure equation the coefficient of price is equal to one plus the price elasticity; thus the price elasticity is the estimated coefficient minus one, or -0.85 for renters and -0.35 for owners.

Minimum Subsidies Necessary to Induce Participation

Mayo and Gross (1985) develop a general procedure for estimating minimum subsidies (benefits) necessary to induce participation by households whose normal housing equilibriums are below those implied by project standards. We have made that procedure operational by assuming a particular utility function, the parameters of which can be approximated using our empirical results, and have examined the impact on required subsidies as they vary with alternative affordability planning assumptions. (The application of this procedure is described in the appendix.)

Figure 2 presents the results of such a procedure based on the actual housing expenditure equations estimated in Malpezzi, Mayo, and Gross (1985) where, for simplicity, it was assumed that a within-city income elasticity of 0.5 and a corresponding price elasticity of -0.4 typified developing country cities, and where a typical target income group for sites and services projects is in the thirty-fifth percentile of the income distribution.⁶ Renters' rather than owners' expenditure functions are used here. The main reasons for this are (1) that often it is "homeless" or renter households that represent the designated sites and services project-target group; (2) owners' current consumption relative to current income reflects on average greater longevity and thus more chance to have upgraded housing services relative to renters; and (3) in some markets, owners' current housing consumption reflects both windfall price appreciation and possible overconsumption due to high transactions costs of moving. To the extent that renters' average propensities to consume housing increase after they become owners, all other things being equal, the estimates of minimum subsidies required for participation will be overstated.

The horizontal axis of figure 2 is monthly household income in 1981 U.S. dollars and the vertical axis is the minimum subsidy necessary to induce participation as a percentage of the market value of the housing provided in a project. The curves shown are subsidy levels required for various values of a , the design affordability ratio. For reference, estimated monthly household incomes in most African countries and countries on the Indian subcontinent were below \$100. Some of the countries with incomes between \$100 and \$200 were Botswana, Cameroon, Egypt, El Salvador, Indonesia, the Philippines, and Thailand; countries between \$200 and \$400 included a number of Latin American and North American countries and Nigeria and Zambia; and those above \$400 included Caribbean, Latin American, and East Asian countries such as the Bahamas, Brazil, Jamaica, Korea, Mexico, and Panama.

The figure indicates clearly the effect that project standards (as derived from design affordability levels) have on target group households' incentive to partici-

6. Our research indicated that on average, income at the thirty-fifth percentile was about 43 percent of city average income. The preproject rent, R_0 , was estimated in two steps, first at average city income using equation 2 and second using the within-city expenditure equation for $y = 0.43 \bar{y}$ (assuming $\epsilon_y = 0.5$) and calculating the corresponding value of R_0 .

Figure 2. Minimum Subsidies Necessary to Induce Participation of Thirty-Fifth Percentile Households at Alternative Values of the Design Affordability Ratio (a)

pate, and on the need to provide subsidies to induce participation when standards are set too high. For example, suppose that it were assumed that households in a typical African country, say Kenya, with a 1981 household income of roughly \$100 per month, would be willing to pay for a unit designed to cost 20 percent of income. According to figure 2, a subsidy of roughly 60 percent of the market value of such a unit would have to be provided in order to induce households in the thirty-fifth percentile of the income distribution to participate. In Burundi, with monthly household income of only about \$70 in 1981, a subsidy of roughly 85 percent would be required to induce thirty-fifth-percentile households to participate if the design standard is based on an affordability assumption of 20 percent of income. Subsidies of these levels are, of course, a reflection of the low average propensities to consume housing indicated by the cross-country expenditure functions presented above.

By contrast, in higher-income developing countries a 20 percent affordability standard may be entirely appropriate. For example, for countries (cities) with an average household monthly income above about \$175, subsidies of less than 20 percent would appear to be adequate to induce target groups to participate. Required subsidies are, however, extremely sensitive to the choice of design standards. While the difference between 20 and 25 percent of income may not sound like much to a project planner, such a difference represents a 25 percentage point difference in monthly shelter costs and can easily mean the difference between required subsidies in the range of 60–70 percent rather than the range of 20–25 percent. Depending on whether subsidies of the required magnitude are forthcoming or not, target income groups may not even participate or, if they do, they may have strong incentives to sell out to higher-income groups. It is useful to examine now the influence of planning standards on the income levels of participants if subsidies are not provided or, alternatively, if initial allottees sell out to higher-income households at the full market price.

*Income Levels of Participants at Zero-Subsidy or upon Resale
by Initial Occupants*

A general procedure for inferring the income level of households that would participate in a project and their place in the income distribution if they pay the full unsubsidized price of the package of services is outlined in the appendix. The expenditure functions that describe across-country and within-city behavior can be used to solve for the income level consistent with a stipulated project design standard and its associated market rent.

Figure 3 illustrates the estimated impact of alternative design standards on the income of households that would participate with no subsidies or that would be likely to purchase from original allottees.⁷ A within-city income elasticity of 0.5

7. The calculations underlying the figure are based on a generalized or average income distribution in developing countries that was estimated by fitting a fourth-degree polynomial to Lorenz curve data presented in Kakwani (1980) for a sample of thirty-three developing countries. Details are provided in Mayo and Gross (forthcoming).

Figure 3. *Income Percentile of Participating Households with No Subsidy at Alternative Design Affordability Ratios*

was assumed, and it was assumed that the design standard was based on the product of income of the thirty-fifth-percentile households and various assumed affordability levels, ranging from 10 to 30 percent of income.

The figure clearly illustrates the effect of alternative design standards of unsubsidized projects on the income of probable participants. Not only does increasing the design affordability ratio increase the income level of likely participants, but it does so with particularly dramatic effect at various thresholds. For example, for households in low-income countries (for example, $y = \$100$ per month), setting the design standards on the basis of an assumption that households are willing to spend 20 percent of income on housing implies that households in approximately the eightieth percentile of the income distribution could afford to participate without subsidies. Dropping the standard to one based on 15 percent of income has only a modest effect, inducing participation down to the sixty-fifth percentile in the absence of subsidies. Dropping the standard still further to one based on just 10 percent of income permits reaching even below the original target group, all the way down to the 15th percentile. Similar thresholds exist at each level of income, which suggests that dramatic improvements can be realized in the ability to reach the poor through sites and services projects by finding the "correct" design standard—the one that reflects true willingness to pay by low-income groups.⁸

Having examined the behavioral evidence on developing country housing demand and having explored some of the implications of that evidence for sites and services project planning, it is useful now to consider the actual planning practice and some of the project outcomes of typical sites and services schemes. Planning practice can be evaluated in terms of its consistency with external evidence on household behavior, and project outcomes can be interpreted in part in light of discrepancies between planning assumptions and actual behavior.

IV. DESIGN AND OUTCOMES OF SITES AND SERVICES PROJECTS

The previous sections have emphasized the important role played by planning parameters in influencing outcomes in sites and services projects. In this section we briefly review important aspects of planning practice in sixty-eight World Bank-financed sites and services projects, particularly affordability assumptions, and examine the consistency of planning parameters with external evidence on willingness to pay for housing. In addition, we look at the magnitude of subsidies that have been provided in a subset of Bank-financed sites and services projects—those which have been completely implemented with all loan

8. This is obviously a simplification in that the goals of projects are numerous and the choice of a design standard must reflect a number of compromises among the various program objectives. A highly simplified way of doing this might be to view the minimum required subsidy to reach the intended target group and the income group percentile likely to be attracted at unsubsidized prices as arguments in a limited social welfare function, with the design standard being chosen to maximize such a function.

amounts disbursed. The implications of these findings for achieving the nominal goals of sites and services are then evaluated.

Among the most important planning criteria in sites and services projects are the income levels of the principal target groups, the affordability ratio, the percentage downpayment, and the interest rate to be charged on project financing. Table 2 presents median values of each of these factors for sixty-eight World Bank-financed sites and services projects initiated between 1972 and 1984, along with median values of three macroeconomic variables that indicate general market conditions. The results are disaggregated by region, by the time period of project initiation, and by 1983 per capita gross domestic product (GDP).

Consider each project design variable in turn. First, the minimum income level intended to be reached by sites and services projects varies considerably among regions, ranging from a low of \$27 per month in South Asian projects to a high of \$141 in European, Middle Eastern, and North African projects. Much of this variation is a result of differences among regions in GDP per capita; the minimum income level of target groups ranges from \$40 per month in countries with per capita GDP less than \$600 per year to \$112 per month in countries with GDP per capita greater than \$1,200 per year. Planned minimum target group incomes have fallen slightly over time, which indicates a concern with reaching relatively farther down the income distribution. Median target groups (not shown here) tend to be at about the thirty-fifth percentile of the income distribution.

Downpayments in sites and services projects have generally been minimal. No region has had a median downpayment greater than 10 percent of the sales price. There are only weak positive associations between downpayment percentages and time and GDP per capita. Further investigation reveals that 79 percent of World Bank projects required downpayments of 10 percent or less. It might be noted that such low downpayment requirements reflect an implicit assumption that the ability to accumulate preproject assets through savings is negligible. In contrast to this, as is discussed below, project designers appear to assume that within-project savings propensities, in the form of mortgage loan repayments, are quite high. The two assumptions do not appear to be consistent.

Affordability ratios reflect assumptions concerning the fraction of income that households are willing to pay for shelter and related services. As table 2 indicates, there is strikingly little variation by region, time, or GDP per capita in the affordability ratios embodied in sites and services project designs. Medians for all substrata are from 20 to 25 percent. Further investigation reveals that 74 percent of all projects initiated by the World Bank have assumed affordability ratios of 20 to 25 percent, with 17 percent below those levels and 9 percent above. This comes close to the very definition of a rule of thumb.

Median interest rates charged for project-related loans have ranged from a low of 7 percent in Europe, the Middle East, and North Africa to a high of 12 percent in South Asia, East Asia, and the Pacific. Interest rates have increased systematically over time and are highest among countries with the highest GDP per capita. It should also be noted that the spread between interest rates charged

Table 2. Median Values of Project Planning and Economic Variables for World Bank-Financed Shelter Projects, 1972-84

Category	Minimum target group income (U.S. dollars per month)	Percentage down payment	Affordability ratio ^a	Project interest rate	Rate of inflation ^b	Per capita GDP ^c	Percentage of income held by lowest income quintile	Number of projects
<i>Region</i>								
East Africa	49	0	20.0	9.00	15.7	339	4.5	15
West Africa	131	10	20.0	9.25	9.0	765	6.1	7
East Asia and the Pacific	60	10	25.0	12.00	15.7	779	5.7	13
South Asia	27	10	20.5	12.00	12.0	249	7.0	7
Europe, Middle East, and North Africa	141	10	20.0	7.00	9.6	944	5.1	8
Latin America and the Caribbean	78	5	20.0	10.00	15.1	1,763	2.9	18
<i>Period^b</i>								
1972-76	74	5	22.5	8.00	16.0	730	3.4	13
1977-80	62	10	20.0	10.00	13.5	835	4.8	29
1981-84	60	10	21.3	11.00	11.1	741	5.3	26
<i>Per capita GDP^c</i>								
Below 600	32	5	20.0	9.00	14.2	296	5.2	25
601-1200	75	10	20.0	10.00	9.4	835	5.2	23
Over 1200	109	10	25.0	11.00	18.0	1,833	3.2	20

a. The fraction of income households are willing and able to pay for shelter.

b. In year of appraisal report.

c. Expressed in 1983 U.S. dollars per year.

in projects and inflation rates has narrowed appreciably over time, which indicates a rise in real interest rates on project loans.

It is useful now to examine in more detail how the planning assumptions inherent in sites and services projects, particularly affordability ratios, correspond to empirical evidence on actual willingness to pay for shelter and services. Figure 4 superimposes two pieces of information concerning World Bank-financed sites and services projects. First, the scattering of triangles indicates for each project its design affordability ratio in relation to estimated average household monthly income expressed in 1981 dollars. Second, the curve indicates how rent-income ratios are estimated to vary with income for typical target groups (assumed to be in the thirty-fifth percentile of the income distribution) based on the cross-country and within-city econometric results presented in section III. The curve is derived from the transpose of the cross-country housing expenditure equation for renters (equation 2), assuming a 0.5 within-city income elasticity of housing demand, and that income at the thirty-fifth-percentile was about 43 percent of average city income (see footnote 6). The figure may be interpreted as follows: if the target group is assumed to have been households at the thirty-fifth income percentile, then the curve shown on the figure indicates our estimate of their actual willingness to pay for housing as a fraction of income.

The figure clearly suggests that affordability, and by implication project design standards, are systematically overestimated in low-income countries. On the assumption that thirty-fifth percentile households represent the lowest-income target group in sites and services projects, then the majority of projects in countries with average household monthly incomes below \$400 appear to have been designed with affordability ratios that were higher than typical ratios of housing expenditures to income—many by substantial margins.⁹

Some of the implications of setting project standards too high have been discussed in section III; among these are exclusion of intended target income groups and participation by higher-income groups, relatively slow consolidation or upgrading of housing and plots beyond initial standards, creation of incentives for subletting or resale by initial allottees, and making subsidies necessary in order to reach intended beneficiaries. The critiques of sites and services projects summarized in section I presented anecdotal evidence that such outcomes have in fact been observed in many sites and services projects. However, other evidence suggests that many of these problems may have been isolated and that project implementation and downstream performance have, by and large, been successful (see, for example, Keare and Parris 1982).

This apparent paradox in the range of outcomes can be partly explained by the existence of subsidies. As section II discussed, subsidies can mitigate a number of possible project problems, in effect purchasing project feasibility at the

9. Were households at the twentieth percentile assumed to have been the lowest-income target group, affordability and design standards would appear to be overstated in countries with monthly incomes below about \$150.

Figure 4. World Bank Sites and Services Projects: Assumed Willingness to Pay versus Empirically Estimated Willingness to Pay for Thirty-Fifth Income Percentile

Key: Δ project assumptions on willingness to pay
 — empirically derived willingness to pay

Note: Based on project appraisal reports. The affordability ratio is that given for the minimum income target group. Income is estimated by adjusting GDP per capita figures to get household disposable income expressed in 1981 U.S. dollars.

Table 3. *Cost and Subsidy Elements in Sites and Services Projects*

<i>Cost element</i>	<i>Subsidy element</i>
Site preparation and housing construction	Cross-subsidies from higher- to lower-income households within project, or government-provided cost writedowns
+ Land acquisition	Government or expropriated private land at below cost
+ Off-site infrastructure and project administration	Ignored or charged at less than marginal cost
+ Recurring costs (utilities, maintenance)	Ignored or charged at less than marginal cost
= Total development and operation costs $\times$ (1 + economic rate of return)	Prices set to recover schedule costs, not resource costs
= Total resource costs $\times$ market rate of interest	Subsidized interest rates
= Annualized resource costs	

expense of longer-term budgetary obligations. To examine the magnitude of subsidies that have actually been provided in World Bank-financed sites and services projects requires first having an understanding of the components of project cost and the sources of potential subsidy elements; these are illustrated in table 3. The table indicates that total development and operating costs are made up of a number of discrete elements, most notably land acquisition costs, site preparation and housing construction, off-site infrastructure and project administration, and recurring costs such as utilities and maintenance. For a proper accounting of resource expenditures, these must be valued at their full opportunity cost. For example, even though government land may be provided at small or no cost, its market value is the proper resource cost measure. The total resource cost of a project, however, should account for the market value of the finished product. Thus the market value of the project, which is equal to the capitalized value of resource costs, is equal to the total development and operation costs multiplied by a factor equal to one plus the project's economic rate of return. The annualized resource cost is then equal to the total resource cost multiplied by the market rate of interest.

Associated with each cost element is a possible subsidy element; these are indicated in table 3. These sorts of subsidies are prevalent in many donor-financed sites and services projects and in government housing schemes that predated the sites and services paradigm.

Consider, for example, only one subsidy element, but an extremely widespread and significant one: interest rate subsidies. As table 2 indicates, median interest rates charged on sites and services project loans ranged from about 7 to 12 percent across regions. At the same time, inflation rates among those regions ranged from 9 to about 16 percent at the time of project appraisal. Assuming that the market rate of interest is, conservatively, from 2 to 3 percentage points above inflation, median interest rate subsidies appear to have ranged from

roughly 20 to 55 percent of annual market interest charges in the six World Bank administrative regions. Were one to calculate real interest rates in World Bank sites and services projects as nominal interest rates less the rate of inflation in the year of project appraisal, 61 percent of Bank projects are estimated to have had negative real rates of interest at the time of project appraisal. Such subsidies are often part of a general pattern of subsidized interest rates that pervade public sector interventions in developing countries, and are thus difficult to eliminate or control at the level of planning an individual sites and services project.

