# THE WORLD BANK ECONOMIC REVIEW

Volume 8 May 1994 Number 2






Public Disclosure Authorized

# THE WORLD BANK ECONOMIC REVIEW



*The World Bank Economic Review* is a professional journal for the dissemination of World Bank–sponsored research that informs policy analyses and choices. It 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. 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 at least one external member 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 Editor.

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. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

Comments or brief notes responding to *Review* articles are welcome and will be considered for publication to the extent that space permits. Please direct all editorial correspondence to the Editor, *The World Bank Economic Review*, The World Bank, Washington, D.C. 20433, U.S.A.

*The World Bank Economic Review* is published three times a year (January, May, and September) by the World Bank. Single copies may be purchased at $10.95. Subscription rates are as follows:

| | *Individuals* | *Institutions* |
|---|---|---|
| 1-year subscription | US$25 | US$45 |
| 2-year subscription | US$46 | US$86 |
| 3-year subscription | US$65 | US$125 |

Orders should be sent to: World Bank Publications, Box 7247–7956, Philadelphia, PA 19170–7956 U.S.A. Subscriptions are available without charge to readers with mailing addresses in developing countries and in socialist economies in transition. Written request is required every two years to renew such subscriptions.


Manufactured in the United States of America
ISSN 0258-6770



This journal is indexed regularly in *Current Contents/Social & Behavioral Sciences*, *Index to International Statistics*, *Journal of Economic Literature*, *Public Affairs Information Service*, and *Social Sciences Citation Index*®. It is available in microform through University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

# THE WORLD BANK ECONOMIC REVIEW

Volume 8 May 1994 Number 2

# Measuring the Restrictiveness of Trade Policy

James E. Anderson and J. Peter Neary

*This article provides an introduction to the trade restrictiveness index (TRI), which equals the uniform tariff that is welfare equivalent to a given pattern of trade protection. Unlike standard measures of trade restrictiveness, the TRI has a solid theoretical basis, can incorporate both tariffs and quantitative restrictions, and can be adapted to construct the trade policy equivalent of domestic distortions. The article compares a number of applications and describes procedures for operationalizing the TRI on a personal computer. The authors conclude that the TRI has considerable potential in empirical work.*

The influence of trade policy on a country's economic well-being is one of the most widely debated topics in economics. Yet the question of how trade restrictiveness should be measured has received little attention in the past. In practice, restrictiveness is typically gauged using an ad hoc measure such as the trade-weighted average tariff, the coefficient of variation of tariffs, or the nontariff-barrier coverage ratio. But all these measures lack any theoretical foundation and are subject to theoretical and practical drawbacks. Some researchers, such as Papageorgiou, Michaely, and Chokski (1991), have attempted to construct subjective country-specific measures of trade restrictiveness. These have the advantage of incorporating important local considerations, but they are inherently difficult to replicate for countries or time periods other than those for which they were designed.

The problem of how trade restrictiveness should be measured is not severe in the world of textbooks, where trade barriers take a single and well-defined form. But in most real-world situations, especially in developing countries, actual systems of trade intervention are pervasive and highly complex. This poses a challenge for analysts and policymakers alike. In the face of a bewildering array of tariffs and quantitative restrictions, it can be extremely difficult to assess the true orientation of a country's trade policy or to evaluate the thrust of a package of policy changes that encourage trade in some product lines but

James E. Anderson is with the Department of Economics at Boston College, and J. Peter Neary is with the Department of Economics at University College Dublin, the London School of Economics, and the Centre for Economic Policy Research, London. This article reports on research that was supported by the World Bank and it was written while the second author was visiting the University of Ulster at Jordanstown. Comments from Will Martin and three anonymous referees are gratefully acknowledged.

discourage it in others. Traditional analysis provides little guidance on methods of aggregating restrictions across different markets. This makes it difficult to evaluate proposals for trade liberalization that form part of a stabilization package or to assess the progress made in moving toward less-restricted trade. A further reason for seeking a framework within which trade policies can be compared consistently is of analytical as well as practical importance. Because the case for free trade is ultimately a scientific hypothesis, theoretically sound but potentially false, some measure of trade restrictiveness is needed to test the impact of trade on growth and economic performance.[1]

This article describes a theoretically satisfactory yet practically implementable approach, developed by the authors, to measuring the restrictiveness of trade policy. Two relatively recent developments have made this approach possible. At a theoretical level, the normative theory of international trade has been formalized in a systematic way and extended to take account of trade policies other than tariffs.[2] At a practical level, the rapid increase in availability of cheap computing power has made it possible to implement models with a disaggregated structure that come closer than ever before to the complexity of real-world protective structures.

Section I provides a nontechnical introduction to the issues, discussing the conceptual problems in measuring trade restrictiveness, the drawbacks of commonly used measures, and the rationale for our proposed alternative measure, the trade restrictiveness index (TRI). Section II then sketches the analytical foundation of the TRI, and section III shows how it can be extended to include quantitative restrictions on trade, as well as tariffs. Subsequent sections review some applications of the TRI that have been carried out to date. Section IV discusses the application of the TRI to measuring the restrictiveness of the Multi-Fibre Arrangement, and section V shows how the TRI has been adapted to measure the trade restrictiveness of domestic policies, specifically, the reform of Mexican agricultural subsidies. Both of those studies are of a partial equilibrium kind; section VI describes a computable general equilibrium model that implements the TRI and that has been applied to an evaluation of Colombian trade policy reform. Section VII gives details of how the TRI can be calculated on a personal computer with only modest data requirements, and section VIII provides conclusions. More technical detail on the calculation of the TRI and on its application to the Multi-Fibre Arrangement is given in our companion article in this issue (Anderson and Neary 1994).

## I. Measuring Trade Restrictiveness in the Presence of Tariffs

What do we mean by a measure of "trade restrictiveness"? In principle, we mean some scalar index number that aggregates the trade restrictions in individ-

1. Leamer (1988) and Edwards (1992) propose and implement tests along these lines, adopting the Heckscher-Ohlin explanation of trade patterns as a maintained hypothesis. Krishna (1991) and Pritchett (1991) review this and other approaches to measuring openness and trade restrictiveness.

2. Dixit (1986) and Anderson (1988; 1994) provide overviews of recent work in the field.

ual markets. Whether a particular method of aggregation is satisfactory depends on the intended uses of the measure of restrictiveness. In a context of trade negotiations, for example, restrictiveness might be defined with reference to the volume of trade in restricted categories. Alternatively, trade restrictiveness might be defined in terms of the effect of domestic trade policies on the welfare of the country's trading partner; such an application is considered in section IV. In many applications, however, the appropriate index number is one that relates trade restrictions to their effect on domestic welfare.

The simplest context in which trade restrictiveness can be measured is when tariffs are the only form of trade policy. The left-hand side of figure 1 illustrates the market for a single good whose world price (assumed given) is $\pi_1^*$ and whose home import-demand curve is $m_1(\pi_1)$. Domestic producers and consumers face a price that is raised by the tariff to $\pi_1^0$. Adopting a partial equilibrium perspective for the moment, the resulting deadweight loss, or cost of protection, is measured by the Marshallian triangle *DCE*. The restrictiveness of trade policy in this one-good context can obviously and unambiguously be measured by the height of the tariff, given by the distance *AB*.

Matters are not so simple, however, when tariffs apply to more than one good. The right-hand side of figure 1 shows the import demand curve for good 2. Demand for good 2 is less elastic than demand for good 1, and good 2 is subject to a higher tariff. The total welfare loss from the two tariffs is the sum of the Marshallian triangles *DCE* and *IHJ*. But how should the "average" level of trade restrictiveness across these two markets be measured? The easiest approach, and the one typically adopted in practice, is to aggregate the two tariffs by weighting them by the imports (valued at border prices) of the two goods, *AC* and *FH*. (For ease of exposition, the world prices of the two goods are assumed to be equal to one.) This leads to the average tariff, denoted by $\bar{t}$. However, this approach immediately runs into difficulties. Consider a change in trade policy that leads to the situation illustrated in figure 2, where the import demand functions are the same as in figure 1 but the configuration of tariff levels is reversed. Now the correlation between demand elasticities and tariff levels is positive rather than negative. On the left, imports of the high-elasticity good 1 are almost eliminated, so its high tariff receives a very low weight in the average tariff. On the right, the low tariff on the low-elasticity good 2 receives a high weight. As a result, the calculated average tariff (again denoted by $\bar{t}$) is low—considerably lower than that in figure 1. Yet it seems intuitively obvious that trade is more restricted in figure 2 than in figure 1, since both welfare (measured by minus the sum of the deadweight loss triangles) and the volume of trade have fallen. The standard index has thus moved in the wrong direction after a change in trade policy.

Another measure of trade restrictiveness that is often used in conjunction with the trade-weighted average tariff is the coefficient of variation of tariffs. This measure is rationalized on the grounds that uniform tariffs minimize the welfare cost of a given constraint on the value of imports, although on other grounds

Figure 1. *Measuring Trade Restrictiveness in the Presence of Tariffs: Negatively Correlated Tariff Rates and Import Demand Elasticities*

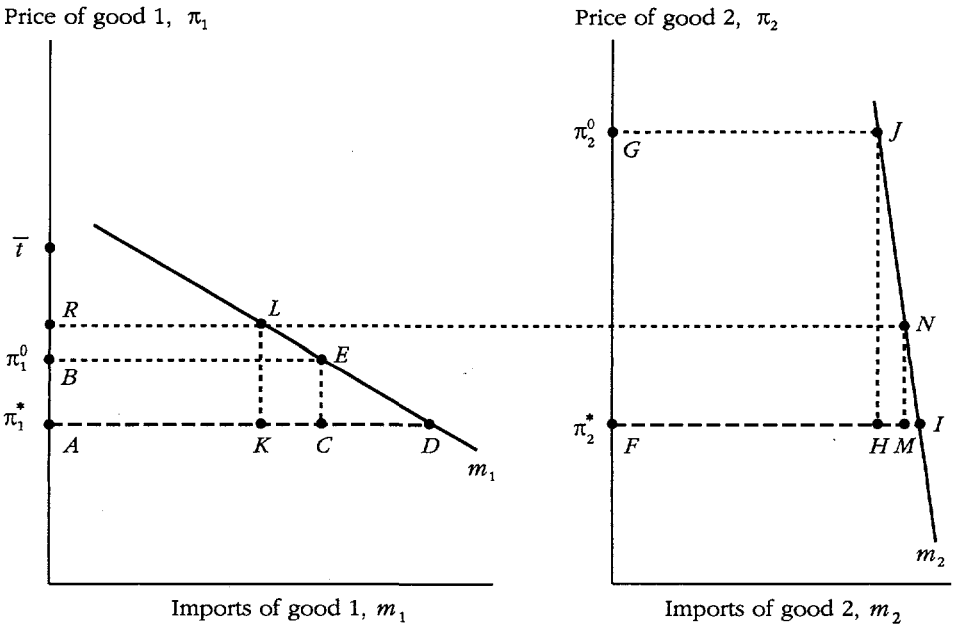
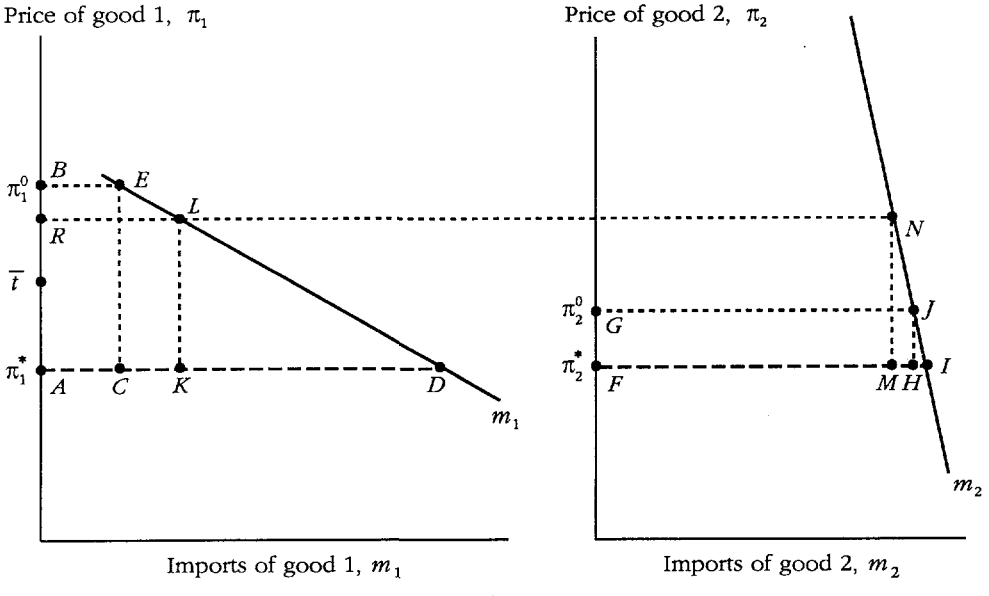



Figure 2. *Measuring Trade Restrictiveness in the Presence of Tariffs: Positively Correlated Tariff Rates and Import Demand Elasticities*



uniform tariffs are not necessarily desirable (see Anderson 1988; Stern 1990). The coefficient of variation is no more satisfactory a measure of trade restrictiveness than the trade-weighted average tariff. For reasonable parameter values, the coefficient of variation of tariffs may even be lower in figure 2 than in figure 1, suggesting once again that trade is less restricted, whereas intuitively this is not the case.