Other subsidy elements, which are more amenable to project-level control, can also be significant, however. Pricing policies for land, off-site infrastructure, administrative costs, utilities, and maintenance are all generally more subject to project-level negotiation than are interest rate policies. Also, decisions about whether or not to price sites and services at a level that recovers full resource costs or simply scheduled costs can be made within the project planning context.

In order to estimate the rough order of magnitude of subsidies in Bank projects, a simplified version of the cost-accounting framework presented in table 3 was applied to seven of the earliest Bank projects, for which it was possible to examine actual pricing policies rather than the hypothetical projections contained in initial project appraisals. Subsidies were calculated as the difference between a measure of the annualized resource cost and actual charges to project beneficiaries. Annualized resource costs were estimated based on reported project development costs not including community facilities and not including recurring costs. These generally include, but are not limited to, stated development costs, land acquisition costs, pro rata shares of off-site infrastructure, and project administration. These development costs are then increased by a percentage equal to the project's projected economic rate of return to get an estimate of the market value of the project. This market value is then multiplied by a 12 percent per year rate of interest to get an annualized resource cost; as is discussed below, this represents in general a conservative estimate of the opportunity cost of money in each of the seven countries.

The resulting calculation measures the economic subsidy as distinct from the financial subsidy in a project. In financial terms, subsidies would be zero if actual project costs were covered by project charges, regardless of whether land prices were written down, interest rates were below the market rate, or units were priced below market selling prices. But in economic terms, subsidies could be considerable even if financial costs were fully covered. By pricing project units so as to recover costs as incurred rather than to recover resources or opportunity costs, project authorities forgo potential revenue and hence decrease opportunities for accumulating surpluses that could be used to replicate projects more broadly.

It could be argued that, from the standpoint of implementing agencies, the best pricing strategy to follow would be one that resulted in no financial subsidy (for example, full recovery of actual costs) but a modest economic subsidy. This latter subsidy would be justified in order to attract participants and to send a

Table 4. *Subsidy Estimates and Interest Rates in Sites and Services Projects*

Country	Subsidy as percentage of annualized resource cost		Interest rate charged in the project (percent)	Compound inflation rate (1973-78) (percent)
	Economic ^a	Financial ^b		
Botswana	72	61	8.0	12.7
Côte d'Ivoire	68	49	13.0	10.5
India	20	-10	12.0	5.8
Indonesia	18	18 ^c	5.0	14.4
Jamaica	62	42	8.0	17.1
Senegal	55	46	7.0	10.7
Tanzania	67	47	6.0	25.0

a. Assumes market interest rate of 12 percent.

b. Subsidy estimate not including pricing of units below market value (for example, using cost-based pricing rather than market-based pricing; see text).

c. Project units were priced at market value; hence, economic and financial subsidies are equal.

message to the private housing sector that shelter solutions for low-income households are capable of both cost recovery and excess demand, leading to profits, thereby stimulating private sector participation in low-income household shelter provision.

Table 4 presents the estimated economic subsidy level as a percentage of annualized resource costs (assuming an opportunity interest rate of 12 percent). With the exception of the Indonesia project, which features market level pricing for all components (but has a subsidized interest rate), and the India (Madras) project, the subsidies range from roughly half to three-quarters of project resource costs. These subsidies have several sources: all projects, except the India project, have subsidized interest rates, and beneficiaries are in general charged less than the market price for their land, building materials, or completed dwellings. Cross-subsidies, which can lower the price charged to low-income households without affecting total project revenue, were implemented only in India and Indonesia.

It should be emphasized that while these subsidy estimates comprise a large part of project costs they probably underestimate the true subsidy. This is particularly true as concerns interest subsidies. As table 4 indicates, for example, interest rates charged to project beneficiaries are generally considerably less than inflation rates and of course less than market interest rates. The only exceptions among the seven early projects are the Côte d'Ivoire and India projects, for which the interest rates are either close to or above the market rate.

What should be made of subsidies of the magnitudes in evidence here? The analytical model of section II (elaborated on in the appendix) suggests a number of implications. It is clear that the subsidies provided in early World Bank projects were large enough to have overcome many of the potential problems that might have resulted from upward biased affordability assumptions and project standards. Comparing the actual subsidies provided with the estimated minimum subsidies necessary to induce participation by typical (say thirty-fifth-

percentile) target groups suggests that subsidies were probably instrumental in inducing many low-income households to participate that might not otherwise have done so.

Conversely, subsidies were sufficiently large for many project beneficiaries (those that were induced to distort their shelter consumption by consuming at overly ambitious project standards) that large relative welfare losses must certainly have occurred. That is, for many households whose participation in sites and services projects required them to spend more out of pocket than before entering the project (even after accounting for subsidies), the net benefit of participation must certainly have been worth considerably less than the resource cost of the subsidy. Also, for such households, incentives to sell out to higher-income households and propensities to default in the face of income shocks would have been substantial. Housing upgrading would have been limited in the early years of the project for such households unless subsidies were generous enough (as, for example, in the case of heavily subsidized building loans) to permit continued housing investment at low cost. The one possibly beneficial side effect for such households and for the housing market in general would have been the encouragement of providing additional space for low-income renters through subletting.

At the same time, many households participated in the projects who would probably have done so without subsidy, either higher-income households or households with stronger than average housing preferences. Each such group would have had incentives to invest in upgrading their units beyond initial standards. For the former group, receipt of large subsidies is inequitable; for the latter group, it is inefficient.

By and large, however, sites and services projects have appeared to have worked as intended: target groups have been reached; affordability by target groups has generally been evident (indeed, some groups have been observed to be spending even more than initial affordability targets); and reported defaults on account of affordability problems have not been overwhelming. The comparatively small scale of many projects has also worked in favor of project success, since households with relatively stronger preferences for the bundle of shelter and services offered by projects tend to enter projects via a process of self-selection. Indeed, much of the evidence concerning project impacts cited in section I appears to paint just such a picture.

Even while project-by-project assessments indicate the appearance of success, much of it purchased at the expense of significant subsidies, a broader view of the sites and services experience is consistent with Cohen's view (1983) that "replicability [can] no longer mean doing more of the same thing." The most obvious concern is that due to budgetary constraints large subsidies will undermine the ability of governments to replicate projects on a large scale and thereby cause them to fail to meet the immense needs of the poor for shelter and services. Another concern is that the subsidies that are provided are not particularly rational. The types of subsidies provided, their magnitude, their distribution

among potential beneficiaries, and their impacts on either beneficiaries or subsidy-granting organizations are rarely addressed in a thorough way when projects are designed. As a result, subsidies are likely to be inefficient, sometimes inequitable, and harmful to the long-term viability of the subsidy-granting institutions.

V. SUMMARY AND CONCLUSIONS

This article has reviewed evidence on the performance of sites and services projects. While projects have been successful in many ways as implemented, there are limits to their effectiveness as a strategy to meet the needs of the poor in developing countries. The single most serious problem in extending the application of the sites and services paradigm is the existence of subsidies that are simply too high to permit large-scale replication of projects.

Reasons for the existence of such subsidies are complex, in part due to the inertia of previous government shelter and services policies that established even larger subsidy amounts. To a considerable degree, however, subsidies are a reflection of overly ambitious planning standards which require that large subsidies be maintained to make projects affordable by intended beneficiaries. In many cases these standards are the result of the inappropriate application of a rule of thumb that households can spend from 20 to 25 percent of their incomes for shelter and related services. As a general rule, this one is wrong.

The empirical analysis of housing demand in developing countries indicates that the actual fraction of income that households are willing to pay for shelter varies with household income, a country's level of economic development, the relative price of housing, tenure status, and a number of other variables. The fraction actually spent on housing generally falls with income within cities but increases with average income across cities, relationships that parallel those observed by Kuznets regarding the way that consumption varies within and across countries.

When one compares actual spending patterns with the affordability assumptions used in World Bank sites and services projects as the basis for project design standards, it is apparent that there has been a systematic upward bias in affordability and hence in design standards. This is acute for low-income countries but largely disappears for better-off developing countries.

The magnitude of the subsidies required to reach typical target groups in World Bank projects (roughly the thirty-fifth percentile of the income distribution) is as high as 80 or 90 percent of project costs in low-income countries with planning standards that are derived from the rule of thumb that households can spend 20–25 percent of income for housing. Alternatively, if no subsidies are provided, households with incomes in the eightieth percentile or above might be the most likely project participants. It was found, however, that required subsidies and income levels of likely participants are highly sensitive to the level of

standards (and hence pricing) of project units. Thus, by choosing standards judiciously, at levels consistent with empirical evidence on housing demand, it should be possible to reach the intended beneficiary groups with small or no subsidies—a clear prerequisite for large-scale replicability in most countries.

Our review of actual planning practice in World Bank sites and services projects found that all of the projects had significant subsidy elements, with a median subsidy equal to about 62 percent of total resource cost. Sources of subsidies included (1) failure to account for all project costs at their true resource cost (for example, the use of written-down government land), (2) pricing units at or below cost rather than at market value, and (3) charging below market interest rates. The latter is widespread, with 61 percent of projects analyzed estimated to have been charging interest rates below the inflation rate during the year of project appraisal. It appears, however, that the real rate of interest being charged in World Bank projects has risen over time, perhaps in response to an explicit goal to rationalize project design policies.

Subsidies of the magnitude of those found in early Bank-financed projects (which are not atypical of similar projects sponsored by other aid organizations) appear to be sufficiently high to have induced target group households to participate, despite requiring them to spend more for shelter and services than they would normally be expected to do. Thus subsidies have mitigated potential problems from overly ambitious planning standards, namely, high levels of defaults, slow rates of housing consolidation or upgrading, turnover among project beneficiaries with higher-income households replacing initial allottees, and lack of participation by intended target groups. While some researchers have noted the existence of these problems in sites and services projects, it seems certain that they have been less severe than might have been the case, since high subsidies in effect purchase the participation and cooperation of project beneficiaries.

There are costs to such subsidies, however, the most severe of which is that they defeat the goal of cost recovery within projects and hence imperil the financial ability of the government to replicate projects on a large scale. Thus they perpetuate some of the problems inherent in the pre-sites and services public housing policies of many developing countries. In addition, the subsidies that are provided are not particularly rational; there is little explicit consideration given in much project planning to the magnitude, type, or distribution of subsidies and little consideration given to their impact on beneficiaries or subsidy-generating institutions. As a result, while subsidies do buy a measure of small-scale success in sites and services projects, they do so at the expense of replicability, equity, and efficiency.

Despite what may seem a critical view of sites and services projects, it is clear that straightforward means exist, at least in principle, to fix many of the problems noted above. For example, there are several ways to reduce subsidies: one is to reduce project standards and hence project costs; another is to change pricing

policies to better reflect true resource costs; another is to provide a bundle of sites and services that increases household willingness to pay while leaving project costs unchanged. We consider each in turn.

The analysis above indicated clearly that project standards are too high in most low-income developing countries. The solution to this problem is simple: start planning on the basis of actual housing demand patterns rather than on the basis of inappropriate rules of thumb about affordability. Actual patterns can be established on the basis of either the sorts of research findings presented in section III or local housing surveys. In the case of the former, a forecasting model has been developed by the authors that uses easily available macroeconomic data (or, if it is available, household survey data) to establish rough willingness to pay estimates in developing countries (Mayo and Gross, forthcoming). In the case of the latter, materials have been developed by Malpezzi, Bamberger, and Mayo (1982) and by Malpezzi (1984) that discuss how to design and analyze urban housing surveys. Simply getting the affordability assumption straight can go a long way toward bringing standards down to an appropriate level in many countries.

To reduce subsidies, pricing policies for sites and services projects must also be revised to reflect true resource costs. This means costing project elements (land, building materials, infrastructure, administration, and recurring costs) at their full resource costs and then charging accordingly. Loans to beneficiaries should also be made at market rates of interest in order to allow loan-granting organizations to generate a sufficient surplus to ensure their institutional viability and growth.

If there are to be departures from resource-cost-based pricing in order to subsidize some households, these should be explicitly recognized, analyzed, and discussed by project planners. Procedures should be established for identifying subsidy elements, quantifying them, and estimating their incidence and impact on project beneficiaries and their consequences for institutions responsible for them. Subsidies should, to the extent possible, be rationalized to serve equity, efficiency, and project impact criteria.

Another way to reduce subsidies is to provide households with a bundle of shelter and services that maximizes the perceived benefit for a given level of cost of provision. For example, if information is available on the trade-offs made by households among different elements of the housing and infrastructure bundle (for example, relative preferences for size and location of plot, size and quality of structure, proximity to community facilities, and quality and type of infrastructure), it may be possible to design packages of shelter and infrastructure characteristics for which households are willing to spend a good deal more than they would do for currently available commercial housing, while costing no more to provide than shelter they already occupy. In some cases, dealing with capital market and land market imperfections by making available long-term finance or secure tenure may induce significant changes in household willingness

to pay for shelter. Research on such trade-offs in developing countries has been conducted recently by Quigley (1982), Follain and Jimenez (1985), and with particular reference to the demand for secure tenure, by Jimenez (1984) and by Friedman, Jimenez, and Mayo (1985). Related work by Gross (1984) looks not only at the trade-offs made by households among shelter and infrastructure attributes but also at the influence of providing different bundles of characteristics on groups most likely to participate in a project. More empirical research into the nature of such trade-offs and their implications for project design could be of great benefit.

The analysis presented here has implications at the level of housing sector policy as well as at the level of project planning and design. Government shelter and services policies for other than sites and services projects can benefit from more vigorous attempts to reduce standards to truly affordable levels, to modify pricing policies to improve cost recovery, and to attempt to package shelter and services in ways that maximize their attractiveness to potential beneficiaries. It is particularly important to rationalize subsidy policies at the sectoral level, because there are often severe limits to the extent of reform that can be accomplished at the project level. This is particularly true in the case of interest rate policies, for which the legacy of subsidized rates for higher-income households makes it politically difficult to charge unsubsidized rates in sites and services projects. Similarly, pricing policies for urban services must often be addressed at a sectoral level before project-level reform can occur successfully.

Two directions in shelter policy that follow from the observation that standards have been upward-biased in low-income countries are (1) to do relatively more upgrading projects than sites and services, and (2) to focus policies relatively more on improving the supply of rented housing rather than focusing predominantly on owner occupied housing. While upgrading or sites and services are and will remain useful complements in shelter policy, it seems clear that the task of reducing standards to affordable levels is often more easily accomplished within the context of slum upgrading projects than in sites and services projects, an approach that would appear warranted on grounds of both equity and efficiency. Evidence that just such a shift has already begun is presented by Ayres (1983, p. 158), who indicates that the World Bank has increased emphasis of slum-upgrading "partly because some of the earlier sites and services projects proved too costly for the urban poor, partly because the number of beneficiaries in sites and services projects tended to be small, [and] partly because it fit better with the Bank's emphasis on realism and lower standards."

Reduction in standards can also be attained by emphasizing production of private rental housing rather than focusing largely on production of housing for owner occupancy. A reasonable strategy would be to encourage existing property owners to intensify development of their properties, to create additional dwelling units for rental occupancy by either horizontal or vertical expansion. Incentives for such development can be created by many of the same instruments

used in upgrading schemes, particularly infrastructure investment and tenure regularization. Such strategies have the potential in many cities to accommodate larger portions of the low-income population in adequately serviced housing than do those that emphasize higher standards, lower densities, and owner occupancy.

In conclusion, it must be noted that the innovative approach of sites and services as a way of improving the lives of poor households in developing countries remains a valid one. The approach has in fact delivered a great deal of shelter and related services to households that might not otherwise have attained them, has by and large reached lower-income households than were typically served by other government-sponsored housing projects, and, despite still-significant subsidy levels, has probably reduced average subsidy levels per household below levels of previous programs. As importantly, the process of undertaking sites and services projects has served as a catalyst and focal point for discussions of housing and land policies, institutional roles and capacities, and training needs, the outcome of which have stimulated reform that goes well beyond the boundaries of the projects themselves.