This example suggests that once we move away from the simple one-good case, purely statistical measures such as the trade-weighted average tariff or the coefficient of variation of tariffs bear no necessary relation to the welfare cost of trade policy. This in turn suggests that the distinction between the welfare cost of protection and the restrictiveness of the protective structure should not be maintained. Any satisfactory measure of trade restrictiveness must take account of the welfare costs imposed on the economy by the pattern of tariffs.

A more satisfactory approach to measuring trade restrictiveness is to find the uniform tariff that would be equivalent—in the sense of yielding the same welfare loss—to the actual tariffs applied.[3] In figure 1, this uniform tariff is equal to *AR*: by construction, the increase in the tariff on good 1 from *AB* to *AR* yields a welfare loss equal to the area *KCEL,* which is equal to the welfare gain of *HMNJ* arising from the reduction of the tariff on good 2. In figure 2 the uniform tariff *AR* now implies a lowering of the tariff on good 1 and an increase in that on good 2. The welfare equivalent uniform tariff is higher in figure 2 than in figure 1 because the welfare cost of restricting trade is higher in figure 2. In both cases, the welfare-equivalent uniform tariff is closer to the actual tariff on the high-elasticity good 1. This result accords with the intuition that tariffs on relatively elastic goods are more restrictive than tariffs on relatively inelastic goods.

## II. The Trade Restrictiveness Index

The concept of the welfare-equivalent uniform tariff can be extended to more general cases than the diagrammatic and partial equilibrium illustration in figures 1 and 2. In principle, the uniform tariff could be defined for any model of the economy, no matter how complex. In practice, our work to date has concentrated on models of perfectly competitive economies, although with very general specifications of technology and factor markets.

To generalize the welfare-equivalent uniform tariff, we make use of recent developments in the theory of trade policy, and especially of some technical tools introduced in Anderson and Neary (1992b; 1992c). Chief among these is the balance-of-trade function, which summarizes in implicit form the general equilibrium of a competitive multigood economy. This function, written as $B(\pi, u)$, is equal to the net transfer required to reach a given level of aggregate national

3. Corden (1966) is an early exploration of this approach.

welfare, denoted by $u$, for an economy with a given vector of domestic prices $\pi$.[4] Implicit in the function are all the variables that characterize the general equilibrium of the economy, including taste and technology parameters, exogenous foreign transfers, the level of world prices, and the price of the numeraire good. The requirement that the economy be in equilibrium is imposed by setting the value of the function equal to zero. Hence, if we wish to compare two situations, indexed by 0 and 1, respectively, the equilibrium conditions in each may be written compactly as

$$B(\pi^0, u^0) = B(\pi^1, u^1) = 0. \tag{1}$$

In some applications, the period 1 equilibrium may be identified with free trade, so that $\pi^1 = \pi^*$, but the techniques discussed here also allow for more general comparisons between two trade-distorted equilibria.

To motivate the derivation of the welfare-equivalent uniform tariff measure using the $B$ function, it is helpful to draw an analogy with the derivation of the true cost of living index for a consumer. This is typically defined as the expenditure needed to attain the utility level of period 0 facing the prices of period 1, $e(\pi^1, u^0)$, scaled by expenditure in the base period, $e(\pi^0, u^0)$:

$$\phi = e(\pi^1, u^0)/e(\pi^0, u^0). \tag{2}$$

Because the expenditure function is homogeneous of degree one in $\pi$, we can divide both sides of equation 2 by $\phi$ to rewrite it in a less conventional way:

$$\phi = [\phi\colon e(\pi^1/\phi, u^0) = e(\pi^0, u^0)]. \tag{3}$$

The interpretation is that the true cost of living index gives the uniform scaling factor by which period 1 prices must be deflated to compensate the consumer for the change in prices from $\pi^0$ to $\pi^1$.

Now, by analogy with equation 3,[5] the TRI is defined as the uniform scaling factor $\Delta$ by which period 1 prices must be deflated to compensate the aggregate consumer for the change in prices from $\pi^0$ to $\pi^1$:

$$\Delta = [\Delta\colon B(\pi^1/\Delta, u^0) = 0]. \tag{4}$$

The larger $\Delta$ is (for given period 0 prices), the more restrictive is the new tariff regime. In the case of a move to free trade ($\pi^1 = \pi^*$), $\Delta$ is less than one, and its inverse equals one plus the uniform tariff rate, which compensates for the abolition of period 0 tariffs. More generally, the inverse of $\Delta$ is the factor of proportionality by which period 1 prices must be multiplied in order to compensate for the change in tariffs. We call this the uniform tariff surcharge factor.

4. More precisely, the balance-of-trade function gives the excess of domestic expenditure, given by an expenditure function $e(\pi, u)$, over domestic income, or $B(\pi, u) = e(\pi, u) - g(\pi) - (\pi - \pi^*)'m - \beta$, where $g(\pi)$ is gross national product, $(\pi - \pi^*)'m$ is tariff revenue, imports $m$ equal $e_\pi(\pi, u) - g_\pi(\pi)$, and $\beta$ is any net transfer from abroad (which may include an exogenous balance of payments deficit).

5. Because of the presence of trade restrictions and the fact that there is an implicit numeraire good, the balance-of-trade function is not homogeneous of degree one in $\pi$, and so there is no step that is analogous to equation 2 in the general equilibrium derivation.

To illustrate further the intuition behind the TRI, consider the effect on $\Delta$ of a tariff change that causes a small change in period 1 prices, holding fixed the reference level of utility in period 0. Totally differentiating the expression on the right-hand side of equation 4 that implicitly defines $\Delta$ gives the proportional change in the TRI, $\hat{\Delta}$:

$$\hat{\Delta} = \Sigma\, \sigma_i \hat{\pi}_i. \tag{5}$$

Equation 5 gives the proportional change in the TRI as a weighted average of the proportional changes in domestic prices caused by the tariff changes. The weights depend on the partial derivatives of the balance-of-trade function with respect to prices, or the "marginal costs of tariffs," $B_i$:

$$\sigma_i = B_i \pi_i / \Sigma_j B_j \pi_j, \tag{6}$$

which are related to the slopes of the general equilibrium import demand functions.[6] Like the change in the trade-weighted average tariff, the change in the TRI is also a weighted average of domestic price changes. The difference is that the weights used, given by equation 6, are marginal welfare weights rather than actual trade shares, $m_i \pi_i / \Sigma_j m_j \pi_j$. This gives another perspective on the theoretical superiority of the TRI: not only does it derive from an explicitly specified model of the economy and so has a firm basis in welfare economics, but changes in it are measured by an aggregate of individual price changes using appropriate marginal welfare weights (as opposed to the ad hoc aggregation by actual trade shares of the trade-weighted average tariff).

Having set up the general theory of the TRI, empirical implementation requires a more precise specification of the model of the economy, which has so far been subsumed inside the black box of the $B$ function.[7] Before proceeding with this, however, we extend the theory of the TRI to include quantitative restrictions, which are a common method of protection in developing countries.

## III. Measuring Trade Restrictiveness with Quotas and Tariffs

The case in which trade is restricted only by quotas lends itself easily to the development of a scalar index of trade restrictiveness. A natural way of posing the problem in this case is, "What is the uniform proportionate change in quotas that would compensate in welfare terms for a given change in quotas?" This

6. From footnote 4, the typical price derivative $B_i$ equals $-\Sigma_j(\pi_j - \pi_j^*)\partial m_j / \partial \pi_i$.

7. For example, in the special case of linear demands illustrated in figures 1 and 2, it is easier to work directly with the (approximate) welfare function given by the sum of the Marshallian triangles rather than with the $B$ function. For any domestic price vector $\pi$, the welfare cost is $u = -\Sigma\, \gamma_i(\pi_i - \pi_i^*)^2$, where $\gamma_i$ is the slope of the import demand curve for good $i$. Hence, the welfare equivalent ad valorem tariff (equal to $\Delta^{-1} - 1$, when $\pi^1$ equals $\pi^*$) is given by the square root of $\Sigma\, \gamma_i(\pi_i - \pi_i^*)^2 / \Sigma\, \gamma_i(\pi_i^*)^2$. Anderson (1992) shows that, in general, the TRI with tariffs only can be written as a function of a weighted average tariff and the generalized variance of tariffs. If all goods are substitutes, the weights are non-negative. If the trade expenditure function is Cobb-Douglas, the weights reduce to trade weights, and the generalized variance collapses to the trade-weighted variance.

question leads to an index that is defined over quantities rather than over prices.[8] The technical development makes use of another function, the distorted balance-of-trade function, which is the analogue of the balance-of-trade function presented in the last section, modified to take account of quota distortions. (See Anderson and Neary 1992b and 1992c for details.) The distorted balance-of-trade function, denoted $B^d(q, u)$, is defined over the permitted import levels of the quota-constrained goods $q$ and the level of utility $u$. As before, a great deal is hidden inside the black box, including the world prices $p^*$ of the quota-constrained goods. The quantity-based TRI for quotas can now be defined as the proportionate change in period 1 quotas required to reach period 0 utility:

$$\Delta^q(q^1, u^0) \equiv [\Delta^q: B^d(\Delta^q q^1, u^0) = 0]. \tag{7}$$

For the case of two goods, this index is illustrated in figure 3, drawn in quota space. Point $A$, with coordinates $(q_1^0, q_2^0)$, represents the base-period equilibrium, and point $D$, with coordinates $(q_1^1, q_2^1)$, represents the period 1 equilibrium (which may be, but need not be, identified with free trade). The curve through $A$ is an iso-utility locus, and it intersects the ray $OD$ from the origin to $D$ at point $E$. The value of $\Delta^q$ is the distance $OE/OD$. As in the case of tariffs, the larger is $\Delta^q$ (for given period 0 quota levels), the more restrictive is the period 1 quota regime.

Finally, we must consider the realistic case in which trade is restricted by both tariffs and quotas. Two alternative approaches are now possible, differing in their intuitive appeal and in their data requirements. The first approach is simply to combine the individual indexes already developed for the cases of tariffs and quotas alone, leading to a mixed quantity- and price-based index:

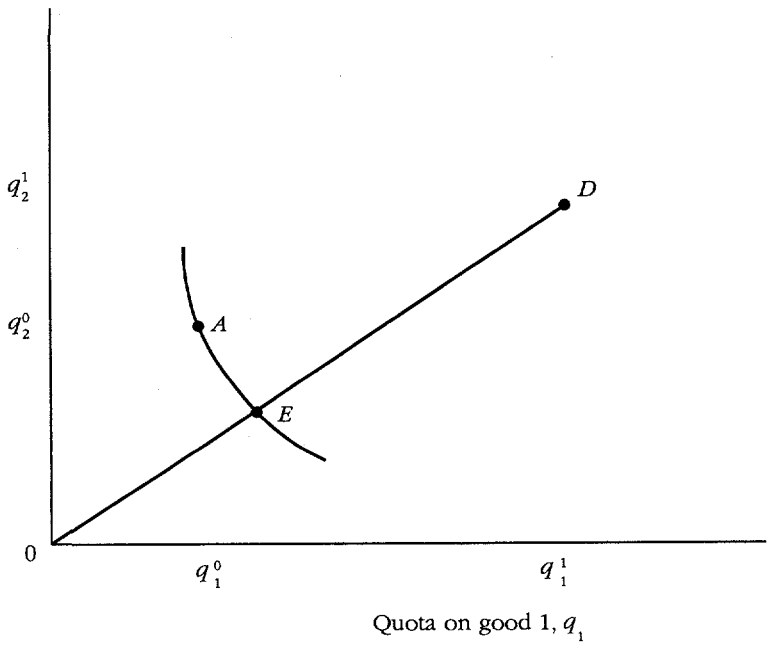
$$\Delta^\lambda(q^1, \pi^1, u^0) \equiv [\Delta^\lambda: B^d(\Delta^\lambda q^1, \pi^1/\Delta^\lambda, u^0) = 0]. \tag{8}$$

The value of $\Delta^\lambda$ is the equal proportionate relaxation of all quota levels and reduction of all tariff-inclusive prices that would be equivalent in welfare terms to a given initial protective structure with an arbitrary pattern of quotas and tariffs. As before, a rise in $\Delta^\lambda$ corresponds to an increase in trade restrictiveness.

The great advantage of the hybrid index is computational: although the level of this index depends on world prices for quota-constrained goods, $p^*$, changes in the index between two distorted situations can be computed without knowledge of these prices (data on which are notoriously difficult to obtain). However, the index has the disadvantage of combining changes in quantities for quota-constrained goods with changes in prices for tariff-constrained goods. This is not a meaningless mixture because the value of the index is a pure number and because changes in the index are a weighted average of changes in the tariffs-only and quotas-only indexes, with the weights measuring the relative contribu-

8. In this form, the TRI is seen to descend from a family of "distance function" measures developed, among others, by Debreu (1951) and Deaton (1979). In our early work, which for the most part considered quota distortions only (Anderson and Neary 1990; Anderson 1991), we called our index the "coefficient of trade utilization," echoing Debreu's "coefficient of resource utilization."

Figure 3. *The Trade Restrictiveness Index with Quotas*




tion of tariffs and quotas to changes in welfare. Nonetheless, the hybrid index is difficult to interpret if we wish to compare the index across countries or time periods in which the mix of goods that are subject to tariffs and quotas differs.

It is desirable, therefore, to develop a second approach, leading to an index based on prices for both categories of goods. In the case of quota-constrained goods, this involves using the tariff equivalents of the quotas. The resulting index is a uniform tariff and tariff-equivalent surcharge factor. The index is the uniform proportionate change in the actual domestic prices, $\pi$, for tariff-constrained goods and the virtual prices, $p^d$, (that is, world prices plus tariff equivalents) for quota-constrained goods that would compensate in welfare terms for the actual change in policy instruments from $(q^0, \pi^0)$ to $(q^1, \pi^1)$:[9]

$$\Delta(q^1, \pi^1, u^0) \equiv [\Delta\colon B(p^d/\Delta, \pi^1/\Delta, u^0) = 0]. \tag{9}$$

In any application, the choice between this index and the hybrid index will depend on the quality of data available and the type of comparative exercise being undertaken.

9. The term "virtual prices" derives from the theory of rationing (Neary and Roberts 1980). For theoretical consistency, the virtual prices, like the value of $\Delta$ itself, must be evaluated at the new instruments but the old level of welfare. See Anderson and Neary (1992b) for details.

## IV. Partial Equilibrium Applications of the TRI

The theoretical approach outlined in sections II and III provides a framework for computing the TRI in a wide variety of applications. To operationalize the approach, it is necessary to have a computable model of the economy under consideration. This raises a whole set of choices: the model may be partial or general equilibrium; it may be linearized around the initial equilibrium or be explicitly nonlinear; and it may be more or less disaggregated at the commodity level. In principle, the TRI can be computed for a model that adopts any combination of choices from this menu. But in practice, its focus on trade policy instruments suggests choosing a highly disaggregated model to capture the fine detail of actual protective policies. This in turn suggests implementing the TRI in either a partial equilibrium model or a general equilibrium model with a tightly specified production structure.

The first set of applications of the TRI adopted a partial equilibrium perspective with quantitative restrictions only, taking the case of imports of textiles and apparel to the United States under the Multi-Fibre Arrangement (MFA).[10] A pilot study (Anderson and Neary 1992b) considered exports from Hong Kong only. This has been extended to exports from six other countries: Bangladesh, India, Indonesia, the Republic of Korea, Mexico, and Thailand (Anderson and Neary 1994, in this issue). In each case, the year-to-year changes in the quantity-based index were estimated. Such changes are weighted averages of the changes in the quotas, where (as in the case of tariffs in equation 6) the weights depend on the quota derivatives of the constrained balance-of-trade function. The partial equilibrium context permits calculation of changes in the TRI using only readily available information on import demand elasticities. It is also straightforward to calculate separately the changes in the TRI from the perspectives of the exporting and the importing countries (the differences reflecting differences in market power and in the shares of quota rents that accrue to each country).

A key issue to be addressed in any empirical study of quantitative trade restrictions is the destination of the resulting rents. If detailed information on the mechanism whereby rents are shared between importing and exporting countries is available, it may be incorporated into the formulas for the marginal costs of tariffs and the shadow prices of quotas that are needed to calculate the change in the TRI, using the general expressions of Anderson and Neary (1992c). Typically, however, such information is not available. (Indeed, even data on export license prices are hard to come by. In the MFA study we were fortunate to have access to estimates of Hong Kong export license prices made by Carl Hamilton.) In the MFA study we assumed that all rents accrued to the exporting country, with one important exception: when the importing country imposes a tariff on a quota-constrained import, the tariff revenue is a transfer of part of the quota rents to the importing country. (Details are given in Anderson and Neary 1994, in this issue.)

10. U.S. imports of cheese have also been considered by Anderson (1991).

Table 1 presents some representative results from the MFA study. These are from the perspective of the importing country (the United States) and refer to imports from Hong Kong only. The first column gives the yearly changes in the TRI. Next, we must take account of a feature peculiar to quotas: unless quotas grow at the rate of growth of domestic excess demand for imports, the severity of the quotas increases. Hence, changes in the restrictiveness of quota policy should not be evaluated in relation to a constant quota policy but rather in relation to a neutral quota policy that would allow all permitted import levels to increase at the rate of growth of excess demand for the quota-constrained goods. Under reasonable assumptions, the growth in excess demand may be approximated by the economy's rate of growth (shown in the second column).[11] This consideration gives rise to a compensated TRI, changes in which are shown in the third column. These compensated TRI changes equal the sum of the uncompensated changes in the first column and the changes in U.S. real disposable income in the second. (Real income growth is added because it increases the restrictiveness of a given set of quotas.) The final column gives the change in the trade-weighted average tariff equivalent, using U.S. import shares as weights. It is clear that changes in this measure bear little relation to those in the TRI. Although some of the assumptions made in calculating changes in the TRI are open to question, this case study clearly demonstrates that using the TRI to evaluate changes in trade restrictiveness yields substantially different results from those obtained using the trade-weighted average tariff equivalent. The superior theoretical properties of the TRI imply that the increased restrictiveness of policy that it reveals, at least when real income growth is taken into account, is a more plausible indicator of the change in trade policy over the period than is the reduction in restrictiveness suggested by the cumulative fall in the average tariff equivalent.

Table 1. *Changes in the Trade Restrictiveness Index (TRI): U.S. Imports of Textiles and Apparel from Hong Kong, 1982–88*
(percentage change)

| *Year* | *Change in TRI* | *Change in real income* | *Change in compensated TRI* | *Change in average tariff equivalent* |
|---|---|---|---|---|
| 1983 | −2.8 | 3.9 | 1.1 | 84.4 |
| 1984 | 4.2 | 6.8 | 11.0 | −8.1 |
| 1985 | −1.7 | 3.2 | 1.5 | −39.2 |
| 1986 | −6.6 | 2.8 | −3.8 | 42.2 |
| 1987 | −1.0 | 2.9 | 1.9 | 12.0 |
| 1988 | −0.9 | 4.5 | 3.6 | −53.0 |
| Cumulative | −8.8 | 26.6 | 15.7 | −22.9 |

*Source:* Anderson and Neary (1994).

11. For a rigorous justification of this procedure, see Anderson and Neary (1992b).

## V. Measuring the Trade Restrictiveness of Domestic Policies

A very different application of the TRI is to the evaluation of the trade restrictiveness of domestic price policies. Such policies distort trade just as much as explicitly trade-focused policies do, a fact that is increasingly recognized in trade negotiations. Domestic price policies have featured prominently, for example, in farm subsidy negotiations in the Uruguay Round and in the North American Free Trade Agreement. Attempts have also been made to quantify their impact on trade (see OECD 1991). But the measures used to do this (known as producer and consumer subsidy equivalent indexes, or PSEs and CSEs) are just as crude as trade-weighted average tariff measures and are subject to the same drawbacks. For example, CSEs are calculated by weighting different subsidy rates by the volume of domestic consumption of each good. This gives a high weight to goods with large consumption and a low price-elasticity of demand, even though a subsidy on such goods has a small effect on either trade volume or welfare.

The theoretical refinements to the TRI required to incorporate domestic policies are complicated in detail but straightforward in principle.[12] Assume that the distortions occur in the markets for traded goods. If $p$ and $q$ represent the domestic producer and consumer prices, respectively, we can once again write the balance of trade as a function of these prices and of the level of utility, $B(p, q, u)$. Now, in comparing two equilibria, the TRI is again defined as the uniform scaling factor that, when applied to both consumer and producer prices, would compensate for a policy change. Formally,

$$\Delta(p^1, q^1, u^0) \equiv [\Delta: B(p^1/\Delta, q^1/\Delta, u^0) = 0]. \tag{10}$$

Once again, when the period 1 equilibrium corresponds to free trade ($p^1 = q^1 = p^*$), the TRI equals the inverse of one plus the uniform tariff, which would have the same welfare effect as the base-period producer and consumer distortions.

The theoretical measure defined in equation 10 was used in Anderson and Bannister (1992) to measure the trade restrictiveness of changes in Mexican agricultural policy between 1985 and 1989 in a partial equilibrium context. This period was one of rapid policy change, with increases in some subsidies and reductions in others. The first row of table 2 shows the yearly changes in the TRI. The index shows large increases in restrictiveness in 1986 and especially 1987, followed by major reductions in restrictiveness in 1988 and 1989. The cumulative effect of these changes is a 40.9 percent fall in trade restrictiveness during the four-year period. These changes may be decomposed into changes in the producer and consumer subsidy components of the TRI, and this decomposition in turn may be compared with the conventional PSE and CSE measures. The comparisons are given in the remaining rows of the table, where $\Delta^p$ and $\Delta^q$ denote the "true" subsidy equivalent indexes for producers and consumers, re-

12. See Anderson and Neary (1992a) for details, where the approach is also extended to distortions in the markets for both factors of production and nontraded goods.

Table 2. *Changes in the Trade Restrictiveness Index (TRI) and Its Components for the Mexican Agricultural Sector, 1985–89*
(percentage change)

| *Index* | *1986* | *1987* | *1988* | *1989* | *Cumulative* |
|---|---|---|---|---|---|
| Changes in TRI | 7.5 | 40.2 | −40.3 | −34.3 | −40.9 |
| Subsidy equivalent indexes | | | | | |
| Ad hoc producer subsidy, PSE | −7.4 | 2.4 | −4.9 | −5.9 | −15.1 |
| True producer subsidy, $\Delta^p$ | 7.1 | 34.4 | −31.6 | −30.1 | −31.2 |
| Ad hoc consumer subsidy, CSE | −6.4 | 15.3 | 32.5 | −31.0 | −1.3 |
| True consumer subsidy, $\Delta^q$ | −79.8 | −11.4 | 69.7 | −7.9 | −72.0 |

*Source:* Anderson and Bannister (1992).

spectively, and PSE and CSE denote the ad hoc (average share-weighted) indexes.[13] As in table 1, there is little or no concordance between changes in the theoretically based measures and the ad hoc measures. Moreover, there is no acceptable procedure for combining the PSE and CSE to form an aggregate index, whereas this is precisely what the TRI is designed to do. Once again, the TRI seems to be a much more satisfactory method of evaluating the effects of policies on international trade.

## VI. General Equilibrium Modeling of the TRI

Both of the applications discussed above have been partial equilibrium in character. Although based on a theoretically consistent framework, this procedure ignores interactions between the markets considered and the rest of the economy that are transmitted through factor markets and markets for nontraded goods. Ignoring these interactions is unlikely to present a problem if trade policy changes in a single sector or in a small group of sectors is being considered, but it is obviously unsatisfactory in the case of a wide-ranging change in trade policy. At the same time, the focus of the TRI on the fine detail of the structure of protection makes it difficult to combine with most existing computable general equilibrium (CGE) models, which tend to be highly aggregated. For example, a typical CGE model distinguishes about 20 or 30 sectors, whereas the application discussed below accommodates more than 2,000 different traded commodities. The price of this disaggregation is the need to restrict significantly the structure of intercommodity and interfactor substitution.

With these considerations in mind, and to implement the TRI in a general equilibrium context, a new CGE model was developed that uses exact functional

13. For example, the "true" producer subsidy equivalent index is defined, by analogy with equation 10, as:

$$\Delta^p(p^1, q^1, u^0) \equiv [\Delta^p: B(p^1/\Delta^p, q^1, u^0) = 0].$$

forms to calculate global changes.[14] The disadvantage of this model is that the effects of any misspecification are likely to be magnified for large changes in trade policy. The advantages are greater theoretical consistency and the ability to calculate explicitly the level of the TRI and changes in it.