These real accomplishments, however, cannot yet be seen as having achieved the goal of large-scale replicability that is so much the object of the sites and services paradigm. For upgrading and for sites and services to move beyond the level of demonstration projects, reforms must be undertaken at the level of both the project design process and sectoral policy. In most cases, simultaneous reform is necessary in order for standards to be set appropriately, for pricing and subsidy policies to be rationalized, and for resource mobilization and cost recovery goals to be met. This will not be easy. But as experience has shown, a great deal can be accomplished by applying technical skills and political will in support of a well-founded sites and services strategy to serve the shelter needs of the poor in developing countries.

APPENDIX

ESTIMATION OF MINIMUM SUBSIDIES

We use a Stone-Geary utility function (discussed in depth in Philips 1974) to approximate household preferences:

$$(4) \quad U = (H - \Theta_h)^{b_h} (X - \Theta_x)^{b_x}$$

where H = housing consumption; X = other goods consumption; and Θ_h , Θ_x , b_h , and b_x ($= 1 - b_h$), are parameters. Applying the procedures described in Mayo and Gross (1985), the following expressions can be derived for the minimum housing necessary to induce participation, $H_{\min}$, and the minimum subsidy (benefit) necessary to induce participation, $S_{\min}$:

$$(5) \quad H_{\min} = \left(\frac{y - R_0 - \Theta_x}{y - R_s - \Theta_x} \right)^{b_x/b_h} (H_0 - \Theta_h) + \Theta_h$$

and

$$(6) \quad S_{\min} = p_h H_{\min} - R_s,$$

where R_0 = initial preproject rent, R_s = within-project rent, H_s = within-project housing, and $p_h H_{\min}$ = market value of shelter and services provided in the project.

If the project is designed according to the typical sites and services paradigm, within-project rent reflects the design affordability ratio, a , and hence $R_s = ay$. Thus, equations 5 and 6 become:

$$(7) \quad H_{\min} = \left(\frac{y - R_0 - \Theta_x}{y(1-a) - \Theta_x} \right)^{b_x/b_h} (H_0 - \Theta_h) + \Theta_h$$

and

$$(8) \quad S_{\min} = p_h H_{\min} - ay,$$

where R_0 is determined by the housing expenditure function.

The parameters of a Stone-Geary utility function may be locally approximated using knowledge of the parameters of a log-linear housing expenditure function; the resulting parameters can then be used to evaluate equations 7 and 8 for various values of a . This approximation requires knowing the price elasticity of demand, ϵ_p , the income elasticity of demand, ϵ_y , and the rent-income ratio at the point where the function is to be approximated. These parameters are successively substituted into the expenditure equation consistent with the Stone-Geary utility function

$$R = \Theta_h (1 - b_h) P_h + b_h y - b_h \Theta_x$$

and the expressions for ϵ_y and ϵ_p , ($\epsilon_y = b_h y/R$ and $\epsilon_p = -\epsilon_y \frac{y - \Theta_x}{y}$), to solve for Θ_h , Θ_x , b_h , and b_x .

ESTIMATION OF PARTICIPANT INCOME LEVELS

The participant income level and rent for a project can also be estimated. By manipulating across-country and within-city expenditure functions, we can derive the income level, $\hat{y}$, of households that would freely enter a project at its full unsubsidized price, where the latter is set equal to the assumed affordability ratio, a , multiplied by the income of the typical target group, defined here as income at the thirty-fifth income percentile, y_{35} :

$$(9) \quad \hat{y} = e^{\frac{1}{\epsilon_y} \{ \ln a y_{35} - [\ln \left(\frac{\hat{R}}{y} \right) - (\epsilon_y - 1) \ln \hat{y}] \}}$$

where ϵ_y is the within-city income elasticity of demand, $\bar{y}$ is average city income, and $(\bar{R}/\bar{y})$ is the rent-income ratio evaluated at average city income (which is estimated using the cross-country expenditure relationship, equation 2). The percentile of the income distribution associated with $\hat{y}$ is found by integrating the estimated income density function to an upper limit of $\hat{y}$.

REFERENCES

- Abt Associates. 1982. *Informal Housing in Egypt*. Cambridge, Mass.
- Ayres, Robert L. 1983. *Banking on the Poor*. Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Chana, T. S. 1984. "Nairobi: Dandora and Other Projects." In G. Payne, ed., *Low-Income Housing in the Developing World*. New York: Wiley.
- Cohen, Michael. 1983. "The Real Challenge of Replicability." *Regional Development Dialogue* 4 (Spring): 90-99.
- Davidson, F. 1984. "Ismailia: Combined Upgrading and Sites and Service Projects in Egypt." In G. Payne, ed., *Low Income Housing in the Developing World*. New York: Wiley.
- Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.
- Follain, James, and Emmanuel Jimenez. 1985a. "Estimating the Demand for Housing Characteristics: A Survey and Critique." *Regional Science and Urban Economics* 15:77-107.
- . 1985b. "The Demand for Housing Characteristics in Developing Countries." *Urban Studies* 22: 421-32.
- Friedman, Joseph, Emmanuel Jimenez, and Stephen Mayo. 1985. "The Demand for Secure Tenure in Developing Countries." Washington, D.C.: World Bank Water Supply and Urban Development Department. Processed.
- Grimes, Orville F., Jr. 1976. *Housing for Low-Income Urban Families: Economics and Policy in the Developing World*. Baltimore, MD.: Johns Hopkins University Press.
- Gross, David J. 1984. "Designing a Suitable Housing Project: Integration of a Demand Module into a Supply Side Planning Model." Washington, D.C.: World Bank Water Supply and Urban Development Department.
- Jimenez, Emmanuel. 1984. "Tenure Security and Urban Squatting." *Review of Economics and Statistics* 66 (November): 556-67.
- Kakwani, Nanak C. 1980. *Income Inequality and Poverty*. New York: Oxford University Press.
- Keare, Douglas, and Scott Parris. 1982. *Evaluation of Shelter Programs for the Urban Poor: Principal Findings*. World Bank Staff Working Paper 547. Washington, D.C.
- Kravis, Irving B., Alan Heston, and Robert Summers. 1982. *World Product and Income: International Comparisons of Real GDP*. Baltimore, Md.: Johns Hopkins University Press.
- Kuznets, Simon. 1961. "Quantitative Aspects of the Economic Growth of Nations. VI. Long-Term Trends in Capital Formation Proportions." *Economic Development and Cultural Change* 9 (July).

- Malpezzi, Stephen. 1984. "Analyzing an Urban Housing Survey: Economic Models and Statistical Techniques." World Bank Water Supply and Urban Development Department Discussion Paper 52. Washington, D.C.
- Malpezzi, Stephen, Michael Bamberger, and Stephen K. Mayo. 1982. "Planning an Urban Housing Survey: Key Issues for Researchers and Program Managers in Developing Countries." World Bank Water Supply and Urban Development Department Discussion Paper 44. Washington, D.C.
- Malpezzi, Stephen, Steven K. Mayo, and David J. Gross. 1985. *Housing Demand in Developing Countries*. World Bank Staff Working Paper 733. Washington, D.C.
- Mangin, William. 1970. *Peasants in Cities*. Boston: Houghton-Mifflin.
- Mayo, Stephen K. 1981. "Theory and Estimation in the Economics of Housing Demand." *Journal of Urban Economics* 10 (July): 95-116.
- Mayo, Stephen K., and David J. Gross. 1985. "Sites and Services—and Subsidies: The Economics of Low-Cost Housing in Developing Countries." World Bank Water Supply and Urban Development Department Discussion Paper 83. Washington, D.C.
- . Forthcoming. "Housing Affordability: Models and Methods for Project Planners." World Bank Water Supply and Urban Development Department Discussion Paper. Washington, D.C.
- Payer, Cheryl. 1982. *The World Bank: A Critical Analysis*. New York: Monthly Review Press.
- Phlips, Louis. 1974. *Applied Consumption Analysis*. Amsterdam: North-Holland.
- Quigley, John. 1982. "Nonlinear Budget Constraints and Consumer Demand: An Application to Public Programs for Residential Housing." *Journal of Urban Economics* 12: 177-201.
- Turner, John. 1972. *Freedom to Build*. New York: Macmillan.
- Tym, Roger. 1984. "Finance and Affordability." In G. Payne, ed. *Low-Income Housing in the Developing World*. New York: Wiley.
- World Bank. 1974. *Sites and Services Projects*. Washington, D.C.
- . 1975. *Housing*. Sector Policy Paper. Washington, D.C.
- . 1983. *Learning by Doing: World Bank Lending for Urban Development, 1972-82*. Washington, D.C.

Multimarket Analysis of Agricultural Price Policies in an Operational Context: The Case of Cyprus

Avishay Braverman, Jeffrey S. Hammer, and Anne Gron

This article presents a method of analysis of agricultural price policies. Its basic features are (1) the incorporation of important supply and demand substitution possibilities; (2) a flexible structure which can account for institutional features of the agricultural sector and its relationships with the rest of the economy; and (3) a focused, policy-oriented approach. An application to Cyprus is presented which highlights the differences between this approach and analyses of single markets. The method incorporates supporting software on personal computers, with the longer-term objective of developing this type of approach as a standard tool in pricing policy analysis.

The idea that the economic world was a general system, with all parts interdependent, seemed (and seems) to me to be an essential of good analysis.

—KENNETH J. ARROW (1964)

The purpose of this article is to report on an operational methodology for analyzing the economic impact of agricultural price (tax/subsidy) policies. This technique accounts for the interaction of markets as relative prices change, a crucial element in agricultural production and pricing. The approach is intended to support the operational economist involved in formulating policy advice within the short time allowed for operational work. We explain below the rationale for using this methodology, briefly compare it with other techniques for analyzing agricultural price policies, and demonstrate its application in reducing animal feed subsidies to livestock producers in Cyprus.¹

1. This method, and similar software for personal computers, has also been applied in conjunction with policy discussions for Hungary and Brazil; see Braverman, Hammer, and Morduch (1985) and Braverman, Hammer, and Brandao (forthcoming). Information about the software is available from the authors on written request. See Braverman, Hammer, and Ahn (forthcoming) and Braverman and Hammer (1986) for methodological discussion and references for other agricultural pricing studies in Cyprus, the Republic of Korea, Senegal, and Sierra Leone. See Kirchner, Singh, and Squire (1984) for an application to Malawi.

Avishay Braverman and Jeffrey S. Hammer are at the World Bank. Anne Gron is a graduate student at the Massachusetts Institute of Technology. The authors thank Angus Deaton, Jonathan Morduch, and many Cypriot officials and researchers for helpful discussions.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

This model is designed to answer questions frequently of concern to policymakers. While the critical questions vary by country, public debate on the impact of changes in price policy usually involves change in the government deficit, foreign exchange earnings (or requirements), and real incomes of various groups within society (for example, urban versus rural residents). Any economist expected to give responsible advice on agricultural pricing policy issues must also be concerned with these issues.

I. ALTERNATIVE METHODS

Standard tools of analysis which are currently in use have serious drawbacks for an operational economist who wants to include these considerations. Studies of comparative advantage—using effective protection rates (EPRS) or domestic resource cost (DRC) calculations—cannot be used either to answer questions about the distributional impact of policies or to determine the adjustment of quantities produced or consumed. Single market studies of producer and consumer surplus can address questions about the short-term impact of price changes on real income (though only on the functional distribution of income—that is, between producers and consumers of products). They are limited, however, in their ability to take into account important interactions between markets. Conversely, linear programming and large computable general equilibrium models may be appropriate as research tools, and for circumstances in which time is not limited and the data base is rich and reliable. These methods are seldom useful, however, for operational economists working on immediate policy prescriptions in developing countries which require incorporation of substantial institutional detail. The multimarket approach is often the most appropriate of these various tools for analysis in the very common circumstances we have described above. We elaborate further on these methods below.

Nominal and Effective Protection Rates

Calculation of these rates shows the relationship between domestic prices and world prices, with EPRS corrected for prices of intermediate goods. While these measures are appropriate for project evaluation, most of the issues mentioned above are not addressed, which limits their usefulness in a discussion of policy.

These numbers are also frequently used to measure “distortions” from a free trade position. There are two problems with this approach. First, completely free trade is often not within the realm of possible policy options, as governments will always intervene somewhere in the economy. In this second best (or *n*th best) world, it is not clear what information is contained in the EPR calculations. For example, with arbitrary distortions in other markets within the economy, it cannot be asserted that reducing the tariff on a good with a positive protection rate will improve welfare. Substitution possibilities with other goods which are protected may exacerbate the distortions in those markets, which

makes the net position worse. By the same token, priorities for liberalization cannot be made on the basis of relative EPRS. Decreasing a tariff on a highly protected good is not necessarily better than decreasing one on a good with a lower EPR, again due to possible differences in elasticities.²

The second problem with these measures is that even if free trade is a possibility, it is unlikely to be a Pareto-improving policy (that is, someone is liable to get hurt). Policymakers may like to know the probable winners and losers in the policy change. If the poor are hurt, this information can be used in designing compensation schemes. These analytic tools do not provide any guidance in these questions.

Single Market Analysis of Consumer and Producer Surplus

Analysis of supply and demand, including calculations of producer and consumer surplus which are done market by market, can address some of the issues mentioned above. In this approach, foreign exchange earnings, government expenditures or revenue, and certain aspects of distributional effects are separately accounted for in each market. However, except in very particular (and in the case of agriculture very peculiar) circumstances, the results of such analyses will be misleading. Single market analysis will give the correct effects on budgets, foreign exchange, and welfare only if the crop in question is not a substitute or complement in supply or demand for any other good which is (a) taxed or subsidized (in budget analysis); (b) traded (in evaluating foreign exchange effects); or (c) subject to any distortion at all, either fiscal or from the private economy (in the study of welfare effects). The chance of such a good existing in agriculture is small. The errors can be substantial, to the extent of predicting even the wrong *direction* of a response with respect to a policy change. This is the risk run by ignoring the substitution effects between markets. While all this is known (for example, Harberger 1971 and Tolley, Thomas, and Wong 1982), the actual use of consumer and producer surplus analysis is often conducted using a single market framework.

Linear Programming and Computable General Equilibrium Models

At the other extreme, large computable general equilibrium (CGE) models and mathematical programming models are also limited as operational tools. Both of these methods deal with the intermarket connections emphasized here. For certain purposes, such as examining the introduction of a new production technique (in the linear programming case), these techniques can be important and appropriate. However, both have serious drawbacks as tools for policy analysis. First, elaborate models are essentially research tools which take a considerable amount of time and data to construct. Usually they cannot be done within the time horizon of operational work.

2. See Bhagwati and Srinivasan (1973) on the inadequacy of effective protection measures in the presence of general equilibrium interactions.

Second, their complexity makes the incorporation of institutional detail more difficult (though not impossible). Similarly, changes in the model parameters for sensitivity analysis purposes are difficult to effect.

Third, and perhaps most important, these models are frequently of such complexity that results are not intuitive—certainly not to a policymaker, and often not to the analyst! This lack of communicability of results can be a serious liability to the analysis. These methods can be considered complementary to the multimarket approach, however, in that ideas generated in one can be used in the construction of the other.

Multimarket Method

The multimarket method can be seen as filling the gap between the single market method and the sophisticated modeling exercises. It is a simulation tool which incorporates both interactions across markets and the economic, political, or technical constraints which policymakers face. Our method is intended to be as simple as possible, while capturing the essential features of the country in question, to allow the economist to understand and explain the consequences of policy. It is not, of course, necessary to have complete information about the extent of market interdependence to make good use of the multimarket approach we describe. The analyst must be sufficiently familiar, however, with a country's agricultural economy to decide which interlinkages are critical and which indirect effects are insignificant. This approach is intended to consolidate the various bits and pieces of knowledge about the sector in a consistent framework, which allows the analyst to draw the logical implications of what is known about the sector and its interrelations with other sectors.³

It is possible to "close" the model and make it a fully general equilibrium system. This generally entails including factor markets which link the sectors. If the important policy questions are essentially intersectoral, this would be the appropriate modeling strategy.

The earlier applications of our method were subject to some of the same drawbacks as the programming and CGE approaches. The studies relied on explicit functional forms for production and demand systems. The forms chosen were so-called flexible functional forms, so that a wide variety of behavioral responses could be accommodated.⁴ It was necessary in constructing the model, however, to reconcile observable data, the assumptions of the model (in terms of

3. Insights from this approach are derived from the farm household economy model (for example, Yotopoulos and Lau 1974; Barnum and Squire 1980).

4. We have used the Almost Ideal Demand system form developed by Deaton and Muellbauer (1980) for demand and the translog form (for example, Lau 1976) for supply. It should be noted that the more restrictive and frequently used functional forms (such as Cobb-Douglas in production and the linear expenditure system in consumption) are often the sole determinants of results rather than the data used (see Deaton 1984 on this point). In particular, the linear expenditure system does not allow for negative income elasticities, that is, inferior goods. Many basic foods (for example, rice and barley) become inferior for relevant income levels.

conditions which characterize an equilibrium), and the use of the particular functional forms used. This required the use of fairly complicated calibration procedures. The method of calibration is explained in Braverman, Hammer, and Ahn (forthcoming, appendix). The essential logic of the procedure was to choose the parameters of the model that correspond as closely as possible to observable levels for the variables and estimated or assumed values for the behavioral relations, while satisfying the features of the model and requirements of economic theory (such as symmetry of the Slutsky matrix in the demand system). This was done by solving an optimization problem in which the objective was to minimize the deviations from known values subject to the necessary constraints.

There are two problems in using this approach in operational work. First, the results are dependent on the particular analyst and are not easily replicable. The choice of weights in the optimization problem, for example, is based on the educated guess of the analyst. Even if such estimates are made explicit, the effects of these judgments on the final results are not obvious. Second, changes in the assumptions of the model often require a complete recalibration. This is sufficiently complicated and time consuming to preclude an operational economist from using the procedure directly.

The purpose of this article is to present a version of the method which avoids the pitfalls of the previous approach while maintaining the fundamental rationale. The new version does not require the calibration described above. It also does not require the use of specific functional forms which impose their own restrictions on the model (unrelated to economic reality) and which involve large numbers of nonintuitive parameter values which obscure the logic of the analysis. The method proceeds by assembling what is known about supply and demand for the important commodities, the institutional structures of government policies, and the mechanisms for market clearing. This information is arranged in a set of equations which is totally differentiated so that changes in the outcomes of interest can be solved in terms of changes in the available policy options (the equations are listed in the appendix).

Since the resulting model is linear, it can easily be solved on a personal computer. User friendly software has been written which facilitates the presentation of results and sensitivity analyses—a clear advantage. The cost is that the analysis is restricted to small changes.⁵ It is hoped that, through experimentation with various policy options and assumptions about the underlying parameters of the model, the analyst can develop a feel for the workings of the agricultural sector and its relation to the urban and external (export and import) sectors. The analysis has two principal limitations. First, it does not take intertemporal decisionmaking on the part of farmers into account. Second, it takes government decisions as exogenous. Full control over the relevant policy instruments is assumed, and any reaction in other parts of the government is ignored.

5. On how to adapt this method to analysis of large changes in prices, see section VI below.

Positive versus Normative Analysis

The multimarket method is essentially a positive approach. In contrast to both Harberger (1971) and the optimal tax tradition (for example, Atkinson and Stiglitz 1980), it emphasizes the positive analysis of consequences over the calculation of various social welfare aggregates. The position is often taken that, by counting dollars accrued to different groups identically, one avoids value judgments. However, the fact remains that this aggregation is a specific welfare evaluation giving identical weights to each group in the population; that is, the impact on the poor is not distinguished from the impact on the rich. The technical economist as policy analyst, in the midst of the political debate over economic policies, is often required to assess the impact of policy on different groups (for example, poor versus rich, rural versus urban). While *ex ante* aggregations often disguise social values for positive analysis, different aggregations can be made explicit *ex post*, including the impact on total real income emphasized by Harberger (see Braverman, Hammer, and Ahn, forthcoming, for Korea).

Besides evaluation of the effect of policy on private incomes, government deficit reduction, particularly in the presence of high inflation, has often been a high policy priority. To attribute a shadow price or *ex ante* "welfare weight" to reduction in the government deficit is difficult (see Stiglitz 1982 for such an attempt). Government revenue, foreign exchange, and other objectives which indirectly affect welfare should be looked at specifically.

II. THE POLICY PROBLEM IN CYPRUS

In order to demonstrate an application of the analysis, the method described above is presented in the context of the problem of feedgrain subsidies in Cyprus.⁶ The Cypriot case is instructive for a variety of reasons. The policy problem is well defined and immediately suggests the market interactions most likely to be important. The government sells barley as an input to the livestock sector at £C33.9 per ton—a price which has not changed since 1978. Most of this grain is imported. In 1982 the price of grain on world markets was £C69 per ton, leading to a deficit of £C12.4 million. By 1984, this deficit had grown to approximately £C25 million, or 2 percent of gross domestic product (GDP).

The rationale for this policy is that the feedgrain subsidy is supposed to be an indirect consumer subsidy as the cost savings in the meat and milk industries are passed on in the final product. It is also argued that, while the foreign exchange

6. In Cyprus, issues of income distribution and the rural-urban distinctions were not considered to be major policy problems by the government or the World Bank. On these issues, see Braverman and Hammer (1986); Braverman, Hammer, and Ahn (forthcoming); and Braverman, Hammer, and Brandao (forthcoming). This application demonstrates the inclusion of a livestock sector in the multimarket analysis.

requirements of the grain trade appear substantial, they are smaller than the requirements implied by importing meat products rather than the inputs.

In addition to the main intervention through the feedgrain sale price, the government also subsidizes barley cultivation through higher prices for the barley it purchases domestically. These prices have been maintained at levels higher than the world prices, though the degree of subsidy has declined over time. Price supports are also provided to wheat producers. The producer subsidies are intended as income support for older, established farmers, and revision is not generally considered to be politically feasible. Finally, as a direct consumer subsidy, bread (of the traditional type) is sold at low prices. We thus consider four components of the government deficit: the feedgrain subsidy (by far the most expensive component), the two producer subsidies, and the direct consumer subsidy on bread.

The statement of the problem alone makes apparent the interaction of related markets. Increasing the cost of inputs to livestock will increase consumer prices. The degree of this rise depends on the elasticities of demand for the various goods whose prices would be affected, as well as their cross price effects. It also depends on how factors substitute in production and more generally on supply elasticities. For the assessment of the government's cost, the increase in the feedgrain price will induce substitution to hay consumption and (as there is no international trade in hay) to hay production. This implies less cultivation of wheat and barley, and hence lower production subsidy payments. By ignoring these intermarket effects, the analyst runs the risk of making serious misjudgments concerning the budgetary and welfare implications of policy change.

III. STRUCTURE OF THE MODEL

In this section we present the structure of the relevant markets and the interactions between them. The markets included in the model were those which would be most affected by increased grain costs. These are primarily the grains themselves, their substitutes, and the livestock products which use them as feed.

It is possible that there are other effects not reflected in the model. There may be substitution between food and manufactured consumer goods which are taxed so that the model would not give a complete description of the fiscal effects of the policies. It is always a matter of judgment—the “art” of modeling—as to where to cut off the analysis to maintain optimal operational simplicity while accurately reflecting the major significant relationships. Our judgment, and that of the Cypriots consulted in the model's formulation, is that the version presented here captures the most important effects.

When price policies change, the principal distinction between the ways the markets clear is whether prices are flexible or fixed, that is, whether prices or quantities adjust to clear the market. The adjustment of fixed price markets will influence the foreign exchange requirements resulting from the policies since the quantity adjustment is through imports and exports. The adjustment of flexible

price markets will determine the cost of living effects of proposed policies. The relative importance of these two types of markets will determine whether the costs of reducing subsidies are borne by producers and the trade balance on the one hand, or consumers on the other.

In our earlier work (Braveman, Hammer, and Jorgensen 1985) explicit functional forms were chosen for the underlying supply and demand functions. In the present application, the starting point is the market clearing conditions, which require fewer restrictions on the behavior of economic agents than the mathematical requirements of the functional forms. While some of the parameter values for the present version were chosen to be consistent with our previous nonlinear study, they are essentially independent exercises.

On the production side, it is assumed that producers are profit maximizers and that a well-behaved profit function can characterize their responses. On the demand side, the comparable concept is the indirect utility function for consumers. From producer theory (see, for example, McFadden 1978), the supply of the commodity is the derivative of the profit function with respect to output price. The demand for factors of production is the derivative with respect to input price. For the case in which there are fixed factors of production, the profit function includes the levels of these factors. With this framework in mind, the arguments in the supply function should be the same as those in the profit function. Also, certain characteristics of well-behaved profit functions, such as symmetry of cross price effects, are incorporated into the supply functions.

Similarly, on the consumer side, demand is a function only of prices, with appropriate restrictions imposed by theory. The income argument in demand is suppressed. The reason for this is that only a single demand function is defined for each commodity market. The only component of income nationwide which could vary endogenously in the model is agricultural profits. This is a very small fraction of aggregate income in Cyprus and can thus be ignored without detriment to the basic analysis.

Besides incorporating basic theory into the analysis, the restrictions imposed serve the additional purpose of limiting the number of free parameters which must be determined by empirical work or expert opinion.

Commodity Market Characteristics

Beef and frozen lamb are traded internationally at fixed world prices, and therefore markets are assumed to clear via changes in imports. In the frozen lamb case there is no domestic production, so imports are equal to demand. Fresh lamb (not a close substitute for frozen) and milk are not traded internationally, so markets clear through domestic prices moving to equate supply and demand.

Beyond this, each of the markets has characteristics specific to it which can be mentioned briefly. Products of ruminant animals (cows and sheep) are characterized by joint production of milk and meat. An increase in milk prices, say, could increase beef production as a by-product of increased herd size. This is modeled

by including a term for size of stock in the profit and supply equation (making the former a “restricted” profit function). Changes in the stock itself are endogenous. We assume that the livestock industry is sufficiently well integrated into the rest of the economy that investments in animals are treated the same as any other investment. We also assume that the livestock sector is small relative to the rest of the economy, so that the rate of return on investments is determined outside this sector. The equilibrium condition for stocks, then, is that the rate of profit on an additional cow or sheep is equal to some constant value. The constant would be a risk-adjusted rate of return to capital. The rate of profit on a marginal animal is composed of the marginal profit in the meat industry plus the marginal profit in the milk industry.⁷

The markets for cow and sheep products are also linked on the output side, as milk from the two sources can be approximately perfectly substituted. The milk supply is thus the sum of production from the two sources.

Two versions of the pork and poultry market were modeled—the first assuming fixed prices with imports clearing the market and the second without trade with prices clearing the market. If costs of production are allowed to rise with the removal of the feed subsidy, the question arises as to whether the government will exclude imports and allow the price of meat to rise or will allow imports, which will limit product price increases with more severe consequences for the pork and poultry industry. Both possibilities are modeled (the without-trade case is shown in table 1 below).

Barley, wheat, and hay can all be grown on the same land, and we assume that farmers maximize profits by allocating land between them. Therefore a single, multiproduct profit function is assumed to underly the supply relations. Wheat and barley producer prices in the supply function are fixed by the government. Similarly, on the demand side, barley feedgrain prices and bread prices are also policy instruments. Demand for barley is a derived demand from the production of animal products. Both barley and wheat markets clear via imports.

Hay demand is also derived from ruminant animal production. On the supply side there are two components: hay, which is planted, and straw, which is a by-product of both wheat and barley. There is no international trade in hay, so its price varies to equate domestic supply and demand.

Policy Factors

The variables can be divided into those which are instruments of policy (exogenous) and those which are determined by the working of the system (endogenous). These elements can be solved in terms of quantities that an economist familiar with the agricultural sector is likely to know. These include the supply elasticities and actual levels of output and prices (on the supply side) and demand elasticities, consumption, and consumer prices (on the demand side). The equations and the form for solving the system are delineated in the appendix.

7. For an alternative formulation of joint products, see Deaton (1984).

Once the prices and quantities of the basic commodities are found from the model, these can be used to determine the effects on the government accounts, the foreign exchange needs, and the welfare of different groups.

The government runs a deficit in its agricultural accounts. This can be divided into four components:

1. Barley consumer deficit: this is the largest component of the deficit and consists of the difference between the sale price of barley to livestock producers and the world price (opportunity cost of acquisition), multiplied by the total feed demand for barley.
2. Barley producer deficit: this is the difference between the producer price and the world price of barley multiplied by domestic production.
3. Bread consumer deficit: this consists of the difference between world and domestic prices of bread, multiplied by total consumer demand.
4. The wheat producer deficit is calculated in the same way as the barley deficit.

Total foreign exchange requirements to support agricultural production and consumption consist of the value, at world prices, of total imports of the above commodities.

Calculation of the net effect of the policy interventions must take into account the fact that government wages are fully indexed to the cost of living. Therefore, if consumer prices rise as a result of subsidy reductions, some of the budgetary savings will be eroded by automatic increases in salaries.

For small changes in prices, the change in welfare of any group in society can be calculated in terms of compensating variations of income. This can be expressed as a fraction of base income necessary to compensate the consumer for any price increase.

Still problematic is the choice of parameter values in both the supply and demand systems. For own price effects, these values may be known or obtainable through simple methods. More difficult to obtain are the cross price effects, which require more carefully conducted econometric studies. For the case of Cyprus, we were fortunate in having such a study by the Agricultural Research Institute (Panayiotou 1983) available to us. This was supplemented by our own econometric analyses presented in the earlier study. In other applications this information may not be available. Sensitivity analysis can be used as a means of overcoming data limitation in such cases and is briefly discussed in section V below.

IV. RESULTS AND COMPARISONS WITH SINGLE MARKET ANALYSIS

The principal policy option to be briefly examined here is the reduction of the subsidy to livestock producers by increasing the sales price of barley.⁸ Table 1

8. A more comprehensive presentation of the detailed results of the Cypriot case is presented in Braverman, Hammer, and Jorgensen (1985).

Table 1. *Effects of a 1 Percent Increase in the Barley Feedgrain Price*
(percentage change)

Parameter assumption	Demand						
	Barley	Pork and poultry ^a	Fresh lamb	Milk	Beef	Bread	Frozen lamb
Base case	-0.46	-0.28	-0.33	-0.27	0.26	0.01	1.14
Alternatives ^b							
I	-0.45	-0.28	-0.34	-0.27	0.26	0.01	1.17
II	-0.48	-0.27	-0.37	-0.30	0	0	0
III	-0.72	-0.27	-0.21	-0.28	0.19	0.01	0.86
IV	-0.37	-0.28	-0.41	-0.31	0.31	0.01	1.41

Parameter assumption	Domestic Supply				Prices			
	Barley	Wheat	Hay	Beef	Fresh lamb	Milk	Pork and poultry ^a	Hay
Base case	-0.03	-0.34	0.98	-1.72	0.13	0.25	0.11	0.08
Alternatives ^b								
I	-0.84	-0.84	0.97	-1.74	0.14	0.26	0.11	0.10
II	-0.03	-0.32	0.93	-1.73	0.13	0.25	0.11	0.07
III	-0.01	-0.11	0.31	-1.03	0.09	0.25	0.11	0.02
IV	-0.03	-0.32	0.94	-1.77	0.17	0.29	0.11	0.30

Parameter assumption	Trade accounts: foreign exchange requirements	Grain commission deficits	Government wage bill	Total government deficit ^c
Base case	0.34	-1.13	0.012	-1.02
Alternatives ^b				
I	0.64	-1.41	0.012	-1.29
II	0.12	-1.14	0.012	-1.03
III	-0.03	-1.25	0.018	-1.14
IV	0.46	-1.09	0.013	-0.96

a. Based on the nontraded scenario.

b. I—maximum crop substitution in supply; II—no substitution in demand; III—livestock supply elasticities are unity; IV—field crop elasticities equal one-fourth the base elasticities.

c. Because percentage changes in the Grain Commission deficits and the government wage bill are calculated from their respective base values, the sum of these two changes will not equal the change in the total government deficit.

presents the results for the base case discussed here (and for four alternative formulations described in section V below). The policy examined is an increase of 1 percent in the price of barley. Entries are percentage changes in the indicated variable and, since the policy is a 1 percent change, these figures can be considered as elasticities. Since the model solutions simulate market equilibrium conditions, changes in supply equal those for demand for the nontraded commodities (hay, milk, fresh lamb, pork, and poultry). These results are therefore shown either under supply or demand in table 1, not under both.

The basic results can be seen in the rows labeled "Base case." The analysis indicates a distinction between traded and nontraded livestock products for the pattern of supply and demand. All animal products suffer a reduction in output. Since nontraded commodities will experience a price rise as the cost of production increases, the reduction in supply is moderated relative to that of the traded commodities. The cost of living rises due to the increase in the price of nontraded goods; foreign exchange requirements rise due to the decline in domestic production of traded livestock products.