The model, discussed more fully in Anderson (1993), makes a number of key assumptions. First, following Armington (1969), every traded good produced at home is assumed to be an imperfect substitute for an imported good, and domestic consumption and production are characterized by functions of the constant elasticity of substitution type.[15] By making these assumptions about the functional forms describing producer and consumer behavior, the values of only a relatively small number of key parameters need be imposed, the remainder being inferred from cross-equation constraints. Second, following Jones (1974), a single, composite, nontraded good is assumed, that good being the only one both produced and consumed at home. Thus, no exports are domestically consumed, and no imports are domestically produced. Eliminating many of the interactions between the consumption and production sides of the economy greatly simplifies the model and allows the relative price of the nontraded good to be interpreted as the real exchange rate. Third, following Jones (1971), a specific-factors structure is assumed for the domestic market: each sector uses a specific type of capital, but all draw on a pool of intersectorally mobile labor. Finally, an assumption is made concerning the destination of the rents that arise from quantitative restrictions on imports. In typical applications (unlike the MFA study discussed in section IV above) data on quota premiums are unlikely to be available. Following Krueger (1974), therefore, the convenient simplification that all the quota rents are dissipated through competitive rent-seeking is made.

Notwithstanding these assumptions, the model is in other respects very general. As noted, it can accommodate detailed information on the trade restrictions affecting a very large number of commodities, whether they are subject to tariff or quota constraints.[16] Moreover, intermediate inputs are explicitly distinguished from final goods, and they too may be subject to both tariff and quota restrictions. A further advantage of the TRI approach is that it permits the consistent aggregation of trade restrictions on final and intermediate goods in a

14. In earlier work, we experimented with a different approach, working directly with the expressions for changes in the TRI and, specifically, with the differentials of equation 8. This procedure is equivalent to calculating the TRI by linearizing the model of the economy embodied in the $B$ function around the initial equilibrium.

15. More specifically, a constant elasticity of substitution (CES) expenditure function has been assumed for consumption; a CES cost function defined over the primary factor, labor, and intermediate inputs (both tariff-constrained and quota-constrained) has been assumed for aggregate output; and a constant elasticity of transformation production frontier has been assumed to determine the allocation of aggregate output between exports and nontraded goods. An advantage of the CES form is that it yields closed-form solutions for the virtual prices.

16. If a good is subject to both a quota and a tariff, and if the quota is binding, then the tariff is nonbinding at the margin. In this case, as already noted in section IV, the tariff serves merely to ensure that the fraction of total quota rents made up of tariff revenue is retained by the importing country.

much more satisfactory way than the traditional approach of using the effective rate of protection.

This model is used in Anderson (1993) to estimate the effects of trade reform in Colombia between 1989 and 1990. Table 3 gives the changes in Colombian trade policy between those two years, using some standard measures of trade restrictiveness. These measures show a confusing pattern. Average tariffs on all goods fell, although it should be recalled that a reduction in tariffs on quota-constrained goods reduces welfare because it lowers the share of quota rents that accrue to the domestic economy. Accompanying this was a greater dispersion of tariffs, as measured by the coefficient of variation, a significant reduction in the coverage of nontariff barriers, and an increase in the (unweighted) average level of quotas. In any case, because the composition of quota- and tariff-constrained categories changes between the two years, the standard indexes are not truly comparable. Clearly, assessing the overall thrust of the trade policy changes is extraordinarily difficult without a consistent framework for aggregating these changes.

The TRI calculations in table 4 attempt to provide just such an assessment, solving the problem of comparability by aggregating in a manner that is fully compatible with the underlying theory. The table gives changes in the TRI (compensated for real income growth) under different combinations of assumptions about the values of three key elasticities: the elasticity of final consumption demand, the elasticity of demand for intermediate inputs, and the elasticity of transformation.[17] The first row of the table shows a change in the TRI of −4.9

Table 3. *Indicators of Trade Reform in Colombia, 1989–90*

| *Item* | *1989* | *1990* |
|---|---|---|
| *Average tariff (index)* | | |
| Final goods | 0.230 | 0.166 |
| Tariff-constrained | 0.092 | 0.100 |
| Quota-constrained | 0.237 | 0.208 |
| Intermediate goods | 0.185 | 0.151 |
| Tariff-constrained | 0.125 | 0.125 |
| Quota-constrained | 0.195 | 0.170 |
| *Coefficient of variation of tariffs* | | |
| Final goods | 0.769 | 0.813 |
| Intermediate goods | 0.596 | 0.625 |
| *Nontariff barrier coverage ratio (percent)* | | |
| Final goods | 90.8 | 57.3 |
| Intermediate goods | 76.3 | 34.7 |
| *Average (unweighted) change in quotas (percent)* | | |
| Final goods | n.a. | 12.0 |
| Intermediate goods | n.a. | 12.0 |

n.a. Not applicable.
*Source:* Anderson (1993).

17. Because of data constraints, the version of the TRI that is implemented is the hybrid index given in equation 8.

Table 4. *Sensitivity Analysis of the Change in the Trade Restrictiveness Index (TRI) for Colombia, 1989–90*

| *Elasticity of final demand* | *Elasticity of intermediate demand* | *Elasticity of transformation* | *Change in compensated TRI (percent)* |
|---|---|---|---|
| 1.5 | 1.0 | 1.5 | −4.9 |
| 2.0 | 0.7 | 2.0 | −4.8 |
| 2.0 | 0.7 | 5.0 | −4.2 |
| 5.0 | 0.5 | 5.0 | −4.4 |
| 5.0 | 0.7 | 5.0 | −3.7 |

*Source:* Anderson (1993).

percent, implying that the overall effect of the policy changes was a modest liberalization, equivalent in welfare terms to a uniform cut in tariffs and a relaxation of quotas by 4.9 percent.[18] The remaining rows show that this conclusion is relatively robust to changes in the assumed values of the elasticities. This robustness has also been found in other applications of the TRI that have been carried out to date. Because it is only an empirical finding, of course, it needs to be replicated extensively on other data sets before it can be regarded as typical.

## VII. Implementing the TRI on a Personal Computer

One additional feature of the CGE model described in the last section is that all the calculations were carried out on a personal computer in a manner that can easily be replicated. The appendix to Anderson (1993) gives a more complete description of how the CGE model has been implemented on an EXCEL spreadsheet. For each commodity, the user needs to enter data on the domestic price, the tariff rate, and the volume of imports, as well as two codes, one indicating whether the commodity is for final or intermediate use and the other indicating whether it is subject to a binding quantitative restriction. This information must be provided for two time periods. The spreadsheet program then calculates the change in the TRI between the two periods. In addition, standard measures such as the trade-weighted average tariff and the nontariff barrier coverage ratio are calculated for comparison. Of course, the second of the two periods for which data are supplied may be a hypothetical one. For example, if estimates are available of world prices and of the free-trade import levels of goods that are currently quota-constrained, the second period could be one in which all trade restrictions have been abolished. In that case, the program calculates the level of the TRI in the initial period.

As far as the underlying model of the economy is concerned, the program specifies default values of the key substitution parameters in production and

18. The uncompensated change in the TRI for these parameter values is −7.8 percent, which indicates that compensating for Colombia's real income growth rate of 3.5 percent significantly reduces the estimated trade liberalization.

consumption. These may be altered by the user, thus permitting an exploration of the sensitivity of the TRI estimates to changes in the underlying parameters. (Table 4 was produced in this way.) The computer program permits estimation of the degree of trade restrictiveness of a given trade policy in a consistent framework. It also makes possible other applications of the approach. For example, the program can easily be adapted to calculate the welfare cost of a given change in trade policy or the welfare cost of an equiproportionate tariff change sufficient to raise a given amount of revenue. Because the program fits on a single 720 kilobyte disk and can be used on a portable computer, it permits an easy assessment of trade restrictiveness with minimal data and computing requirements.

## VIII. Conclusion

In this article, we have outlined a new approach to measuring the restrictiveness of trade policy. Our starting point was the observation that the appropriate method of aggregating individual trade restrictions to construct a single scalar index depended on the uses to which the index would be put. Because many applications of measures of trade restrictiveness deal with the impact of trade restrictions on domestic welfare, we argued that a satisfactory measure of trade restrictiveness must be closely related to standard measures of the welfare cost of protection. This led to our approach, which is a generalization of the welfare-equivalent uniform tariff, that is, the uniform tariff that would have the same welfare cost as a given nonuniform tariff structure. This approach is firmly based in economic theory and so avoids the ad hoc nature of the measures typically used in practice, such as the trade-weighted average tariff or the non-tariff-barrier coverage ratio. Of course, to implement our measure, it is necessary to assume a particular model of the economy: all our applications to date have assumed that the economy is perfectly competitive. Implementing our approach requires more data than traditional measures of trade restrictiveness, but we have developed some empirical procedures the data requirements of which are likely to be met in most developing countries.

Beyond the specific models we have developed, the TRI perspective draws attention to a number of key general issues that should be borne in mind in any empirical study of trade policy. One is that simple averages of tariff rates are unlikely to be helpful guides to the true extent of trade restrictiveness. Another is that the destination of the rents that arise from quantitative restrictions is an important determinant of their welfare impact and their restrictiveness. A third is that the restrictiveness of quotas depends crucially on the environment in which they apply and hence on the values of the exogenous variables that determine the economy's equilibrium. Because the methods described in this article attempt to deal with these issues in a consistent framework, we claim that these methods, however crude, represent a significant advance over any others that are available. Moreover, because of recent developments in computer technology, our methods can readily be implemented for practical problems.

## References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Anderson, James E. 1988. *The Relative Inefficiency of Quotas.* Cambridge, Mass.: MIT Press.

———. 1991. "The Coefficient of Trade Utilisation: The Cheese Case." In Robert E. Baldwin, ed., *Empirical Studies of Commercial Policy.* University of Chicago Press.

———. 1992. "Tariff Index Theory." WPS 1023. World Bank, International Economics Department, Washington, D.C. Processed.

———. 1993. "Measuring Trade Restrictiveness in a Simple CGE Model, with Appendix: A Manual for Using the TRI Spreadsheet Model." Boston College, Department of Economics, Boston. Processed.

———. 1994. "The Theory of Protection." In David Greenaway and L. Alan Winters, eds., *Surveys in International Trade.* Oxford: Basil Blackwell.

Anderson, James E., and Geoffrey Bannister. 1992. "The Trade Restrictiveness Index: An Application to Mexican Agriculture." WPS 874. World Bank, International Economics Department, Washington, D.C. Processed.

Anderson, James E., and J. Peter Neary. 1990. "The Coefficient of Trade Utilization: Back to the Baldwin Envelope." In Ronald W. Jones and Anne O. Krueger, eds., *The Political Economy of International Trade: Essays in Honor of Robert E. Baldwin.* Oxford: Basil Blackwell.

———. 1992a. "Domestic Distortions and International Trade." Boston College, Department of Economics, Boston. Processed.

———. 1992b. "A New Approach to Evaluating Trade Policy." WPS 1022. World Bank, International Economics Department, Washington, D.C. Processed.

———. 1992c. "Trade Reform with Quotas, Partial Rent Retention and Tariffs." *Econometrica* 60(1):57–76.

———. 1994. "The Trade Restrictiveness of the Multi-Fibre Arrangement." *The World Bank Economic Review* 8(2):171–89.

Armington, Paul S. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *International Monetary Fund Staff Papers* 16(1):159–78.

Corden, W. Max. 1966. "The Effective Protective Rate, the Uniform Tariff Equivalent and the Average Tariff." *Economic Record* 42:200–16.

Deaton, Angus S. 1979. "The Distance Function in Consumer Behaviour with Applications to Index Numbers and Optimal Taxation." *Review of Economic Studies* 46(3):391–405.

Debreu, Gerard. 1951. "The Coefficient of Resource Utilization." *Econometrica* 19(2):273–92.

Dixit, Avinash K. 1986. "Tax Policy in Open Economies." In Alan Auerbach and Martin Feldstein, eds., *Handbook of Public Economics.* Amsterdam: North-Holland.

Edwards, Sebastian. 1992. "Trade Orientation, Distortions and Growth in Developing Countries." *Journal of Development Economics* 39(1):31–57.

Jones, Ronald W. 1971. "A Three-Factor Model in Theory, Trade and History." In Jagdish N. Bhagwati, Ronald W. Jones, Robert Mundell, and Jaroslav Vanek, eds., *Trade, Balance of Payments and Growth: Papers in International Economics in Honor of Charles P. Kindleberger.* Amsterdam: North-Holland.

———. 1974. "Trade with Non-traded Goods: The Anatomy of Interconnected Markets." *Economica* 41(162):121–38.

Krishna, Kala. 1991. "Openness: A Conceptual Approach." Harvard University, Department of Economics, Cambridge, Mass. Processed.

Krueger, Anne O. 1974. "The Political Economy of the Rent-Seeking Society." *American Economic Review* 64(3):291–303.

Leamer, E. E. 1988. "Measures of Openness." In Robert E. Baldwin, ed., *Trade Policy Issues and Empirical Analysis*. University of Chicago Press.

Neary, J. Peter, and K. W. S. Roberts. 1980. "The Theory of Household Behaviour under Rationing." *European Economic Review* 13(1):25–42.

OECD (Organization for Economic Cooperation and Development). 1991. *Producer Subsidy Equivalent and Consumer Subsidy Equivalent Tables, 1979 to 1990*. Paris.

Papageorgiou, Demetrios, Michael Michaely, and Armeane M. Chokski, eds., 1991. *Liberalizing Foreign Trade*. 7 vols. Oxford: Basil Blackwell.

Pritchett, Lant. 1991. "Measuring Outward Orientation in Developing Countries: Can It Be Done?" WPS 566. World Bank, Country Economics Department, Washington, D.C. Processed.

Stern, Nicholas H. 1990. "Uniformity Versus Selectivity in Indirect Taxation." *Economics and Politics* 2(1):83–108.

# The Trade Restrictiveness of the Multi-Fibre Arrangement

James E. Anderson and J. Peter Neary

*This study uses the trade restrictiveness index (TRI) of Anderson and Neary (1990) to evaluate U.S. policy toward seven major exporters of textiles and apparel under the Multi-Fibre Arrangement (MFA). The period covered is 1982–88. The MFA controls the shipment of most textile and apparel items to the United States through a system of bilaterally negotiated export quotas that allow for annual growth. The arrangement itself was renegotiated in 1986, with an expansion in the number of items covered and countries included. In addition to these policy changes, changes in economic conditions during the data period altered the restrictiveness of the MFA. The TRI in principle permits all these influences to be accounted for in a consistent manner. The TRI results are contrasted with the standard trade-weighted average tariff equivalent of the quotas. The correlation of the two measures is not significantly different from zero.*

Measuring the restrictiveness of trade policy in a way that allows international and intertemporal comparability is important for trade negotiations and other comparative policy evaluation purposes. In the absence of a theory of index numbers for trade restrictions, analysts have used ad hoc indexes of restrictiveness such as the trade-weighted average tariff and the coefficient of variation of tariffs, despite their apparent flaws. For quota restrictions, the trade-weighted average tariff equivalent of the quota is standard. In Hamilton and Kim (1990), for example, this measure is used along with five others in an attempt to evaluate the restrictiveness of the Multi-Fibre Arrangement (MFA).

In our companion article in this issue (Anderson and Neary 1994), we review our recent work that attempts to resolve these problems by developing a scalar measure of trade restrictiveness consistent with economic theory and yet feasible to implement in practice. This trade restrictiveness index (TRI) is used here to evaluate U.S. policy toward seven major exporters of textiles and apparel under

James E. Anderson is in the Department of Economics at Boston College, and J. Peter Neary is in the Department of Economics at University College Dublin, the London School of Economics, and the Centre for Economic Policy Research, London. The research for this study was supported by the Research Committee of the World Bank and was carried out while the second author was visiting the University of Ulster at Jordanstown. The authors are grateful for able assistance from Christopher Holmes and Ulrich Reincke; for helpful discussions with Refik Erzan, Kala Krishna, and Will Martin; and for the comments of three anonymous referees.

the MFA from 1982 to 1988. The MFA controls the shipment of most textile and apparel items to the United States through a system of bilaterally negotiated export quotas, or voluntary export restraints (VERs). There is an allowance for annual growth in the quotas, and the arrangement itself was renegotiated in 1986 (MFA-III was replaced by MFA-IV), with an expansion in the number of items covered and countries included. In addition to these policy changes, economic conditions changed over the data period in both the United States and the exporting countries, as well as the rest of the world. All these influences altered the restrictiveness of the MFA. The TRI in principle permits all these influences to be accounted for in a consistent manner.

The trade restrictiveness of a VER may be viewed from either the importer's or the exporter's point of view, and both are considered here. In each MFA product category, we assume that all exporters produce perfect substitutes. Nevertheless, each exporter nation has monopoly power because of the restrictions on other suppliers to the VER-constrained market. Specifically, a license to sell an additional unit adds to total rent by the amount of the quota premium but depresses the premium and causes a loss on inframarginal quota licenses. Hong Kong is the main exporter of textile and apparel products to the United States, and for a low assumed elasticity of U.S. demand, it has sufficient market power in trade with the United States that a marginal fall in trade restrictiveness helps the United States while hurting Hong Kong.[1] This is related to the finding of Trela and Whalley (1990) that a reversion to free trade would hurt Hong Kong. For higher elasticities of U.S. demand, Hong Kong's monopoly power is less, and a fall in trade restrictiveness helps both nations. This issue does not arise for the remaining exporters, which have market shares too small to create a large inverse elasticity, but in all cases the measure of trade restrictiveness is quantitatively altered when the point of view shifts from exporter to importer because the weights in the index are altered.

Section I reviews the theory of the TRI and outlines how changes in it are calculated from the perspective of importing and exporting countries. Section II discusses the data used. Section III presents and evaluates the results, and section IV offers conclusions. The appendix details the calculation of the shadow prices of quotas, the key step in operationalizing changes in the TRI.

## I. The Theory of the TRI

If a single quota changes, an unambiguous measure of the change in trade restrictiveness is, trivially, the change in the quota. An alternative measure is the change in the domestic price of the quota-constrained good, equal to the change in the quota premium (the "tariff equivalent") if the foreign price is constant. The index number problem arises when many quotas shift simultaneously. The

1. The shadow price of the export quota is negative from Hong Kong's point of view. In the model we use, Hong Kong is assumed to have constant marginal cost of production, so the only welfare effect of a change in the quota comes through the total quota rent.

standard, atheoretic response to the problem is to use the trade-weighted average of the changes in tariff equivalents.[2]

A key problem with the tariff equivalent approach is that quota rent is internationally shared, so that the quota premium does not generally represent a marginal welfare gain to either importer or exporter (see Anderson and Neary 1992b for a recent treatment of the shadow price of quotas, applied here in the appendix; for other problems with the tariff equivalent approach, see Anderson 1988). Even setting this point aside, the problem remains that there is no theoretical basis for aggregating tariffs with a trade-weighted average. A standard complaint is that the trade weight is affected by the distortion, most dramatically when a zero quota receives a zero weight, whereas it appears that a zero quota should get a large weight.

In the case of a uniform proportional shift in all quotas, the problem collapses back to the trivial case because in effect there is only one quota. The theoretical index number solution given by Anderson and Neary (1990) builds on this insight to define the TRI as the uniform scaling factor that, applied to the new set of quotas, yields the same level of welfare as that obtained with the initial set of quotas. Defined in this way, the scaling factor rises when all quotas fall (tighten), signifying a rise in restrictiveness. The novelty of the TRI is that it uses appropriate weights for forming a weighted sum of the quota changes. Each quota change receives a weight equal to the marginal welfare contribution of a 1 percent rise in the quota relative to the marginal welfare contribution of a 1 percent rise in all quotas. The weights work out to be the trade shares evaluated in terms of the shadow prices. Because these weights are grounded in welfare theory, they are preferred to atheoretic weights such as trade shares.

We implement the TRI in an economy that abstracts from many complications. The economy is assumed to be competitive and to face no distortions other than those arising from the trade policy. A representative agent consumes and produces unconstrained tradable goods at domestic price $\pi$ and purchases or sells the goods at international prices $\pi^*$, with the difference, $\pi - \pi^*$, made up by tariffs. The representative agent also consumes and produces quota-constrained goods, for which the importer's price is $p$, the exporter's price is $p^*$, and the quota itself is $q$. (All these variables are vectors; elements of them will be denoted by a subscript.) Within a given product category all exporters' products are perfect substitutes. Each VER-constrained exporter's supply price is assumed to be constant.[3] In contrast, the importer's price is endogenous. We radically simplify the true complexity of the pricing of quota-constrained goods in order to focus on the difference our methods make to quota evaluation even in a simple model. Thus we assume that trade is executed once a year and that all quotas bind. The latter extreme assumption is partially defensible by appealing

2. For example, computable general equilibrium models usually treat quotas in this way.

3. Marginal suppliers not constrained by VERs are assumed to have capacity constraints that result in increasing marginal cost.

to the value of quota licenses as options (Anderson 1987), resulting in protection even when the quota appears not to bind. Moreover, many utilization rates for MFA quotas exceed 90 percent (Hamilton 1988).

### *Welfare Foundations*

To define the TRI, we must first characterize the equilibrium of the economy. We do this in terms of the balance-of-trade function, which equals the economy's total net expenditure on traded goods less the revenue from trade restrictions, all of which is assumed to be returned in a lump sum to the representative consumer. (Further details on the properties of the balance-of-trade function may be found in Anderson and Neary 1992b.) For the importing country, the balance-of-trade function is

$$(1) \qquad B(\pi, q, u; \pi^*, p^*) = E(\pi, q, u) - E_q' q - (\pi - \pi^*)' \, E_\pi - R.$$

(Here and in the remainder of the article, subscripts other than $i, j$, and $k$ denote partial derivatives.) The left-hand side of equation 1 states that the balance of trade depends on the home prices of unconstrained goods, $\pi$; the permitted import levels of quota-constrained goods, $q$; the utility of the representative consumer, $u$; and the world prices of both sets of goods, $\pi^*$ and $p^*$. Besides these active arguments, the balance-of-trade function depends on variables such as technology and factor endowments, which are inactive here and are hence suppressed. The first term on the right-hand side of equation 1 is equal to the "distorted" trade expenditure function, giving the net expenditure on the unconstrained goods. (Expenditure on the unconstrained goods is distorted by the presence of binding quotas.) The second term equals the total expenditure on the quota-constrained goods because $-E_q$ equals $p$, the vector of the prices of these goods in the importing country. The third term equals tariff revenue from unconstrained goods because $E_\pi$ equals net imports of those goods. The fourth and final term is the importing country's portion of the rents generated by the quota constraints, to which we return below.

For the typical exporting country, the balance-of-trade function is

$$(2) \qquad B^* (\pi^*, q, u^*; \pi, p) = E^*(\pi^*, q, u^*) - E_q^{*\prime} q - (\pi^* - \pi)' \, E_\pi^* - R^*.$$

The description of the individual terms in equation 2 is identical to that of the corresponding terms in equation 1, with $R^*$ denoting the exporting country's portion of the quota rents. The assumption that $p^*$ is constant means that $-E_q^*$ is constant in equation 2. The exporting country is assumed to be a price taker for its imports of non-MFA products, so $\pi$ is a constant. In practice, $\pi$ denotes a vector of prices for a different set of non-quota-constrained goods from those in the importing country, so that $E_\pi$ and $E_\pi^*$ refer to different sets of goods.

Turning next to the division of quota rents, we assume that rents accrue to the exporting country except for that portion retained by the importing country by imposing on the quota-constrained goods a vector of ad valorem tariffs, $\tau$. (Actual U.S. tariffs on MFA imports average about 20 percent.) The usual prac-

tice is to levy the tariff on a base value equal to the exporter's price, $p^*$, plus the market price of a unit export license, $\rho$.[4] Arbitrage implies that for each good $i$,

$$p_i = (1 + \tau_i)(p_i^* + \rho_i). \tag{3}$$

Equation 3 implies that the unit quota rent, $r_i$, equals

$$r_i = p_i - p_i^* = \rho_i + \tau_i(p_i^* + \rho_i). \tag{4}$$

Of this, the exporting country receives the license price, $\rho_i$, while the importing country receives the tariff revenue, $\tau_i(p_i^* + \rho_i)$. Hence the U.S. portion of quota rents, $R$ in equation 1, equals $\Sigma\tau_i(p_i^* + \rho_i)q_i$, while the exporting country's portion, $R^*$ in equation 2, equals $\Sigma\rho_i q_i$.