The hay market provides the link between the livestock and field crop sectors. As hay replaces barley in the mix of feedstuff, its price rises, causing the production of the competing grains to be reduced. The higher price of hay also moderates the reduction in demand for barley.

The principal effect on the deficit is the savings in the barley subsidy. Lower per unit subsidies as well as the reduction in feedgrain used lead to a 1.13 percent savings in Grain Commission costs. This main effect can be captured using a single market analysis, but some of the indirect effects would be overlooked. Using the factor demand elasticities assumed in the multimarket model, calculations were made for a single market analysis which predict a deficit reduction of -1.21 percent from a 1 percent increase in barley prices. In a multimarket framework, the two effects operating through the hay market influence the deficit calculations. The higher hay price reduces the elasticity of demand for barley, thereby limiting the Grain Commission savings. Counteracting this effect, the increased production of hay reduces the output of both wheat and barley, which are subsidized. The incorporation of the hay market effects creates a net reduction in budgetary savings from the 1.21 found in the single market analysis to 1.13 percent. Due to the complete indexation of government wages to the cost of living, however, the increased prices in fresh lamb, milk, pork, and poultry lead to automatic pay increases, which further reduce the savings in government revenue to the 1.02 percent shown in table 1. Therefore, ignoring multimarket effects would lead to an overstatement of government revenue savings of nearly 20 percent.

In the foreign exchange calculation, the differences between the single and multimarket approaches are more pronounced. In the single market analysis, attention would be limited to the barley market. The 1 percent increase in price of barley would reduce total foreign exchange requirements in agriculture by 0.43 percent. The reduction in domestic production of meat, wheat, and barley, however, increases imports of these goods and increases foreign exchange requirements by 0.34 percent. It should be noted that this analysis does not give a complete account of foreign exchange earnings for the whole economy. The reduction in farm output releases labor and investment funds to the rest of the economy. Some substantial fraction of that resource flow will be converted into export earnings, although precise calculations of this effect require more detailed information on the nonagricultural sector.

The principal arguments against reducing the livestock subsidies were (1) that

they were really an effective consumer subsidy, and (2) that Cyprus would either import grain or import meat, which was more expensive than grain. This analysis indicates that the first reason is ill-founded since the effect on the cost of living is quite small (0.012 percent). There does appear to be some justification for the second claim, although overall export earnings are sensitive to the export earning capabilities of labor in nonagricultural sectors. Studies by others show a large number of part-time farmers, and substantial integration of the rural and urban labor force (see Ansell, Bishop, and Upton 1983). Therefore, one would expect released labor to find alternative employment, possibly in the export sector. If these released workers earn the net average export earnings of workers in the nonagricultural sectors, then the reservation concerning foreign exchange becomes much weaker.⁹

V. SENSITIVITY ANALYSIS

All of the behavioral parameters (elasticities of supply and demand) are subject to error due to econometric estimation error or inadequacies of the data. Therefore, it is important to examine the sensitivity of the results to changes in assumptions concerning these values. Four variants of the model are described here which either are based on assumptions which might be used in the absence of data (versions I and II) or which probe aspects of the base model that, while based on data, may be questioned on empirical grounds (versions III and IV).

The first two scenarios were developed because cross price elasticities are notoriously hard to estimate and simple rules of thumb may have to be used at times in the choice of parameters. The first variant uses the "base case" values of the own price elasticities of supply of field crops but alters the cross price terms. Under the assumption of fixed yields per acre, it is assumed that an increase in supply and therefore acreage of one crop reduces the acreage and supply of competing crops in equal proportions. This yields the maximum possible substitutability in crops for a given supply elasticity consistent with the convexity of the profit function. Since barley, wheat, and hay are close substitutes, this special case is more relevant than the alternative of minimum (that is, zero) substitution between crops.

The second scenario concerns the demand side of the model but makes the opposite assumption concerning cross price effects. In this case, all off-diagonal elements of the matrix of price elasticities are set equal to zero, which implies *no* substitution between these goods in demand. This minimizes all intercommodity effects on the demand side and would be a natural assumption to use if all econometric information were lacking.

Due to the limited importance of fixed factors and thus the low reported profit levels in the livestock sector, the base model has relatively high supply elasticities. Discussions with experts in the Ministry of Agriculture in Cyprus indicated

9. See Braverman, Hammer, and Jorgensen (1985) for calculations to this effect.

that these parameters may be incorrect due to underreporting and data problems. The third variant of the model therefore examines a case with supply elasticities of unity for all livestock products, with consistent cross term effects.

The base model also has relatively high own and cross price supply elasticities in the field crops. As mentioned, these three crops are close substitutes. However, the fourth scenario reduces all of the elasticities to one-fourth of their base values to offer some contrast to variant I and to suggest the extent of change likely if intercrop substitution is more difficult.

Rather than discuss each version in detail, a few overall comments are possible. Many aspects of the model do not appear especially sensitive to quite large changes in the parameters. The structure of the model and the specific linkages between markets appear to determine much of the outcome of policy experimentation. In those places where the sensitivity of the model is indicated, a few interesting features of the model are highlighted.

The interaction of field crop and livestock production is the most interesting channel of multimarket effects. When the supply elasticities of field crops are low (version IV), the increased demand for hay due to reduced subsidies results in higher hay prices which are not strongly offset by subsequent increases in hay output. Prices of nontraded meat (fresh lamb) and imports of traded meats rise faster, so the cost of living and the foreign exchange requirements both rise relative to the base. The hay price increase also moderates the reduction in barley demand, which makes the deficit saving smaller as well. Conversely, when cross price elasticities are high (variant I), the deficit saving is higher due to greater reduction in the production of subsidized grains. The policy of livestock subsidy reduction looks best (that is, has fewer negative consequences on the cost of living and the government deficit) when the field crop sector is most flexible (that is, has high own and cross price elasticities).

The most sensitive aspect of the results is the effect on foreign exchange requirements. This varies from quite high extra requirements (variant I) to a slight reduction (variant III). The versions giving lower requirements than the base are alternatives II and III. In variant II there is no substitution in demand and thus no spillover demand from high-priced, nontraded meats to imported meats. In the low supply elasticity case (version III), there is a much smaller supply decline in response to the input price rise so that much less of the meat demand needs to be met through imports. The two versions yielding higher total import values are (1) the case with high cross price elasticity for field crops (version I), in which large reductions in barley and wheat output are good for the deficit but require further imports, and (2) the case of low overall supply elasticity of field crops, which generates high hay prices and more substitution of traded for nontraded meats.

The results for foreign exchange earnings vary substantially among the different versions. Because the experts consulted believe that the supply elasticities derived are higher than actual values, variant III, which estimates the effects of livestock elasticities of unity, is likely to most closely approximate the actual

Cypriot case. As shown in table 1, the reduction in the barley subsidy in that scenario slightly *reduces* foreign exchange requirements. Added to these savings are the potential foreign exchange earnings of displaced agricultural labor in alternative employment. Therefore, the support for reducing the barley subsidy is also strong on the foreign exchange account. (See Braverman, Hammer, and Jorgensen 1985 for further details of the Cyprus case.)

VI. ANALYSIS OF LARGE CHANGES IN POLICIES

All of the above results are valid for small changes in policy instruments but may give obviously wrong answers when large changes are contemplated. This is due to the linear nature of the model. For example, if the price of barley were to be raised to actual world price levels, the supply of beef would be negative in the above model. In the earlier, nonlinear version of the model, this problem does not arise since the functional forms prevent nonnegative values from occurring. This brings out a fundamental dilemma in the use of economic theory in policy analysis. We generally know very little about supply and demand functions other than their behavior near observed equilibria. Each of the functional forms used in empirical analysis—even the so-called flexible ones—has inherent implications for the behavior of demand over large price movements which have little to do with actual behavior (see Caves and Laurits 1980). Therefore, while a nonlinear model may give superficially more plausible results, it is not based more firmly on any economic principles. In either case, confidence can only be placed in small deviations from the status quo, or in other words, in the direction of change of the economy.

Conversely, policymakers often may want or need to know something about the outcome of policy reforms which create large changes in important economic variables. In Cyprus, the consumer price of feed is often less than half the world price, and the consequences of completely removing the subsidy might be just such a change. In this case, only very approximate answers should be given (and even then, not by the faint of heart). One possible way of handling large changes in policy is in the context of the sensitivity analysis. For large changes in prices, the relevant elasticities would be the arc elasticities between relatively distant points on the supply or demand curves. While direct evidence on these values is difficult to come by, a variety can be used. The choice may be guided by specific functional forms. The simplest example would be the constant elasticity form by which the arc elasticity, e , corresponding to a point elasticity, ϵ , over a price range of P_0 to P_1 would be calculated as

$$e = \frac{\left(\frac{P_1}{P_0}\right)^\epsilon - 1}{\frac{P_1}{P_0} - 1}$$

In the Cyprus case, if the price of barley doubles, the arc elasticity of demand should be changed to 0.52 in order to correspond to a point elasticity of 0.6 estimated at the new (doubled) price. Other functional forms might be used or, alternatively, generous bounds might be placed on the parameters in the sensitivity analysis. A second way to approximate larger policy changes is to solve the model iteratively. A small change can be made and followed by a recalculation of levels of variables and possibly of elasticities corresponding to a particular chosen functional form before a second step is taken. This can be repeated for more small steps until the desired policy change is reached. Subsequent work will examine the possibility of implementing this method in the software.

VII. USE IN POLICY DIALOGUE

The discussion of sensitivity analysis and of large policy changes underscores the importance of being able to do repeated policy experiments in order to get the feel of the effects of policy. This intuitive understanding is essential in policy analysis both as a check on the realism of calculated effects and as a way to explain the policy consequences to noneconomists. We have developed user friendly software for use on personal computers which allows these sensitivity analyses to be performed quickly and easily without requiring great familiarity with computers. These models should not and are not to be perceived as forecasting tools.

A valuable use of this type of model is to facilitate an informed discussion of policy outcomes between people with different points of view. If the basic structure of the economy can be agreed upon, the different assumptions necessary to support different expected outcomes can be made explicit. This can be used to focus discussion between World Bank economists, between ministries within the country, or between government and Bank economists. Of course, numbers can be used out of context, and the abuse of models by political advocates is frequent. However, our judgment is that the advantages of our approach in providing a consistent framework and quantitative structure for discussion greatly outweigh the drawbacks from such possible misuse.

APPENDIX

Market Equilibrium Conditions

Variables for equations which follow:

Superscripts and subscripts:

ℓ = lamb (or sheep)

b = beef (or cows)

m = milk

cm = cow's milk

sm = sheep's milk

p	= pork and poultry
w	= wheat (or bread)
F	= frozen lamb
B	= barley
H	= hay
P_i	= price of commodity i
$\bar{P}_i$	= producer price of i (if different from consumer price)
P_i^*	= world price of i
D_i	= consumer demand for i
S_i	= supply of i
M_i	= imports of i
K_i	= stock of animal i
D_i^j	= factor demand for commodity i in the production of j
π_i	= total profit in activity i
θ	= returns to scale parameter
r	= rate of return to capital (constant)
q	= conversion factor of wheat and barley to straw
a	= conversion factor of sheep to cow's milk equivalent
a_i	= Laspeyre index weights in government cost of living calculation for commodity i
w_i	= share of consumer budget spent on i
WB	= government wage bill
dC/Y	= compensatory variation as proportion of income.

Beef:

$$S_b(P_b, P_B, P_H, K_b) + M_b = D_b(P_\ell, P_b, P_m, P_p, P_w, P_F)$$

Fresh lamb:

$$S_\ell(P_\ell, P_B, P_H, K_\ell) = D_\ell(P_\ell, P_b, P_m, P_p, P_w, P_F)$$

Frozen lamb:

$$M_F = D_F(P_\ell, P_b, P_m, P_p, P_w, P_F)$$

Milk:

$$S_{cm}(P_m, P_H, P_B, K_\ell) + aS_{sm}(P_m, P_H, P_B, K_\ell) = D(P_\ell, P_b, P_m, P_p, P_w, P_F)$$

Livestock equilibrium:

$$\frac{\theta \pi_b}{K_b} + \frac{\theta \pi_m}{K_b} = r$$

Pork and poultry:

version I:

$$S_p(P_p, P_B) = D_p(P_t, P_b, P_m, P_p, P_w, P_F)$$

version II:

$$S_p(P_p, P_B) + M_p = D_p(P_t, P_b, P_m, P_p, P_w, P_F)$$

Wheat:

$$S_w(\bar{P}_B, \bar{P}_w, P_H) + M_w = D_w(P_t, P_b, P_m, P_p, P_w, P_F)$$

Barley:

$$S_B(\bar{P}_B, \bar{P}_w, P_H) + M_B = D_B^l(P_t, P_B, P_H, K_\ell) + D_B^b(P_b, P_B, P_H, K_b) + D_B^{cm}(P_m, P_B, P_H, K_b) \\ + D_B^{sm}(P_m, P_b, P_H, K_\ell) + D_B^g(P_p, P_B)$$

Hay:

$$S_H(\bar{P}_B, \bar{P}_w, P_H) + qS_B(P_B, P_w, P_H) + qS_w(P_B, P_w, P_H) = D_H^l(P_t, P_B, P_H, K_\ell) \\ + D_H^b(P_b, P_B, P_H, K_b) + D_H^{cm}(P_m, P_B, P_H, K_b) + D_H^{sm}(P_m, P_b, P_H, K_\ell)$$

The system is solved by totally differentiating the equations with respect to both exogenous and endogenous variables.

Government Accounts

Barley consumer deficit:

$$(P_B^* - P_B)(D_B^l + D_B^b + D_B^{cm} + D_B^{sm} + D_B^g)$$

Barley producer deficit:

$$(\bar{P}_B - P_B^*)S_B$$

Wheat consumer deficit:

$$(P_w^* - P_w)D_w$$

Wheat producer deficit:

$$(\bar{P}_w - P_w^*)S_w$$

Change in government wage bill:

$$\Delta WB = \left(\frac{\sum a_i P_i}{\sum a_i P_i^0} - 1 \right) WB$$

Cost of living:

$$\sum w_i \frac{dP_i}{P_i} = \frac{dC}{Y}$$

REFERENCES

- Ansell, David, Claire Bishop, and Martin Upton. 1983. "The Role of Part Time Farming in the Economic Development of Cyprus." Reading, England: University of Reading. Processed.
- Arrow, Kenneth J. 1964. "The Role of Securities in the Optimal Allocation of Risk Bearing." *Review of Economic Studies* 31, April: 91-96.
- Atkinson, A. B., and Joseph E. Stiglitz. 1980. *Lectures in Public Economics*. New York: McGraw-Hill.
- Barnum, Howard N., and Lyn Squire. 1980. *A Model of an Agricultural Household: Theory and Evidence*. World Bank Staff Occasional Paper 27. Baltimore, Md.: Johns Hopkins University Press.
- Bhagwati, J., and T. N. Srinivasan. 1973. "The General Equilibrium Theory of Effective Protection and Resource Allocation." *Journal of International Economics* 3, no. 3: 259-82.
- Braverman, Avishay, and Jeffrey S. Hammer. 1986. "Multimarket Analysis of Agricultural Pricing Policies in Senegal." In Inderjit Singh, Lyn Squire, and John Strauss, eds., *Agricultural Household Models: Extensions, Applications, and Policy*. Baltimore, Md.: Johns Hopkins University Press.
- Braverman, Avishay, Jeffrey S. Hammer, and Choong Yong Ahn. Forthcoming. "Multimarket Analysis of Agricultural Pricing Policies in Korea." In David Newbery and Nicholas Stern, eds., *The Theory of Taxation for Developing Countries*. New York: Oxford University Press.
- Braverman, Avishay, Jeffrey S. Hammer, and Erika Jorgensen. 1985. *An Economic Analysis of Reducing Input Subsidies to the Livestock Sector in Cyprus*. World Bank Staff Working Paper 782. Washington, D.C.
- Braverman, Avishay, Jeffrey S. Hammer, and Antonio Salazar Brandao. Forthcoming. "Economic Analysis of Agricultural Pricing Policies in Brazil: The Wheat and Soybean Case." *Revista Brasileira de Economia*.
- Braverman, Avishay, Jeffrey S. Hammer, and Jonathan J. Morduch. 1985. "Economic Analysis of Wheat and Maize Price Policies in Hungary." World Bank Agriculture and Rural Development Department Discussion Paper. Processed.
- Caves, Douglas, and Christian Laurits. 1980. "Global Properties of Flexible Functional Forms." *American Economic Review* 70, no. 3 (June): 422-32.
- Deaton, Angus. 1984. "Issues in the Methodology of Multi-Market Analysis of Agricultural Pricing Policies." Woodrow Wilson School, Princeton University. Processed.
- Deaton, Angus, and John Muellbauer. 1980. "An Almost Ideal Demand System." *American Economic Review* 70, no. 3 (June): 312-26.
- Harberger, Arnold C. 1971. "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay." *Journal of Economic Literature* 9, no. 3 (September): 785-97.
- Kirchner, J., I. Singh, and L. Squire. 1984. "Agricultural Pricing and Marketing Policies in Malawi." World Bank Country Policy Department Discussion Paper. Processed.
- Lau, Lawrence. 1976. "A Characterization of the Normalized Restricted Profit Function." *Journal of Economic Theory* 12, no. 1 (February): 161-63.
- McFadden, Daniel. 1978. "Cost, Revenue and Profit Functions." In D. McFadden and

- M. Fuss, eds., *Production Economics: A Dual Approach to Theory and Applications*. Amsterdam: North-Holland.
- Panayiotou, George S. 1983. "Demand on Price Analysis of Livestock Production in Cyprus." Nicosia, Cyprus: Agricultural Research Institute, Ministry of Agriculture.
- Stiglitz, Joseph E. 1982. "Taxation and Agricultural Pricing Policies, Cost Benefit Analysis, and the Foreign Exchange Constraint." Princeton, N.J.: Princeton University Department of Economics. Processed.
- Tolley, George S., Vinod Thomas, and Chung Ming Wong. 1982. *Agricultural Price Policies and the Developing Countries*. Baltimore, Md.: Johns Hopkins University Press.
- Yotopoulos, P. A., and L. J. Lau. 1974. "On Modelling the Agricultural Sector in Developing Countries: An Integrated Approach of Micro and Macroeconomics." *Journal of Development Economics* 1, no. 2: 105-28.