### *The Trade Restrictiveness Index*

The trade restrictiveness index is defined as follows. Starting with the new setting of the quota policy $q^1$, scale the quotas by a uniform expansion or contraction factor so that with utility at the old level, the balance of trade is unchanged and equilibrium prevails. In effect, the TRI gives the uniform quota contraction equivalent of the actual policy change. A rise in the index implies a rise in trade restrictiveness and a tightening of quotas.[5] Formally, the TRI is defined from the importer's point of view as

$$\Delta(\pi, q^1, u^0) = [\Delta \mid B(\pi, q^1\Delta, u^0) = 0].^6 \tag{5}$$

The simplest interpretation arises when the new quota vector $q^1$ rises to the free-trade level. The TRI is then less than one and is the uniform quantity contraction factor that destroys as much utility (lowers utility to the initial distorted-level $u^0$) as the initial quota vector $q^0$. That is, the initial quota policy $q^0$ is equivalent in restrictiveness to a uniform reduction in free-trade quantities by the factor $\Delta$. For other values of $q^1$, $\Delta$ is the uniform contraction factor that is equivalent to the change in restrictiveness implied in the move to $q^1$. If policy becomes more restrictive, $\Delta$ is greater than one. In principle, it is of course possible to always measure policy in relation to the free trade benchmark, but in practice, the free-trade quantities are calculated with considerable error.

4. This is true for the majority of cases in which the importing agent contracts with an exporter who already owns a license—the invoice price (which is the dutiable base) includes the value of the license. For the minority of cases in which the exporter must first purchase a license, the dutiable base in U.S. trade case law excludes the license price. Because the exporter avoids taxes in the latter case, it should be the dominant form: all exporters should buy and sell their licenses in the market as needed. The reason that is not done is that exporters lose future rights to receive licenses if they transfer them beyond a small allowance. The behavior of exporters suggests that the value of this penalty is large. In any case, our results are not sensitive to whether the dutiable base excludes or includes the value of the license.

5. In our earlier work (Anderson and Neary 1990; Anderson 1991), the index was called the coefficient of trade utilization (in recognition of its link to Debreu's coefficient of resource utilization), and the sign convention was reversed: a rise in the index denoted a fall in restrictiveness.

6. This definition of $\Delta$ is "partial" because it is for quotas only. See Anderson and Neary (1992a) for a definition of the TRI that includes variations in tariffs as well as quotas. The equilibrium value of $B$ is assumed to be zero here, but it could be any constant value.

An intuitive understanding of the TRI is helped by considering its rate-of-change form. The rate-of-change form is also easier to implement and is the one we apply in later sections. Differentiating equation 5 and using ^ to denote percentage change,

$$\hat{\Delta} = -\sum_{i}^{n}\left(\frac{-B_i q_i}{\sum - B_i q_i}\right)\hat{q}_i. \tag{6}$$

Here, $-B_i$, which is minus the derivative of the balance-of-trade function with respect to the quota on good $i$, is the shadow price of the quota on good $i$. Thus the rate of change of the TRI is equal to minus the weighted sum of the changes in individual quotas. If all weights are positive (which need not be the case), equation 6 is a weighted average. The structure of the weights is intuitive. The denominator is the marginal social value (the foreign exchange savings) of a 1 percent expansion of all the quotas. The numerator is the marginal social value (the foreign exchange savings) of a 1 percent rise in the quota on good $i$. Thus the weight is the marginal value of a 1 percent rise in quota $i$ in relation to the marginal value of a 1 percent rise in all quotas. A rise in quota $q_i$ will cause a fall in the TRI (trade restrictiveness falls) when its weight is positive, as happens when the corresponding shadow price of the quota $(-B_i)$ is positive and the denominator is positive.

For exporting countries the same operation as in equation 5 defines the TRI, except that the relevant balance-of-trade function is that of the exporting nation. Similarly, the same operation as in equation 6 defines the rate of change of the TRI for the exporting nation, except that the shadow prices of quotas are those for the exporting nation.

## II. Operationalizing the TRI

Equation 6 shows that the key step in operationalizing the TRI is to calculate the shadow prices of quotas. The details of how this is done in the present application are set out in the appendix, but two crucial simplifications should be noted. The first is the adoption of a partial equilibrium perspective. Because we are concerned with trade restrictions on a single industry only, it seems reasonable to ignore changes in the prices of nontraded goods or factors (prices of traded goods have already been held constant with the small-country assumption). The second simplifying assumption is that the importing country's expenditure function is separable with respect to the partition between MFA products and other goods and has a subexpenditure function for MFA products that has constant elasticity of substitution (CES) within the MFA group. This allows explicit calculation of the shadow prices with feasible data requirements.

The TRI is operationalized here with the formula in equation 6. This requires information on the elasticities and rates of change of constrained quantities. The substitution parameter of the CES form is inferred from an estimate of aggregate demand elasticity for textiles and apparel. The standard Divisia average is used to approximate the discrete change.

Two important issues arise in the evaluation of the MFA quota policy over time. First, the number of textiles and apparel products covered by VERs changes over time. Second, there is annual growth in quota allocations along with annual growth in the importing and exporting economies. As for the first issue, the balance-of-trade function that is the basis of equation 6 is viewed as having a very large list of quota arguments, many of which are set so high that the shadow price of the quota is equal to zero. (A nonbinding quota has no effect on the importer's price and hence no effect on any of the relevant variables.) Then, as the quota is tightened, the TRI rate-of-change formula picks it up with a nonzero shadow price (in numerator and denominator). Of course, the shadow price may jump discretely as the quota starts to bind, but this is qualitatively the same issue as a discrete change with quotas that already bind. The changing coverage phenomenon shows up in the present study in the inclusion of Bangladesh in MFA-IV, starting in 1986. (Lack of coverage in the license price data forces us to exclude some products newly subject to VER.)

As for the issue of growth, it is obvious that an unchanging quota in a growing economy is becoming more restrictive. The solution to this problem offered by Anderson and Neary (1992a) is to define the quota policy as the actual quota in relation to a quota that would leave the importer's price unchanged. In the special "neutral growth" case, the defined quota policy is equal to the actual growth rate of quotas minus the growth rate of real income. Thus, if the growth rate of quotas falls short of the growth rate of real income, quota policy has become more restrictive. In the results below, the rate of change of the TRI based on the actual policy changes is termed the uncompensated TRI rate of change, whereas the growth-adjusted change (that is, equation 6 minus the rate of change of real income in the relevant country) is termed the compensated TRI rate of change.

### *The Data*

The basic trade data for the study are the World Bank's data on shipments under the MFA. The free on board export value ($p^*q$), quantity ($q$), and export unit value ($p^*$) data for shipments from an exporter (denoted by $EX$) to the United States are extracted from the larger MFA data base. The seven exporters considered here are Bangladesh, Hong Kong, India, Indonesia, the Republic of Korea, Mexico, and Thailand. The bilateral trade data are combined with license price data ($\rho$) from two sources: for Hong Kong, assembled from work done by Carl Hamilton; and for Indonesia, from data supplied to us by the Bursa authorities for two years relevant to the study.

Following Hamilton (1988), a crude means of simulating license prices where they are missing is applied. By assuming perfect substitutability and perfect arbitrage, it can be inferred that the export supply price to the U.S. market must be the same for each good:

$$p^{*EX} + \rho^{EX} = p^{*\text{HK}} + \rho^{\text{HK}} \tag{7}$$

for any exporter $EX$ where HK denotes Hong Kong. The right-hand side of

equation 7 is data, so the license price $\rho^{EX}$ can be inferred if the variable $p^{*EX}$ can be identified. Hamilton's method is to construct marginal costs of exporters on the basis of relative wage data adjusted for labor productivity and then to equate these marginal costs to $p^*$ because of the assumption of competition. For Taiwan (China), Hamilton (1988) reported that the implied estimates of $\rho$ did not differ much from the limited license price information available. But in Krishna, Martin, and Tan (1992), Hamilton's method was checked against two years of Indonesian license price data and was found to give a large overestimate of the value of a license. One explanation of this divergence is that Hamilton's method of projecting costs is in error.[7] Fortunately, the TRI results are insensitive to across-the-board adjustments of the type implied by the productivity story, as noted below. A more problematic explanation for the divergence of the inferred from the direct license data is differences in quality, the implication being that the arbitrage assumption is not applicable. If differences in quality are important, it would mean that the results for exporters other than Hong Kong are suspect.

Commodities are aggregated within textiles and apparel to the 27 categories for which license price data for Hong Kong are available, and the years studied (1982–88) are similarly those for which complete license price data are available. Not all exporters ship in all categories, of course, and Bangladesh is subject to constraint only for 1986–88, that is, after MFA-IV. The TRI calculations also require the U.S. tariff rate on each category in each year of the study. In practice, 1986 data for the tariffs are used. As noted above, the quota policy is defined as the actual rate of change of shipments less the rate of growth of real disposable income. Because of uncertainty about the proper value of U.S. demand elasticity, we report results based on assumed U.S. substitution elasticities of 0.5, 1.0, and 2.0, with the unitary case apparently being the literature's consensus, following a suggestion of Trela and Whalley (1990).

For comparison purposes, the trade-weighted average tariff equivalent of the quota is also calculated. The (ad valorem) tariff equivalent formula for year $t$ is

$$\tilde{\nu}^t = \sum_{i=1}^{27} \left(\frac{p_i^* q_i}{\sum p_i^* q_i}\right)^t \left(\frac{\rho_i}{p_i^*}\right)^t . \tag{8}$$

The primary focus of this study is on annual rates of change in restrictiveness based on a Divisia index, so that formula 8 is used to construct the arc rate of change:

$$\frac{(\tilde{\nu}^t - \tilde{\nu}^{t-1})}{(\tilde{\nu}^t + \tilde{\nu}^{t-1})/2} . \tag{9}$$

7. Effectively, a Cobb-Douglas framework is assumed. See Krishna, Martin, and Tan (1992) for details.

## III. Results

The results of applying the TRI concept to the MFA in the 1982–88 period are presented in tables 1 to 3. The results for the base case of a U.S. substitution elasticity equal to one are in table 1, and tables 2 and 3 replicate the calculations for elasticity values of 0.5 and 2.0, respectively. For the seven exporters, the tables present the rate of change of the TRI from the exporter's point of view (that is, equation 6 with shadow prices given by appendix equation A-15) and compare it with the change in the average tariff equivalent. The last two columns of the tables present the bilateral TRI from the U.S. point of view (that is, equation 6 with shadow prices given by appendix equation A-13). The second and fourth columns are in bold face to facilitate comparison between the changes in the exporter's compensated TRI and the tariff equivalent.

The results are quite striking in their implications for the use of the TRI. First, in the base case, the yearly changes in the tariff equivalent of the quota have the opposite sign from the rate of change of the TRI in 21 of the 38 observations. The correlation coefficient between the two series is equal to 0.01, which is not significantly different from zero. Second, the results are not very sensitive to changes in the assumed elasticity of demand. Examining tables 1–3, sensitivity appears only in the results for Hong Kong from the Hong Kong point of view. This arises because Hong Kong has significant monopoly power in the low-elasticity case. Third, the results are not sensitive to variations in the method used to impute the license prices for exporters other than Hong Kong. A simple test of sensitivity to the method is performed by inferring relative unit costs from relative wages without adjusting for productivity. The imputed license prices rise by several hundred percent under this alternative method, but the TRI rates of change (not shown) are altered by less than 5 percent of the values in tables 1 to 3.

Fourth, application of the TRI concept to the Hong Kong point of view reveals very interesting properties of the TRI. In the base case, its movement over time reveals more restrictive policy in four of the six years. Hong Kong, as well as the United States, has been losing from this more restrictive policy direction. For a low value of the elasticity of U.S. import demand, however, the shadow prices of the quotas are almost all negative, whereas for the base case, about half the shadow prices are negative. A special issue of interpretation arises when $-\sum B_i q_i$, the deflator in $\hat{\Delta}$, is negative. This occurs in four of the six years for Hong Kong when the elasticity is at its lowest setting. The cases in which $-\sum B_i q_i$ is negative in table 2 are indicated by an asterisk. If $-\sum B_i q_i$ is positive, then a uniform proportional rise in $q$ is welfare-improving. If it is negative, a fall in all quotas is welfare-improving. Thus, for the low value of the demand elasticity, Hong Kong benefits from the restrictiveness of the quota policy because of its monopoly power in trade. Hong Kong's market share ranged from 13 to 20 percent in the sample years.

These results may be compared with a finding of Trela and Whalley (1990), who show that a reversion to free trade would hurt Hong Kong because of its

Table 1. *Changes in the Trade Restrictiveness Index and the Tariff Equivalent Index (Base Case: U.S. Demand Elasticity = 1.0), Selected Countries and Years*
(annual percentage change, except as specified)

| *Exporter and year* | *Change in uncompensated TRI* | *Change in compensated TRI* | *Average tariff equivalent (percent)* | *Change in average tariff equivalent* | *Change in U.S. uncompensated TRI* | *Change in U.S. compensated TRI* |
|---|---|---|---|---|---|---|
| *Bangladesh* | | | | | | |
| 1987 | −15.0 | −10.9 | 189.9 | −3.8 | −17.0 | −14.1 |
| 1988 | −19.8 | −17.0 | 182.8 | −3.8 | −17.6 | −13.1 |
| *Hong Kong* | | | | | | |
| 1983 | −2.4 | 3.1 | 30.9 | 84.1 | −2.8 | 1.2 |
| 1984 | 5.4 | 16.1 | 28.5 | −8.1 | 4.2 | 11.0 |
| 1985 | −3.9 | −1.4 | 19.2 | −39.0 | −1.7 | 1.5 |
| 1986 | −3.9 | 4.7 | 29.4 | 42.0 | −6.6 | −3.8 |
| 1987 | −0.9 | 11.8 | 33.2 | 12.1 | −1.0 | 1.9 |
| 1988 | 0.1 | 6.5 | 19.3 | −53.0 | −0.9 | 3.5 |
| *India* | | | | | | |
| 1983 | −38.4 | −31.6 | 80.0 | −14.6 | −32.8 | −28.8 |
| 1984 | 5.7 | 9.3 | 73.2 | −8.8 | 5.4 | 12.2 |
| 1985 | −12.4 | −6.4 | 127.1 | 53.8 | −12.7 | −9.5 |
| 1986 | 26.7 | 31.1 | 225.3 | 55.7 | 27.5 | 30.4 |
| 1987 | −66.8 | −62.3 | 140.6 | −46.3 | −73.1 | −70.1 |
| 1988 | −3.1 | 5.8 | 154.2 | 9.2 | −3.4 | 1.0 |
| *Indonesia* | | | | | | |
| 1983 | −7.9 | 0.5 | 23.4 | 225.4 | −8.3 | −4.4 |
| 1984 | −10.0 | −3.0 | 65.5 | 94.8 | −12.9 | −6.0 |
| 1985 | −8.1 | −4.6 | 71.8 | 9.1 | −8.3 | −5.1 |
| 1986 | −25.3 | −18.5 | 127.1 | 55.7 | −25.3 | −22.5 |
| 1987 | −2.7 | 3.2 | 168.3 | 27.9 | −4.1 | −1.2 |
| 1988 | −12.3 | −5.6 | 175.6 | 4.2 | −13.1 | −8.6 |
| *Korea, Rep. of* | | | | | | |
| 1983 | −3.8 | 8.4 | 90.8 | 40.1 | −3.9 | 0.1 |
| 1984 | 2.5 | 11.2 | 67.9 | −28.9 | 2.6 | 9.4 |
| 1985 | −21.6 | −14.8 | 96.0 | 34.3 | −16.7 | −13.5 |
| 1986 | −5.9 | 6.1 | 74.7 | −24.9 | −4.8 | −1.9 |
| 1987 | −0.9 | 10.9 | 56.1 | −28.5 | −0.6 | 2.3 |
| 1988 | 2.0 | 13.5 | 27.3 | −69.1 | 2.9 | 7.4 |
| *Mexico* | | | | | | |
| 1983 | −23.7 | −26.9 | 67.7 | 218.9 | −25.4 | −21.5 |
| 1984 | −23.6 | −19.4 | 40.5 | −50.3 | −22.3 | −15.5 |
| 1985 | 50.2 | 53.9 | 52.7 | 26.2 | 34.9 | 38.1 |
| 1986 | −73.2 | −78.4 | 56.2 | 6.3 | −61.3 | −58.5 |
| 1987 | −39.6 | −36.8 | 80.0 | 35.0 | −31.9 | −29.0 |
| 1988 | −10.6 | −8.3 | 77.9 | −2.7 | −9.6 | −5.1 |
| *Thailand* | | | | | | |
| 1983 | −19.7 | −11.5 | 72.8 | −16.7 | −20.3 | −16.4 |
| 1984 | −5.3 | 1.1 | 38.3 | −62.2 | −9.3 | −2.5 |
| 1985 | 33.7 | 36.3 | 67.5 | 55.2 | 34.9 | 38.1 |
| 1986 | −37.9 | −33.4 | 48.7 | −32.3 | −31.0 | −28.2 |
| 1987 | 8.3 | 17.7 | 50.9 | 4.4 | 12.3 | 15.3 |
| 1988 | −3.3 | 9.4 | 45.5 | −11.2 | 0.0 | 4.5 |

*Source:* Authors' calculations.

Table 2. *Changes in the Trade Restrictiveness Index and the Tariff Equivalent Index (U.S. Demand Elasticity = 0.5), Selected Countries and Years*
(annual percentage change, except as specified)

| *Exporter and year* | *Change in uncompensated TRI* | *Change in compensated TRI* | *Average tariff equivalent (percent)* | *Change in average tariff equivalent* | *Change in U.S. uncompensated TRI* | *Change in U.S. compensated TRI* |
|---|---|---|---|---|---|---|
| *Bangladesh* | | | | | | |
| 1987 | −15.0 | −10.9 | 189.9 | −3.8 | −16.8 | −13.9 |
| 1988 | −19.8 | −17.0 | 182.8 | −3.8 | −17.5 | −13.0 |
| *Hong Kong* | | | | | | |
| 1983* | −4.1 | 1.4 | 30.9 | 84.1 | −2.6 | 1.4 |
| 1984* | 3.7 | 14.3 | 28.5 | −8.1 | 4.1 | 10.9 |
| 1985* | −0.5 | 2.0 | 19.2 | −39.0 | −1.9 | 0.0 |
| 1986 | 0.0 | 8.6 | 29.4 | 42.0 | −6.5 | 1.4 |
| 1987 | 8.0 | 20.8 | 33.2 | 12.1 | −1.1 | −3.7 |
| 1988* | 0.9 | 7.3 | 19.3 | −53.0 | −0.8 | 1.9 |
| *India* | | | | | | |
| 1983 | −38.5 | −31.7 | 80.0 | −14.6 | −33.9 | −30.0 |
| 1984 | 5.8 | 9.3 | 73.2 | −8.8 | 5.2 | 12.0 |
| 1985 | −12.4 | −6.4 | 127.1 | 53.8 | −12.5 | −9.3 |
| 1986 | 26.7 | 31.1 | 225.3 | 55.7 | 27.1 | 29.9 |
| 1987 | −66.7 | −62.3 | 140.6 | −46.3 | −71.1 | −68.1 |
| 1988 | −3.1 | 5.8 | 154.2 | 9.2 | −3.4 | 1.0 |
| *Indonesia* | | | | | | |
| 1983 | −8.3 | 0.2 | 23.4 | 225.4 | −8.3 | −4.4 |
| 1984 | −9.9 | −2.9 | 65.5 | 94.8 | −13.0 | −6.2 |
| 1985 | −8.1 | −4.7 | 71.8 | 9.1 | −8.0 | −4.8 |
| 1986 | −25.2 | −18.4 | 127.1 | 55.7 | −25.5 | −22.7 |
| 1987 | −2.6 | 3.2 | 168.3 | 27.9 | −4.0 | −1.1 |
| 1988 | −12.3 | −5.6 | 175.6 | 4.2 | −13.0 | −8.6 |
| *Korea, Rep. of* | | | | | | |
| 1983 | −3.8 | 8.4 | 90.8 | 40.1 | −3.9 | 0.0 |
| 1984 | 2.6 | 11.2 | 67.9 | −28.9 | 2.6 | 9.4 |
| 1985 | −23.2 | −16.4 | 96.0 | 34.3 | −16.5 | −13.3 |
| 1986 | −6.3 | 5.8 | 74.7 | −24.9 | −4.9 | −2.0 |
| 1987 | −1.0 | 10.8 | 56.1 | −28.5 | −0.6 | 2.3 |
| 1988 | 7.6 | 19.1 | 27.3 | −69.1 | 2.9 | 7.4 |
| *Mexico* | | | | | | |
| 1983 | −24.7 | −27.9 | 67.7 | 218.9 | −25.7 | −21.8 |
| 1984 | −23.6 | −19.5 | 40.5 | −50.3 | −22.5 | −15.6 |
| 1985 | 51.0 | 54.6 | 52.7 | 26.2 | 36.0 | 39.3 |
| 1986 | −73.7 | −78.8 | 56.2 | 6.3 | −62.4 | −59.6 |
| 1987 | −39.9 | −37.1 | 80.0 | 35.0 | −32.3 | −29.4 |
| 1988 | −10.7 | −8.4 | 77.9 | −2.7 | −9.8 | −5.3 |
| *Thailand* | | | | | | |
| 1983 | −19.7 | −11.5 | 72.8 | −16.7 | −20.5 | −16.6 |
| 1984 | −5.1 | 1.2 | 38.3 | −62.2 | −8.9 | −2.0 |
| 1985 | 33.7 | 36.3 | 67.5 | 55.2 | 34.9 | 38.1 |
| 1986 | −38.1 | −33.6 | 48.7 | −32.3 | −31.7 | −28.9 |
| 1987 | 8.2 | 17.6 | 50.9 | 4.4 | 12.4 | 15.3 |
| 1988 | −3.4 | 9.3 | 45.5 | −11.2 | −0.2 | 4.3 |

*Note:* An asterisk (*) denotes negative shadow value of distorted trade.
*Source:* Authors' calculations.

Table 3. *Changes in the Trade Restrictiveness Index and the Tariff Equivalent Index (U.S. Demand Elasticity = 2.0), Selected Countries and Years*
(annual percentage change, except as specified)

| *Exporter and year* | *Change in uncompensated TRI* | *Change in compensated TRI* | *Average tariff equivalent (percent)* | *Change in average tariff equivalent* | *Change in U.S. uncompensated TRI* | *Change in U.S. compensated TRI* |
|---|---|---|---|---|---|---|
| *Bangladesh* | | | | | | |
| 1987 | −15.0 | −10.9 | 189.9 | −3.8 | −17.4 | −14.5 |
| 1988 | −19.7 | −17.0 | 182.8 | −3.8 | −17.7 | −13.3 |
| *Hong Kong* | | | | | | |
| 1983 | 2.7 | 8.2 | 30.9 | 84.1 | −3.0 | 1.0 |
| 1984 | 4.6 | 15.2 | 28.5 | −8.1 | 4.3 | 11.1 |
| 1985 | −3.0 | −0.6 | 19.2 | −39.0 | −1.6 | 1.7 |
| 1986 | −5.1 | 3.5 | 29.4 | 42.0 | −6.7 | −3.9 |
| 1987 | −1.1 | 11.6 | 33.2 | 12.1 | −0.9 | 2.0 |
| 1988 | −0.1 | 6.3 | 19.3 | −53.0 | −1.1 | 3.4 |
| *India* | | | | | | |
| 1983 | −38.4 | −31.5 | 80.0 | −14.6 | −30.8 | −26.9 |
| 1984 | 5.7 | 9.3 | 73.2 | −8.8 | 5.7 | 12.5 |
| 1985 | −12.4 | −6.4 | 127.1 | 53.8 | −13.2 | −9.9 |
| 1986 | 26.7 | 31.1 | 225.3 | 55.7 | 28.3 | 31.2 |
| 1987 | −66.8 | −62.4 | 140.6 | −46.3 | −76.5 | −73.5 |
| 1988 | −3.1 | 5.8 | 154.2 | 9.2 | −3.4 | 1.1 |
| *Indonesia* | | | | | | |
| 1983 | −7.6 | 0.8 | 23.4 | 225.4 | −8.3 | −4.4 |
| 1984 | −10.1 | −3.1 | 65.5 | 94.8 | −12.7 | −5.8 |
| 1985 | −8.1 | −4.6 | 71.8 | 9.1 | −8.7 | −5.5 |
| 1986 | −25.3 | −18.5 | 127.1 | 55.7 | −24.9 | −22.1 |
| 1987 | −2.7 | 3.1 | 168.3 | 27.9 | −4.2 | −1.3 |
| 1988 | −12.3 | −5.6 | 175.6 | 4.2 | −13.1 | −8.6 |
| *Korea, Rep. of* | | | | | | |
| 1983 | −3.8 | 8.4 | 90.8 | 40.1 | −3.8 | 0.1 |
| 1984 | 2.5 | 11.2 | 67.9 | −28.9 | 2.7 | 9.5 |
| 1985 | −21.0 | −14.3 | 96.0 | 34.3 | −16.9 | −13.7 |
| 1986 | −5.8 | 6.2 | 74.7 | −24.9 | −4.6 | −1.8 |
| 1987 | −0.9 | 11.0 | 56.1 | −28.5 | −0.5 | 2.4 |
| 1988 | 2.4 | 13.9 | 27.3 | −69.1 | 2.9 | 7.4 |
| *Mexico* | | | | | | |
| 1983 | −22.7 | −26.0 | 67.7 | 218.9 | −25.0 | −21.0 |
| 1984 | −23.5 | −19.4 | 40.5 | −50.3 | −22.1 | −15.3 |
| 1985 | 49.9 | 53.6 | 52.7 | 26.2 | 33.1 | 36.3 |
| 1986 | −73.0 | −78.2 | 56.2 | 6.3 | −59.5 | −56.7 |
| 1987 | −39.5 | −36.7 | 80.0 | 35.0 | −31.3 | −28.4 |
| 1988 | −10.6 | −8.3 | 77.9 | −2.7 | −9.3 | −4.8 |
| *Thailand* | | | | | | |
| 1983 | −19.8 | −11.5 | 72.8 | −16.7 | −19.9 | −16.0 |
| 1984 | −5.3 | 1.0 | 38.3 | −62.2 | −10.0 | −3.2 |
| 1985 | 33.7 | 36.4 | 67.5 | 55.2 | 34.8 | 38.0 |
| 1986 | −37.8 | −33.3 | 48.7 | −32.3 | −30.1 | −27.2 |
| 1987 | 8.4 | 17.8 | 50.9 | 4.4 | 12.3 | 15.2 |
| 1988 | −3.3 | 9.4 | 45.5 | −11.2 | 0.3 | 4.8 |

*Source:* Authors' calculations.

large terms of trade loss. The results are not fully comparable because free trade entails a loss of all the quota rent and because of other modeling reasons. We use Trela and Whalley's U.S. demand elasticity for our base case. For our low-elasticity case, we obtain the similar result that the tighter restrictiveness shown in the rising license prices of the 1980s does indeed benefit Hong Kong.