Devaluation, Fiscal Deficits, and the Real Exchange Rate

Mohsin S. Khan and J. Saul Lizondo

This article examines the use of fiscal policies to sustain the effects of a nominal devaluation on the real exchange rate. It is shown that the magnitude of the change in the real exchange rate depends not only on the size of the devaluation and the degree of fiscal adjustment but also on the means by which the fiscal deficit is reduced. The change in the nominal exchange rate necessary to maintain the depreciation of the real exchange rate will depend on whether the fiscal deficit is eliminated by increasing taxes or by reducing government expenditures on traded and nontraded goods. The required depreciation of the domestic currency will be larger if the fiscal deficit is reduced by increasing taxes than it will be if the deficit is cut by lowering government expenditures. Further, the depreciation would be smaller if the cuts in expenditure fell on traded rather than nontraded goods. This result implies that the authorities must ensure consistency between exchange rate action and policies to reduce fiscal imbalances in order to achieve a desired level of the real exchange rate necessary to attain balance of payments equilibrium.

The real exchange rate represents a key relative price in the economy and policies to change it are often the centerpiece of adjustment programs designed to improve international competitiveness and shift resources toward the production of tradable goods. Consequently, it is critical for policymakers to have some idea of the magnitude and time path of the likely response of the real exchange rate to nominal exchange rate action. It is a well-accepted proposition, however, that a nominal devaluation will only have a transitory effect on the real exchange rate. In the long run domestic wages and prices will rise by the full amount of the devaluation and the real exchange rate will return to its original level. To alter the real exchange rate on a permanent basis, therefore, devaluation has to be supplemented by policies that restrain the increase in domestic factor costs that results from a devaluation.

Generally speaking, the effects of a devaluation on the real exchange rate can be decomposed into two separate effects. First, there is the short-run response

Mohsin S. Khan is at the World Bank. J. Saul Lizondo is at the University of Tucumán, Argentina. The authors are grateful to Mario Blejer, Willem Buiter, Jonathan Eaton, Nadeem U. Haque, Ricardo Martin, Peter Montiel, and Assaf Razin for helpful comments.

Copyright © 1987 by the International Bank for Reconstruction and Development / THE WORLD BANK.

that reflects the initial increase in the domestic price level. This "first-order" effect has been shown to depend on the share of traded goods in total expenditure, and the relative price elasticities of the demand and supply of nontradable goods. As a first approximation, assuming that the price of nontraded goods is constant in the short run, the impact of a devaluation on domestic prices can be estimated as simply the product of the exchange rate change and the share of traded goods in total expenditures. The depreciation in the real exchange rate would therefore be equal to the nominal devaluation adjusted for the increase in domestic prices. Second, there is the longer-run response of the real exchange rate which would depend on how prices of nontraded goods respond to shifts in resources away from the nontraded sector. The more rapid is the increase in nontraded goods prices, other things being equal, the more quickly would the effects of a devaluation on the real exchange rate be dissipated.

This two-stage response of the real exchange rate to devaluation has been discussed in a number of theoretical studies, and it is clear from these that a sustainable change in the real exchange rate requires policies that bring about a change in *real* macroeconomic aggregates (see, for example, Dornbusch 1974; Krueger 1974; Connolly and Taylor 1976; Rodriguez 1978; Blejer 1979; and Edwards 1985). Since overvaluation of the exchange rate and the accompanying balance of payments problems can typically be traced back to imbalances in the government budget, measures to reduce the fiscal deficit are one example of policies that can have a long-run effect on the real exchange rate. Generally speaking, the extent to which a devaluation will affect the real exchange rate, as well as the length of time over which the effects persist, will be a direct function of the supporting fiscal policies that are put in place (see Berglas and Razin 1973 and Montiel 1986).

The purpose of this article is twofold. First, it examines the relationship between nominal and real exchange rates. This is done by formulating a relatively simple theoretical model that combines several arguments that have been made on the subject into one convenient framework. With this model it is possible to determine the real exchange rate response to a devaluation in both the short and long run. Second, we use this model to analyze the effects of a reduction in the fiscal deficit on the real exchange rate. It is shown that the effects will depend on *how* the fiscal deficit is reduced—whether through increases in taxes or reductions in expenditures. Finally, one can also ascertain the effects of combining devaluation with deficit-reducing measures within this relatively simple framework. Devaluation speeds up the process of adjustment and thus reduces the loss of reserves that would typically accompany policies that focused exclusively on eliminating fiscal deficits. Our general results point clearly to the need for close coordination between exchange rate and fiscal policies in any attempt to alter the real exchange rate in a predictable fashion and to achieve balance of payments equilibrium.

In the following section we describe the model and use it to determine the effects on the real exchange rate of devaluation and policies to reduce the fiscal

deficit. We also briefly discuss the case where the two types of policies are combined, as they would be in any real situation. The concluding section discusses some of the policy implications of the exercise.

I. FRAMEWORK OF ANALYSIS

In this section we start by formulating a simple two-good, dependent economy model that has come to be used extensively in the study of open-economy macroeconomic issues (see, for example, Dornbusch 1974; Rodriguez 1978; and Liviatan 1979). We then trace out the effects of a devaluation in isolation on the real exchange rate and the domestic price level. Here we are essentially interested in reproducing the main theoretical results obtained in previous work on the subject. The role of policies designed to reduce the fiscal deficit are then examined. The final part of this section deals with the joint effects of devaluation and fiscal policies on the real exchange rate and the balance of payments.

Theoretical Model

Consider an economy that produces and consumes traded and nontraded goods. The domestic currency price of traded goods, P_T , is equal to the foreign currency price of these goods in the world market, P_T^W , times the nominal exchange rate, E , defined as the domestic currency price of foreign currency. Since the economy is assumed to be small, it does not affect world market prices. The assumption that the terms of trade are given to the country allows us to aggregate importable and exportable goods into a single composite traded good. Assuming also that world prices are constant, units can be chosen so that $P_T^W = 1$. Thus, the domestic currency price of traded goods is equal to the nominal exchange rate, E . The domestic currency price of nontraded goods, P_N , conversely, is endogenously determined by the condition of equilibrium in the nontraded goods market. The relative price between traded and nontraded goods, E/P_N will be referred to as the "real exchange rate" and denoted by e . An increase in e represents a real depreciation, and vice versa.

It is assumed that the supply of factors of production is fixed and that factor prices are flexible. Because of this perfect price flexibility, full employment is always maintained. These assumptions imply that production of traded and nontraded goods, Y_T and Y_N , respectively, depends only on the real exchange rate (e):

$$(1) \quad Y_T = Y_T(e) \quad Y_N = Y_N(e) \quad \frac{dY_T}{de} > 0; \quad \frac{dY_N}{de} < 0$$

Private sector consumption of traded and nontraded goods, C_T and C_N , respectively, depends on the real exchange rate and on the real stock of wealth, a , which is measured in terms of the domestic currency price of traded goods, E .

$$(2) \quad C_T = C_T(e, a) \quad C_N = C_N(e, a)$$

$$\text{with } \frac{\partial C_T}{\partial e} < 0; \quad \frac{\partial C_T}{\partial a} > 0; \quad \frac{\partial C_N}{\partial e} > 0; \quad \frac{\partial C_N}{\partial a} > 0.$$

Public sector total expenditure, g , and taxes, t , are defined in terms of traded goods. Allocation of total government expenditures between traded goods, g_T , and nontraded goods, g_N , is exogenous and given as

$$(3) \quad g_T + g_N = g$$

There are only two assets in the economy, namely, domestic money, M , and foreign money, f . Traded goods producers earn foreign exchange when production, Y_T , is greater than domestic consumption, C_T , and government purchases, g_T . Neither domestic nor foreign money is assumed to be interest bearing, and we assume that individuals do not acquire physical assets or bonds. All savings, therefore, is in the form of domestic and foreign money. The real stock of domestic money in terms of traded goods, M/E , will be denoted by m . Therefore, total real wealth is given by:

$$(4) \quad a = m + f$$

Private sector domestic money holdings are a fraction, λ , of total assets, determined by the expected rate of depreciation of the nominal exchange rate, which is assumed to be exogenous and is denoted by $\hat{E}^*$.¹ (The symbol “^” over a variable signifies a percentage change, $\hat{x} \equiv \Delta x/x$.)

$$(5) \quad m = \lambda(\hat{E}^*)a$$

with $d\lambda/d\hat{E}^* < 0$ and $0 < \lambda(\hat{E}^*) < 1$.

The sign restrictions imply that the higher the expected rate of depreciation of the domestic currency, the smaller will be the fraction of wealth that is held in the form of domestic money. Since it is assumed that the nominal exchange rate is fixed by the central bank, changes in the expected rate of depreciation of the domestic currency cause the private sector to exchange foreign money for domestic money, and vice versa. The transactions take place with the central bank at the fixed exchange rate, thereby affecting the composition but not the level of private sector real wealth.

The short-run position of the economy is determined by the condition of equilibrium in the nontraded goods market:

$$(6) \quad Y_N(e) = C_N(e, a) + eg_N$$

1. Models with endogenous expected rates of depreciation are examined in the literature on balance of payments crises (see, for example, Flood and Garber 1984 and Blanco and Garber 1986). Edwards and Khan (1985) also discuss this issue and point out the difficulties involved in incorporating endogenous expectations of devaluation.

Figure 1. *Determination of Long-Run Equilibrium*

Equation 6 is represented by the downward sloping curve, NN , in Figure 1. The higher the level of real wealth, the greater is the private sector demand for nontraded goods, and thus the lower must be the real exchange rate (that is, the relative price of traded goods) in order to reduce demand and increase the supply of nontraded goods so as to maintain equilibrium. Similarly, a reduction in real wealth lowers the demand for nontraded goods and this requires a higher real exchange rate.

The nontraded goods market is always in equilibrium; that is, the economy is always on curve NN . In the short run the level of real wealth is predetermined and curve NN indicates the resulting real exchange rate. In the steady state, however, the level of real wealth is endogenous. Hence, it is necessary to examine the evolution of real wealth in order to determine the steady-state equilibrium. From equation 4, we can obtain the rate of change of wealth,

$$(7) \quad \dot{a} = m + f$$

where a dot over a variable indicates its derivative with respect to time. As long as there is no change in the nominal exchange rate, we have

$$(8) \quad \dot{m} = \frac{\dot{M}}{E}$$

From the balance sheet of the banking system, the change in the nominal stock of domestic money will be

$$(9) \quad \dot{M} = \dot{R} + \dot{D}$$

where $\dot{R}$ is the change in the stock of international reserves, measured in terms of domestic currency, and $\dot{D}$ is domestic credit creation. The change in the stock of international reserves is described by

$$(10) \quad \dot{R} = E[Y_T(e) - C_T(e, a) - g_T] - E\dot{f}$$

where the first term on the right hand side is the current account of the balance of payments, and the second term is the capital account reflecting the fact that an increase in private sector holdings of foreign currency will draw down central bank reserves. We assume that the government cannot finance its deficit by borrowing from the nonbank public or from abroad. We further assume that there is no credit extended to the private sector, although introducing this variable into equation 11 would not affect the analysis. Domestic credit creation is used exclusively to finance the public sector budget deficit, so

$$(11) \quad \dot{D} = E(g_N + g_T - t)$$

Using equations 10 and 11 to replace $\dot{R}$ and $\dot{D}$ in equation 9, and then using the resulting expression and equation 8 to substitute for $\dot{m}$ in equation 7, we obtain

$$(12) \quad \dot{a} = Y_T(e) - C_T(e, a) + g_N - t$$

Equation 12 describes the evolution of real wealth. For the economy to be in equilibrium, the level of wealth must be constant, which requires that

$$(13) \quad Y_T(e) - C_T(e, a) = t - g_N$$

Equation 13 is represented by $\dot{a} = 0$ in Figure 1. To the left of curve $\dot{a} = 0$, C_T is lower and savings is positive so that real wealth increases ($\dot{a} > 0$). Correspondingly, to the right of $\dot{a} = 0$ savings would be negative and real wealth would be declining ($\dot{a} < 0$). Therefore, for any given initial level of real wealth, the economy moves along curve NN as indicated by the arrows until it reaches the intersection with the $\dot{a} = 0$ curve at point A , where equilibrium would be achieved.

Point A in Figure 1 represents a short-run (quasi) equilibrium of the economy. Whether this position is sustainable or not depends on the size of the country's reserve holdings. At point A there is a balance of payments deficit and the country is losing reserves; obviously there is a limit to the amount of reserves a country can lose. In equilibrium, the capital account of the balance of payments

(equation 10) is in balance, since a constant level of real wealth implies a constant desired stock of foreign money, unless there is a change in the expected rate of depreciation, $\hat{E}^*$, which will change desired relative holdings of m and f . In other words, once the level of real wealth is constant ($\dot{a} = 0$), then persistent capital outflows ($\dot{f} > 0$) require an ever increasing expected rate of depreciation to induce an offsetting decline in $m(\hat{E}^*)$. Conversely, persistent capital inflows require an ever increasing expected rate of appreciation.

Since we are interested in identifying sustainable long-run equilibrium positions, we assume that in such positions the expected rate of depreciation is equal to zero. Therefore, we need to focus only on the current account of the balance of payments, denoted by CA :²

$$(14) \quad CA = Y_T(e) - C_T(e, a) - g_T$$

Equilibrium in the current account requires

$$(15) \quad Y_T(e) - C_T(e, a) = g_T$$

Equation 15 is represented by curve $CA = 0$ in Figure 1. To the left of this curve there is a current account surplus, and to the right of this curve there is a current account deficit. Comparing equations 13 and 15, it is clear that the position of the curve $CA = 0$ with respect to the curve $\dot{a} = 0$ depends on whether g_T is higher or lower than $(t - g_N)$, that is, on whether there is a government budget deficit or surplus. In Figure 1 it is assumed that the budget is in deficit to start with, and therefore the curve $CA = 0$ is to the left of curve $\dot{a} = 0$. Hence, at point A the current account, and therefore the overall balance of payments, is in deficit. Using equation 13, which holds at point A, to replace $(Y_T - C_T)$ in equation 14, it follows that

$$(16) \quad CA = t - g_N - g_T$$

Therefore, at point A, the current account deficit, and thus the overall balance of payments deficit, is equal to the deficit in the government budget.³ The long-run stationary equilibrium of the economy would be reached when the two curves $\dot{a} = 0$ and $CA = 0$ collapse into one curve, which implies equilibrium in both the fiscal and balance of payments accounts.⁴

2. Note that as we have assumed that services flows are zero, the current account balance is identical to the trade balance.

3. If the exchange rate were crawling, instead of being fixed, the long-run rate of inflation would be positive and equal to the rate of crawl. As a result, the current account (and thus the overall balance of payments) deficit at the stationary equilibrium would be equal to the public sector budget deficit minus the inflation tax. This also implies that changes in the expected rate of devaluation would affect the balance of payment not only in the short run due to the immediate change in portfolio composition but also in the long run due to the change in the inflation tax.

4. In this model with a zero fiscal deficit and a positive growth rate, the debt-income ratio would steadily fall.

Effects of Devaluation

The effects of a devaluation can be examined with the aid of Figure 1. A nominal devaluation reduces the real stock of domestic money, and thus the level of real wealth. This causes a decline in demand, which in turn requires a depreciation of the real exchange rate to increase the demand and reduce the supply of nontraded goods so as to maintain equilibrium. For example, starting from point A, a devaluation would reduce the level of real wealth from a_0 to a_1 , causing a depreciation of the real exchange rate from e_0 to e_1 . After this initial impact, however, real wealth increases with reserves until it returns to a_0 , while the real exchange rate appreciates steadily until it comes back to e_0 .

The effect of a devaluation on the real exchange rate can be obtained from equations 5 and 6. Since the nominal stock of domestic money is predetermined at the time of the devaluation, the real stock of domestic money, (M/E) , declines by the amount of the devaluation. Therefore, as a fraction of wealth, $\lambda(\hat{E}^*)$, is held in the form of domestic money, the change in real wealth is given by:⁵

$$(17) \quad \hat{a} = -\lambda(\hat{E}^*)\hat{E}$$

where $\hat{E}$ is the actual rate of devaluation. In Figure 1, equation 17 would determine the extent of the decline in real wealth shown here as the movement from a_0 to a_1 . The magnitude of the depreciation of the real exchange rate (the rise from e_0 to e_1) is found from equation 6, which defines the curve NN. Differentiating equation 6 and using equation 17 to substitute for $\hat{a}$, we obtain

$$(18) \quad \hat{e} = \lambda(\hat{E}^*) \frac{\beta \eta_N^e}{\epsilon_N + \beta \eta_N^e + (1 - \beta)} \hat{E}$$

where

- η_N^a = nontraded goods demand elasticity with respect to real wealth (positive)
- η_N^e = nontraded goods demand elasticity with respect to the real exchange rate (positive)
- ϵ_N = nontraded goods supply elasticity with respect to the real exchange rate, defined to be positive ($\epsilon_N = -dY_N/de \cdot e/Y_N$)
- β = share of the private sector in total consumption of nontraded goods, (C_N/Y_N).

As indicated in equation 18, there are several factors that affect the initial depreciation of the real exchange rate. Other things being equal, the higher the share of domestic money in real wealth, the larger is the decline in real wealth resulting from a devaluation, and therefore the larger must be the real depreciation that is necessary in order to maintain equilibrium in the nontraded goods market. Similarly, the larger the demand elasticity of nontraded goods with respect to wealth, the larger is the decline in the demand for nontraded goods

5. This equation is obtained by differentiating equation 4 with respect to E and then using equation 5 to replace m/a .

brought about by the reduction in real wealth, and thus the greater is the real depreciation that is required to maintain equilibrium. Conversely, the higher the supply elasticity and the relative price demand elasticity of nontraded goods, the smaller is the real depreciation that is needed to maintain equilibrium.

As mentioned above, after the initial effect, the real exchange rate appreciates steadily until it returns to its original level. The long-run neutrality of the real exchange rate with respect to a nominal devaluation follows from equations 6 and 13, which determine the long-run stationary equilibrium. Since both equations are independent of the level of the nominal exchange rate, the long-run real exchange rate does not change as a result of a devaluation. The evolution of the real exchange rate (e) through time after a devaluation is described in Figure 2, and most of the models in the literature imply a path for the real exchange rate similar to one described here (see, for example, Dornbusch 1974; Krueger 1974; Boyer 1977; Rodriguez 1978; Blejer 1979; Liviatan 1979; and Montiel 1986).

There are some studies, however, that yield different results for the effect of devaluation described by equation 18. For example, Jones and Corden (1976) show that in a Heckscher-Ohlin model with wages kept constant, a devaluation may produce a real appreciation if traded goods are labor-intensive. They also

Figure 2. *Behavior of the Domestic Price Level and the Real Exchange Rate*

show that this result cannot arise in a capital-specific model, and they then argue that a capital-specific model is better suited for the analysis of a devaluation. Lapan and Enders (1978), using a two-country model, show that a devaluation in the home country may appreciate its real exchange rate if the share of home country wealth held in home currency is lower than the share of the foreign country wealth held in the home country currency. Under these conditions, a devaluation of the home currency would redistribute wealth in the "wrong" direction, namely from the foreign country to the home country, leading to an appreciation of the real exchange rate in the home country.⁶ Kyle (1978), using a model with money and bonds in the utility function, shows that it is possible for a nominal devaluation to produce a short-run real appreciation due to substitution effects between bonds, goods, and money in the utility function. Finally, Stockman (1983), using an equilibrium two-period model, shows that a devaluation has no effect on the real exchange rate since the reduction in private real wealth arising from the "tax" on domestic money is compensated by the increase in the interest-bearing reserves of the government. Generally speaking, however, the results obtained here are consistent with those obtained by the majority of papers on the subject.

The effect of the devaluation on the price level can also be derived from our previous results. Assume that the price level is defined as:

$$(19) \quad P = E^{\alpha_T} P_N^{\alpha_N}$$

where α_T and α_N are the respective shares of traded and nontraded goods in total expenditures, $\alpha_T + \alpha_N = 1$. Therefore,

$$(20) \quad \hat{P} = \alpha_T \hat{E} + \alpha_N \hat{P}_N$$

but we know that

$$(21) \quad \hat{e} = \hat{E} - \hat{P}_N$$

In order to obtain the effect of a devaluation on the price level, we use equation 21 to substitute for $\hat{P}_N$ in 20 and then use 18 to substitute for $\hat{e}$ in the resulting expression and obtain

$$(22) \quad \hat{P} = \left\{ \alpha_T + \alpha_N \left[1 - \lambda(\hat{E}^*) \frac{\beta \eta_N^e}{\epsilon_N + \beta \eta_N^e + (1 - \beta)} \right] \right\} \hat{E}$$

If the expression in square brackets is less than one, the price level initially increases proportionally less than the exchange rate. This implies that the nominal price of nontraded goods either increases proportionally less than the rate of devaluation or declines.⁷ From equation 22 the effect on the price level is larger,

6. This result is valid only in the case of a large economy.

7. The nominal price of nontraded goods declines if the expression in square brackets in equation 22 is negative. Models that allow for this possibility include Krueger (1974), Connolly and Taylor (1976), Boyer (1977), and Blejer (1979). An initial decline in the nominal price of nontraded goods could theoretically result in an initial decline in the price level. However, this case is not plausible since it would require unusual values for the various parameters.

the higher are the share of traded goods in the price index (α_T), the nontraded goods supply elasticity (ϵ_N), and the nontraded goods demand elasticity with respect to the real exchange rate (η_N^s); and the lower are the nontraded goods demand elasticity with respect to wealth (η_N^w) and the share of domestic money in real wealth (λ).

After the initial impact, as real wealth rises, domestic prices increase steadily until they reach a new long-run level which is higher than the initial one by the same proportion as the nominal devaluation. Formally, the long-run effect on the price level can be obtained by using equation 21 to replace $\hat{P}_N$ in 20, and then setting $\hat{e} = 0$, a condition that was shown to hold in the long run. As a result, we obtain $\hat{P} = \hat{E}$, and the price level rises in the same proportion as the nominal exchange rate. Figure 2 shows the evolution of the price level following a devaluation.

Clearly the time path of the price level does not accord with the idea that a devaluation produces a once-and-for-all increase in the price level equal to the share of traded goods in the price index times the rate of devaluation. This would be the effect only if the price of nontraded goods does not change with a devaluation, which is equivalent to assuming that the various price and income elasticities in equation 18 (or the expression in brackets in equation 22) are equal to zero. Furthermore, as shown in Figure 2, following the initial jump the price level keeps rising for a period of time before it reaches a new stationary level.

Another implication of this model is that the effectiveness of a devaluation in producing a (transitory) real depreciation, and thus a (transitory) improvement in the current account of the balance of payments, depends on the extent to which the devaluation was anticipated. The higher the expected rate of depreciation, the lower the share of wealth held in domestic money, and thus the smaller the reduction in real wealth brought about by the devaluation. In the limit, if the private sector holds its entire portfolio in foreign currency, a devaluation would be immediately reflected in a proportional increase in all prices without any effect on the real exchange rate. These results can be obtained formally in equations 18 and 22 by setting $\lambda(\hat{E}^*)$ equal to zero.⁸

Effects of Reducing Fiscal Deficits

Since a devaluation by itself produces only a transitory depreciation of the real exchange rate and a transitory improvement in the balance of payments, other policies are needed in order to make these effects permanent. As was shown above, under the assumption of this model the long-run steady state balance of payments deficit is equal to the public sector budget deficit. Therefore, a perma-

8. Since most of the early models included only domestic money, the expected rate of depreciation played no role (see, for example, Dornbush 1974; Krueger 1974; and Connolly and Taylor 1976). Isard and Porter (1977) stress the need to take into account the anticipatory actions of the private sector diversification of portfolios. Models that do incorporate foreign assets include, among others, Boyer (1977) and Lapan and Enders (1978).

nent improvement in the balance of payments necessarily requires a reduction in the public sector deficit, and in particular, long-run balance of payments equilibrium requires government budget equilibrium.

This result can also be illustrated by using Figure 1. Recall that the initial equilibrium of the economy is determined by the intersection of curves NN and $\dot{a} = 0$, and that long-run balance of payments equilibrium only holds for points on the $CA = 0$ curve. Therefore, unless NN and $\dot{a} = 0$ intersect at some point on the $CA = 0$ curve, the initial equilibrium will be inconsistent with balance of payments equilibrium. Curve $\dot{a} = 0$ is defined by equation 13, while curve $CA = 0$ is defined by equation 15. If the public sector deficit is eliminated, the right hand side in both equations will be the same, and curves $\dot{a} = 0$ and $CA = 0$ will collapse into one curve. Under this condition, the intersection of curves $\dot{a} = 0$ and NN will necessarily be at a point on the curve $CA = 0$.

Although the objective of long-run balance of payments equilibrium is attained with the elimination of the fiscal deficit, irrespective of the specific fiscal policies used, the mix of fiscal policies is important for the determination of the long-run real exchange rate. Different combinations of increases in taxes and reductions in public sector expenditure on traded and nontraded goods, all of which create long-run balance of payments equilibrium, will imply different values for the long-run real exchange rate. This can be shown in the framework of Figure 3, which reproduces the original curves NN , $\dot{a} = 0$, and $CA = 0$.

Let us assume first that the fiscal deficit is eliminated entirely by increasing taxes. From equations 6, $Y_N(e) = C_N(e, a) + eg_N$, and 15, $Y_T(e) - C_T(e, a) = g_T$, it is clear that curves NN and $CA = 0$ do not shift, while equation 13, $Y_T(e) - C_T(e, a) = t - g_N$, indicates that curve $\dot{a} = 0$ shifts to the left. From our previous results we know that curve $\dot{a} = 0$ will collapse into the curve $CA = 0$. Thus, the new long-run equilibrium will be at a point such as B , where the NN and the new $\dot{a} = 0$ curves intersect. Therefore, if the fiscal deficit is eliminated by increasing taxes, real wealth will decline from a_A to a_B and the real exchange rate will depreciate from e_A to e_B .

Alternatively, let us assume that the fiscal deficit is eliminated entirely by reducing public sector expenditure on traded goods. From equations 6 and 13, curves NN and $\dot{a} = 0$ do not shift, while from equation 15 it follows that curve $CA = 0$ shifts to the right until it collapses into curve $\dot{a} = 0$. Thus, the new long-run equilibrium will be point A , which is the same as the previous equilibrium. Therefore, eliminating the fiscal deficit by reducing public sector expenditures on traded goods does not require either a reduction of wealth or a depreciation of the real exchange rate.

Finally, let us assume that the fiscal deficit is eliminated by reducing public sector expenditures on nontraded goods. Equation 15 indicates that curve $CA = 0$ does not shift, while equations 6 and 13 indicate that curve NN shift upward to $N'N'$ and curve $\dot{a} = 0$ shifts upward until it collapses into curve $CA = 0$. Since curve $\dot{a} = 0$ shifts more than curve NN , the new intersection point C must

Figure 3. Effects of Different Policies to Reduce the Fiscal Deficit and the Real Exchange Rate

be to the northwest of point A.⁹ Therefore, eliminating the public sector deficit by reducing public sector expenditures on nontraded goods implies a decline in real wealth from a_A to a_C and a depreciation of the real exchange rate from e_A to e_C .

9. The upward shift in curve $\dot{a} = 0$ is obtained by differentiating equation 13 with respect to g_N while keeping a constant. The result is

$$-\frac{de}{dg_N} = \left(\frac{dy_T}{de} - \frac{\partial C_T}{\partial e} \right)^{-1}.$$

Following the same procedure with equation 6, the upward shift in curve NN is

$$-\frac{de}{dg_N} = \left(\frac{\partial C_N}{\partial e} \frac{1}{e} - \frac{\partial y_N}{\partial e} \frac{1}{e} + \frac{g_N}{e} \right)^{-1}.$$

Since producers' profit maximization implies $dy_N/de = -e \cdot dy_T/de$, and consumers' utility maximization implies $\partial C_N/\partial e = -e \cdot \partial C_T/\partial e$, it follows that curve $\dot{a} = 0$ shifts up by more than curve NN .

The various means of eliminating the fiscal deficit have different implications regarding the level of wealth and the real exchange rate for the following simple reason. In each case the current account (and balance of payments) deficit is eliminated, which can be done either directly through a reduction in public sector expenditures on traded goods or indirectly through a reduction in wealth and a depreciation of the real exchange rate, which would reduce private sector demand and increase supply of traded goods. If the deficit is eliminated directly through a reduction in public sector expenditures on traded goods, there is no need for changes in relative prices and wealth that would change the behavior of the private sector. If the deficit is eliminated in some other way, however, a combination of reduction in real wealth and depreciation of the real exchange rate is required. The elimination of the deficit through a reduction in public sector expenditures on nontraded goods leads to a larger real depreciation (and a smaller reduction in real wealth) than an increase in taxes because in the former case the exogenous reduction in demand falls entirely on nontraded goods, which requires a larger decline in its relative price.¹⁰

Obviously, combinations of the three policies examined above could also be used to eliminate the deficit. For example, a combination of a reduction in public sector expenditure in traded and nontraded goods would imply a new equilibrium along the segment *AC* in Figure 3. A combination of an increase in taxes and a reduction in public sector expenditure on traded (nontraded) goods would produce a new equilibrium along the segment *AB* (*BC*). Finally, a combination of the three policies would imply a new equilibrium somewhere within the triangle *ABC*. In other words, for each combination of the three policies mentioned above there is a corresponding long-run equilibrium position in triangle *ABC*.¹¹ This implies that when setting targets for a reduction in the fiscal

10. Other researchers have also examined the relationship between fiscal policy and the real exchange rate and have reached similar conclusions (see, for example, Berglas and Razin 1973; Bruno 1976; and Greenwood 1984). Berglas and Razin note that restrictive fiscal policy may lead to an appreciation of the real exchange rate if nontraded goods are inferior goods. This could also happen in our model, since inferior nontraded goods imply a positively sloped *NN* curve. However, we do not consider the case of inferior nontraded goods because it also has some fairly implausible implications. For example, a nominal devaluation by itself would cause a transitory real appreciation—see equation 18 with a negative η_k . For the same reason, an exogenous increase in the money supply would cause a real depreciation. Furthermore, if the *NN* schedule becomes steeper than curve $\dot{a} = 0$, the system is divergent; if the economy is not in point *A*, it will move continuously away from *A*. Kapur (1981), using a growth model, shows that changes in government expenditures on traded capital goods may affect the steady state real exchange rate since this would change the steady state stock of capital per capita. Capital accumulation is not allowed for in our model. Kimbrough (1985) examines the relationship between fiscal policy and real exchange rates for countries under freely floating exchange rates.

11. It is also possible to obtain a long run equilibrium position outside the triangle *ABC* if the policy package that eliminates the fiscal deficit includes some expansionary policy. For example, if the policy package includes an increase in taxes but also a (smaller) increase in public sector expenditure in traded goods so that the fiscal deficit is eliminated, the long-run equilibrium position would be located on the *NN* curve to the northwest of point *B*, outside the triangle *ABC*. The consideration of this type of policy packages does not affect the main conclusions of the article.

deficit and a real depreciation in order to attain a given balance of payments improvement, the target for the real exchange rate cannot be set independently from the specific way in which the fiscal deficit is to be reduced.

Combining the Effects of Devaluation and Fiscal Adjustment

We have shown that long-run balance of payments equilibrium can be attained by eliminating the fiscal deficit, which generally implies a reduction in real wealth and a depreciation of the real exchange rate. For this reduction in real wealth and depreciation of the real exchange to take place, it is not necessary to devalue the domestic currency. If the fiscal deficit is eliminated, but the domestic currency is not devalued, there will be an adjustment period that will endogenously bring about the necessary changes in real wealth and the real exchange rate. For example, if the fiscal deficit is eliminated entirely by increasing taxes, the economy will adjust from point *A* to point *B* along the *NN* curve, causing a gradual reduction in real wealth and a gradual depreciation of the real exchange rate. If the deficit is eliminated entirely by reducing the expenditure on non-traded goods, the real exchange rate will jump from *A* to *D*, and then both the real exchange rate and real wealth will adjust gradually along the *NN* curve until they reach point *C*. If the deficit is eliminated entirely by a reduction in public sector expenditure in traded goods, however, there is no need for an adjustment process and the economy will attain a long-run equilibrium position immediately at the same point *A*. Combinations of the three types of fiscal policies will produce an initial jump in the real exchange rate, followed by a gradual depreciation.

Although the new long-run equilibrium can be attained without a devaluation of the domestic currency, the process of adjustment in which the real exchange rate depreciates gradually is accompanied by balance of payments deficits until the economy reaches the new long-run equilibrium. A devaluation of the domestic currency, however, can make the economy reach the new long-run equilibrium immediately, thereby preventing the loss of international reserves during the process of adjustment. Furthermore, in models that assume rigidities in prices and wages, a devaluation also helps to prevent unemployment, since a real depreciation without a devaluation would require a decline in the price of non-traded goods. If nontraded goods prices are sticky, output will decline and unemployment may result (see, for example, Dornbusch 1974 and Bruno 1978).

The magnitude of the devaluation that is required to immediately reach a new long-run equilibrium depends on the new equilibrium point; that is, it depends on the specific policies that are used to eliminate the fiscal deficit. For example, if only an increase in taxes is used, the magnitude of the devaluation must be such that real wealth falls from a_A to a_B . Conversely, if only a reduction in public sector expenditure in nontraded goods is used, real wealth must fall from a_A to a_C , which requires a smaller devaluation. In general, the lower the level of real wealth in the new long-run equilibrium, the larger is the devaluation needed to attain equilibrium immediately. This implies that the greater the reliance on

increasing taxes with respect to reducing public sector expenditures, the greater the needed devaluation. Conversely, the larger the reliance on reducing public sector expenditure on traded goods relative to the other policies, the smaller the required devaluation.

A devaluation that moves the economy immediately to equilibrium causes a once-and-for-all effect on the price level. The magnitude of this effect depends on the specific policies that were used to eliminate the fiscal deficit. From equation 20 and the definition of the real exchange rate, it follows that

$$(23) \quad \hat{P} = \hat{E} - \alpha_N \hat{e}$$

Equation 23 and Figure 3 can be used to compare the effect of the various policies on the price level. For example, the price level increases by more if the deficit is eliminated by increasing taxes, rather than reducing public sector expenditure in nontraded goods, under the assumption that both policies are accompanied by the devaluation required to attain long-run equilibrium immediately. This follows because to move from *A* to *B* (when increasing taxes) requires both a larger devaluation and a lower real depreciation than when moving from *A* to *C* (when $g_N = 0$) in Figure 3.

II. CONCLUSIONS

One of the familiar results that is captured in the framework of this paper is that devaluation by itself is not sufficient to achieve a real exchange rate objective, except perhaps in the very short term. Devaluation does have an initial effect on the real exchange rate, with the size depending on the substitution elasticities between traded and nontraded goods in consumption and production, and the share of domestic money in real wealth. The effects can only be sustained if supporting policies, such as measures to reduce the fiscal deficit and control of the growth of money, are adopted. These results are well known and what we have done here is to provide a unified treatment of the relationships between devaluation, fiscal deficits, and the real exchange rate.

The analysis in this article also highlighted a further aspect of the policy combination that is not as well-appreciated in the literature. That is, the magnitude of the nominal devaluation necessary to attain real exchange rate and balance of payments targets will depend on whether the fiscal deficit is reduced by increasing taxes, or by cutting government expenditures on traded and nontraded goods. In broad terms, the required depreciation of the domestic currency will be larger if the fiscal deficit is reduced by increasing taxes than by lowering government expenditures. Furthermore, the required nominal devaluation will be smaller if the cuts in expenditure fall on traded goods rather than nontraded goods.

The results here have an obvious bearing on the design of policy packages that are aimed at improving international competitiveness and the external balance of trade. It is shown that the size of the nominal devaluation and the way the

fiscal deficit is reduced cannot be considered as independent policy measures. The authorities have to ensure that consistency is maintained between the two types of policies to achieve a desired depreciation of the real exchange rate. If this is not done then the level of the real exchange rate and thus the impact on wealth would be indeterminate, even though the balance of payments deficit is eliminated. Simply proposing that the fiscal deficit should be reduced when devaluing the exchange rate is not enough to achieve a particular real exchange rate target; one also has to spell out the combination of methods by which the fiscal deficit is to be reduced.

REFERENCES

- Berglas, E., and A. Razin. 1973. "Real Exchange Rate and Devaluation." *Journal of International Economics* 3, no. 2 (May): 179-91.
- Blanco, H., and P. Garber. 1986. "Recurrent Devaluation and the Timing of Speculative Attacks." *Journal of Political Economy* 94, no. 1 (February): 148-66.
- Blejer, M. I. 1979. "Devaluation, Inflation, and the Balance of Payments: A Short-Run Monetary Approach." *Economic Record* 55, no. 148, (March): 33-40.
- Boyer, R. S. 1977. "Devaluation and Portfolio Balance." *American Economic Review* 67, no. 2 (March): 54-63.
- Bruno, M. 1976. "The Two-Sector Open Economy and the Real Exchange Rate." *American Economic Review* 66, no. 4 (September): 566-77.
- . 1978. "Exchange Rates, Import Costs, and Wage-Price Dynamics." *Journal of Political Economy* 86, no. 3 (June): 379-403.
- Connolly, M., and D. Taylor. 1976. "Adjustments to Devaluation with Money and Nontraded Goods." *Journal of International Economics* 6, no. 3 (August): 289-98.
- Dornbush, R. 1974. "Real and Monetary Aspects of the Effects of Exchange Rate Changes." In R. Z. Aliber, ed., *National Monetary Policies and the International Financial System*. Chicago: University of Chicago Press.
- Edwards, S. 1985. "Real Exchange Rate Misalignment in Developing Countries: Analytical Issues and Empirical Evidence." World Bank Country Policy Department. Processed.
- Edwards, S., and M. S. Khan. 1985. "Interest Rate Determination in Developing Countries: A Conceptual Framework." *IMF Staff Papers* 32, no. 3 (September): 377-403.
- Flood, R. P., and P. Garber. 1984. "Collapsing Exchange Rate Regimes: Some Linear Examples." *Journal of International Economics* 17, nos. 1-2 (August): 1-13.
- Greenwood, J. 1984. "Non-traded Goods, the Trade Balance, and the Balance of Payments." *Canadian Journal of Economics* 17, no. 4 (November): 806-23.
- Isard, P., and M. G. Porter. 1977. "A Note on the Monetarist Analysis of Devaluation." *Journal of International Economics* 7, no. 4 (November): 407-09.
- Jones, R. W., and W. M. Corden. 1976. "Devaluation, Non-flexible Prices, and the Trade Balance for a Small Country." *Canadian Journal of Economics* 9, no. 1 (February): 150-61.
- Kapur, B. K. 1981. "Traded Goods, Nontraded Goods, and the Balance of Payments: A Steady-State Analysis." *International Economic Review* 22, no. 1 (February): 167-78.

- Kimbrough, K. P. 1985. "An Examination of the Effects of Government Purchases in an Open Economy." *Journal of International Money and Finance* 4, no. 1 (March): 13–33.
- Krueger, A. O. 1974. "The Role of Home Goods and Money in Exchange Rate Adjustments." In Willy Salleskaerts, ed., *International Trade and Finance: Essays in Honor of Jan Tinbergen*. London: Macmillan.
- Kyle, J. F. 1978. "Finance Assets, Non-Traded Goods, and Devaluation." *Review of Economic Studies* 45, no. 1 (February): 155–63.
- Lapan, H., and W. Enders. 1978. "Devaluation, Wealth Effects, and Relative Prices." *American Economic Review* 68, no. 4 (September): 601–13.
- Liviatan, N. 1979. "A Diagrammatic Exposition of the Monetary Trade Model." *Economic Record* 55, no. 150 (September): 222–29.
- Montiel, P. 1986. "Long-Run Equilibrium in a Keynesian Model of a Small Open Economy." *IMF Staff Papers* 33, no. 1 (March): 28–59.
- Rodriguez, C. A. 1978. "A Stylized Model of the Devaluation-Inflation Spiral." *IMF Staff Papers* 25, no. 1 (March): 76–89.
- Stockman, A. C. 1983. "Real Exchange Rates under Alternative Nominal Exchange-Rate Systems." *Journal of International Money and Finance* 2, no. 2 (August): 147–66.

Subscription Information

As a professional journal, *The World Bank Economic Review* seeks to provide a wide distribution among economists and other social scientists of research results and lessons from World Bank experience, and thereby to contribute to a more fruitful and factually well-founded dialogue on issues in development economics.

To further this aim, subscriptions to the *Review* are being offered free of charge for the first two years of its publication. Readers are encouraged to circulate this volume among their colleagues and others interested in development issues. To have your name entered on the subscription list, please complete and submit one of the attached subscription request forms. You must renew your subscription in writing each year if you wish to remain on the distribution list.

The subscription request form includes a list of occupation codes and subject interest codes. Please check the appropriate codes on your subscription request form. This information is important to us in the management of the *Review's* distribution.

Request Form for One-Year Subscription to *The World Bank Economic Review*

Volume 1, Number 2, January 1987

Send to: World Bank Publications
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

OR World Bank Publications
66, avenue d'Iéna
75116 Paris, France

Name _____

Title _____

Organization name _____

Organization address _____

Country _____ Postal code _____

Occupation Codes

Please check only one.

- ☐ 01 International, regional organizations (including field offices), that is, WHO, UNDP
- ☐ 02 Central banks, finance and planning ministries
- ☐ 03 Other government agencies (including embassies)
- ☐ 04 Nongovernmental organizations (NGOs) and other associations
- ☐ 05 Research institutions
- ☐ 06 University faculty
- ☐ 07 Libraries
- ☐ 08 Bookstores
- ☐ 09 Commercial banks and other financial institutions (including savings and loan associations, security dealers, and so forth)
- ☐ 10 Business—manufacturing (goods) } Where both apply,
- ☐ 11 Business—consulting (services) } check “goods”
- ☐ 12 Student
- ☐ 13 News media
- ☐ 14 Other (please specify) _____

Subject Interest Codes

- ☐ 000 General economics: theory, history, systems
- ☐ 100 Economic growth: development, planning, fluctuations
- ☐ 200 Economic statistics
- ☐ 300 Monetary and fiscal theory and institutions
- ☐ 400 International economics
- ☐ 500 Administration, business finance, marketing, accounting
- ☐ 600 Industrial organization; technological change; industry studies
- ☐ 700 Agriculture, natural resources, rural development
- ☐ 800 Labor, population and human resources
- ☐ 900 Welfare programs; consumer economics; urban and regional economics
- ☐ A00 Related disciplines (please specify) _____

Request Form for One-Year Subscription to *The World Bank Economic Review*

Volume 1, Number 2, January 1987

Send to: World Bank Publications
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

OR World Bank Publications
66, avenue d'Iéna
75116 Paris, France

Name _____

Title _____

Organization name _____

Organization address _____

Country _____ Postal code _____

Occupation Codes

Please check only one.

- ☐ 01 International, regional organizations (including field offices), that is, WHO, UNDP
- ☐ 02 Central banks, finance and planning ministries
- ☐ 03 Other government agencies (including embassies)
- ☐ 04 Nongovernmental organizations (NGOs) and other associations
- ☐ 05 Research institutions
- ☐ 06 University faculty
- ☐ 07 Libraries
- ☐ 08 Bookstores
- ☐ 09 Commercial banks and other financial institutions (including savings and loan associations, security dealers, and so forth)
- ☐ 10 Business—manufacturing (goods) } Where both apply,
- ☐ 11 Business—consulting (services) } check “goods”
- ☐ 12 Student
- ☐ 13 News media
- ☐ 14 Other (please specify) _____

Subject Interest Codes

- ☐ 000 General economics: theory, history, systems
- ☐ 100 Economic growth: development, planning, fluctuations
- ☐ 200 Economic statistics
- ☐ 300 Monetary and fiscal theory and institutions
- ☐ 400 International economics
- ☐ 500 Administration, business finance, marketing, accounting
- ☐ 600 Industrial organization; technological change; industry studies
- ☐ 700 Agriculture, natural resources, rural development
- ☐ 800 Labor, population and human resources
- ☐ 900 Welfare programs; consumer economics; urban and regional economics
- ☐ A00 Related disciplines (please specify) _____

A WORLD BANK POLICY STUDY

Population Growth and Policies in Sub-Saharan Africa

Now Available in English and French

- Describes population policies and programs, assesses the prospects for policy changes to help slow population growth.
- Examines the consequences of population growth on labor productivity, the environment and social spending, and the health of mothers and children.
- Includes a population data supplement, bibliographical note, statistical annex, and forty-nine tables and figures.

112 pages, 8¹/₂" × 11"

Order your copy of *Population Growth and Policies in Sub-Saharan Africa* with the coupon.

New from the World Bank and Oxford University Press

Listen to the People

Participant-Observer Evaluation of Development Projects

Lawrence Salmen writes about the human side of economic development and describes the lives of the poor in the tradition of William Whyte, Robert Coles, and Oscar Lewis. But Salmen uses the tools of the social anthropologist for a very practical purpose—to enlist the participation of slum dwellers in the development process.

200 pages, 6¹/₄" × 8¹/₂"

Of Related Interest

Putting People First

Sociological Variables in Rural Development

Describes a culturally sensitive approach to the preparation, planning, and implementation of rural development projects.

1985, 444 pages, 6¹/₈" × 9¹/₄"

Order *Listen to the People* and *Putting People First* by completing the coupon that follows.

Please send me

- _____ copies of *Listen to the People* (OX520545) US\$17.95 each
_____ copies of *Putting People First* (BK520465) US\$24.95 each
_____ copies of *Industrialization and Growth: A Comparative Study* (OX520547)
US\$29.50 each

Please send me *Population Growth and Policies in Sub-Saharan Africa*.

- _____ copies in English (BK0773) US\$7.50 each
_____ copies in French (BK0833) US\$7.50 each

Subtotal _____

Air Mail outside the U.S. (US\$4.50 a copy) _____

Total _____

Enclosed is my check for US\$ _____
(or equivalent currency).

Charge my: ☐ VISA ☐ MasterCard ☐ American Express ☐ Choice (Credit cards
accepted only for orders addressed to Washington, D.C.)

Number _____ Expiration _____

Signature _____

Name _____

Title _____

Address _____

State/Province _____ Postal Code _____

Country _____ Telephone _____

Return to:

World Bank Publications

Department No. 0552
Washington, D.C. 20073-0552

or

World Bank Publications

66, avenue d'Iéna
75116 Paris, France

Orders outside the United States will be forwarded to local distributors. Prices may vary from country to country.