Finally, a natural use of the TRI is to compare the evolution of trade policy across countries. The relative treatment of exporters in the MFA is a subject of considerable international political interest and pressure, peaking at negotiation times. The results show a large degree of nonuniformity across exporters in the evolution of MFA policy, broadly favoring the lower-cost exporters. Tables 1 to 3 show that MFA trade policy became more restrictive toward Korea and Thailand, as well as Hong Kong. In contrast, policy became less restrictive toward Indonesia, Mexico, and Bangladesh. Toward India, policy evolved more ambiguously but on balance became somewhat less restrictive. Significantly, this policy difference across exporters is not well described by the evolution of the average tariff equivalent index.

Each bilateral policy might naturally be evaluated from either the exporter's or the importer's point of view. The main emphasis here is on the exporter's point of view, but political pressure by exporters on importers might well be countered with use of a TRI based on the importer's point of view. Comparing the two TRIs, the adjustment of quota policy for growth makes a significant difference to the policy conclusion. With high-growth exporters and low-growth importers, the change in point of view is enough to make the evaluation of quota policy using the compensated TRI reverse sign in four cases of small monopoly power (two for Indonesia, one for Korea, and one for Thailand), besides two cases for Hong Kong that mix this effect with the effect of monopoly power. In such cases, a more appropriate index is the uncompensated TRI for the exporter or importer. The uncompensated TRIs differ in sign in only one observation in the base case, that of Hong Kong for 1988.

We conclude our discussion of results with some speculation as to why the results for the TRI and the average tariff equivalent results are uncorrelated. The trade-weighted average tariff equivalent has no theoretical foundation, of course, but its common use (along with the even more common use of average tariffs for non-quota-constrained goods) makes it a natural benchmark. The lack of correlation is perplexing in light of intuition based on only one quota-constrained good. In this case the rate of change of the license price and minus the rate of change of the quota necessarily have the same sign.[8] Where does the difference in the two indexes come from?

First, license prices change for many reasons other than shifts in the quota. This can explain an absence of strong negative correlation between changes in

8. The rate of change of the domestic price is equal to the elasticity of demand multiplied by minus the rate of change of the quota. The rate of change of the license price is equal to the rate of change of the domestic price times the ratio of the domestic price to the license price.

license prices and changes in quotas over time for each good. On the supply side of the market, export prices shift with costs as well as with the level of the importer's quotas. On the demand side of the market, domestic prices change over time because of changes in national income and in the prices of unconstrained goods, as well as changes in quotas. It should also be noted that the generalized law of demand, even for the CES case, does not guarantee perfect negative correlation of $p$ and $q$, other things being equal.

Second, there is a difference in practice between the TRI weights and the trade weights used to form the indexes. For the importer this difference disappears in the CES case only under the (empirically false) restriction that the rent-retaining tariff is uniform over quota-constrained goods.[9] For the exporter, much more extreme false assumptions are required to reduce TRI weights to trade weights.[10] It is possible to construct an argument that the weights of the TRI and the tariff equivalent tend to be negatively correlated in the cross-section of quota-constrained goods because of the dispersion of rent-retaining tariffs and of ad valorem license prices.[11]

We conclude that theory gives no reason to expect that the average tariff equivalent will behave similarly to the TRI. Where it diverges, the TRI is a properly weighted index of changes in the actual policy, the vector of quotas, whereas the average tariff equivalent is an atheoretically weighted index of changes in the license prices (or price differentials), which themselves can be the result of changes in many factors other than policy.

## IV. Conclusion

This article has applied the TRI to one of the most important trade-distorted industries: textiles and apparel. Use of the TRI is strongly indicated because its implications are very often opposed to those based on standard atheoretic methods, and its implementation requires only the elements needed in any trade policy model of quotas.

The results show wide annual swings in the TRI as well as very significant differences in the movement of the index across countries. Toward Hong Kong, Korea, and Thailand, U.S. trade policy became more restrictive. Toward Bangladesh, Indonesia, and Mexico, policy became more liberal. Toward India, the policy evolution was more mixed but on balance liberal.

9. Under the assumptions, applying equation A-13 of the appendix, $-B_j$ is proportional to $p_j$. The TRI weights of equation 6 contain $-B_j$ in both numerator and denominator, hence the factor of proportionality cancels, and the TRI weights are equal to trade weights.

10. For an exporter with a tiny market share, the shadow price of a quota is approximately equal to its license price. If by chance the initial quota allocation (over the commodity categories, for all countries) is such that all the individual exporter's license prices are uniformly proportional to the domestic price (a uniform tariff equivalent), the factor of proportionality cancels out of the numerator and denominator of the weights, and the form of the TRI in rates of change is that of the analogue (for exporters) of equation 6.

11. The argument is rather technical, so we do not pursue it here.

Most significantly, the TRI and the average tariff equivalent are independent in the sense of descriptive statistics (the correlation coefficient for the two series is equal to 0.01), and their behavior is so significantly different as to lead to qualitatively different answers more than half the time (21 out of 38 observations for the base-case value of the U.S. demand elasticity). This replicates a finding of Anderson (1991) for imports of U.S. cheese, but for a much more important industry.

Finally, the results show that the TRI measure is not very sensitive to the main sources of potential measurement error: the value of the license price and the value of the elasticity of demand. Inference about trade restrictiveness from tariff equivalents depends on obtaining reliable data on imputed quota license premiums where these are not directly observable. Krishna, Martin, and Tan (1992) show that imputation of MFA quota license prices from constructed costs, a method that has become standard, can be seriously misleading. In contrast, although the TRI requires license prices, it turns out to be quite insensitive to error in the imputation process in the application. The main problematic value is the appropriate adjustment for growth. Here, the analyst can at least provide a range of reasonable values by using the compensated and uncompensated measures of the TRI.

There are two important implications of this study for future work on quota policy evaluation in general and MFA policy evaluation in particular. First, inferences based on average tariff equivalents are very seriously misleading. Second, feasible alternatives exist, and they can be used to form aggregate indexes that can be used for many modeling purposes.

Despite these achievements, we do not claim that our methods have definitively pinned down the subject of our title. The excellent survey of Trela and Whalley (1990) emphasizes the many dimensions of textile and apparel markets that have so far eluded modelers. Phenomena such as shipment growth in excess of allowed quota growth rates, positive license prices with low utilization rates, and quality upgrading are important and are not explained by the model used here.

## Appendix. The Shadow Prices of Quotas in the CES Case

This appendix sets out an operational solution for the calculation of the MFA shadow prices of quotas. We assume a constant elasticity of substitution (CES) subexpenditure system for imports that is automatically separable into all MFA categories and all nonconstrained imports, thus permitting the simplicity of the separable structure of Anderson and Neary (1992b). The first subsection derives the distorted expenditure system and then the inverse elasticity system for this CES case. The second subsection derives the shadow prices.

### *Inverse Demand Elasticity System in the CES Case*

Let the subexpenditure function for U.S. imports be written

$$e(p, \pi, u) = (\sum \beta_k p_k^{1-\sigma} + \sum \alpha_j \pi_j^{1-\alpha})^{1/(1-\sigma)} u \tag{A-1}$$

where $u$ is the level of utility, $p$ is the domestic price of quota-constrained goods (MFA textiles and apparel), and $\pi$ is the domestic price of non-quota-constrained imports. The Armington assumption is imposed here so that there is no domestic production of a perfect substitute for imports. The elasticity of substitution is equal to the parameter $\sigma$, and the $\alpha$s and $\beta$s are share parameters for the non-quota-constrained goods and the quota-constrained goods, respectively. For empirical work it is convenient to pick a base year for which prices are initially set to one. This implies that the $\alpha$s and $\beta$s are the initial values of the expenditure share values in the base data and that the initial level of expenditure is equal to $u$. The true (that is, welfare-consistent) cost of living index is

$$P = (\sum \beta_k p_k^{1-\sigma} + \sum \alpha_j \pi_j^{1-\sigma})^{1/(1-\sigma)}. \tag{A-2}$$

The quota-constrained imports are subject to fixed binding quotas equal to $q_k$ for all $k$. This results in a distorted expenditure function for the unconstrained goods. The distorted expenditure function is defined (as in Anderson and Neary 1992b) by

$$E(\pi, q, u) = \max_p \; [e(p, \pi, u) - p'q].$$

The vector of maximizing prices that solves this program is a virtual price vector, each virtual price being the consumer's marginal willingness to pay for one more unit of the constrained good (Neary and Roberts 1980). In the context of quotas, virtual prices are also market-clearing prices. Using Shephard's lemma, and solving the first-order (market-clearing) condition for the virtual and market price of each quota-constrained good $k$, we obtain

$$p_k = P\left(\frac{u\beta_k}{q_k}\right)^{1/\sigma}. \tag{A-3}$$

Substituting equation A-2 into equation A-3, the vector of virtual prices $p$ is implicitly defined as a function of the $\pi$s and the quotas. Fortunately, an explicit solution is available. First, substitute equation A-3 into equation A-2. Next, raise both right- and left-hand sides to the power $1 - \sigma$. Then, solve the resulting expression for $P^{1-\sigma}$. Finally, raise both sides to the power $1/(1 - \sigma)$. The reduced-form true cost of living price index is

$$P = P(\pi, q, u) = \left(\frac{\sum \alpha_j \pi_j^{1-\sigma}}{u^{(1-\sigma)/\sigma} \sum \beta_k^{(1-\sigma)/\sigma} q_k^{-(1-\sigma)/\sigma}}\right)^{1/(1-\sigma)}. \tag{A-4}$$

The connection of equation A-4 to equation A-2 is clear: if consumers face fixed prices $p$ at the level of virtual prices $p$ defined by equation A-3, their cost of living is the same as when constrained by quotas $q$.

The distorted-expenditure function is obtained by substituting equations A-3 and A-4 into the definition of $E$:

$$\text{(A-5)} \qquad E(\pi, q, u) = P(\pi, q, u)u - p'q$$
$$= Pu\left(1 - u^{(1-\sigma)/\sigma} \sum_k \beta_k^{(1/\sigma)} q_k^{1-(1/\sigma)}\right).$$

Using equation A-4, this can be factored into

$$\text{(A-6)} \qquad E(\pi, q, u) = \left(\sum_j \alpha_j \pi_j^{1-\sigma}\right) P(\pi, q, u)^{\sigma} u$$

where $P$ is given by equation A-4. The constrained demand for non-quota-constrained imports is obtained from use of Shephard's lemma:

$$\text{(A-7)} \qquad E_{\pi_j} = \alpha_j \left(\frac{\pi_j}{P}\right)^{-\sigma} u.$$

The virtual price vector is obtained as

$$\text{(A-8)} \qquad -E_{q_k} = p_k = E \frac{u^{(1-\sigma)/\sigma} \beta_k^{(1/\sigma)} q_k^{-(1/\sigma)}}{1 - u^{(1-\sigma)/\sigma} \sum \beta_k^{(1/\sigma)} q_k^{-(1-\sigma)/\sigma}}.$$

Now consider the inverse demand elasticity system. Differentiating equation A-8 with respect to $q_i$, using equations A-6 and A-4, and multiplying by $q_i/p_k$ to form the elasticities,

$$\text{(A-9)} \qquad \frac{q_i \partial p_k}{p_k \partial q_i} = -\frac{1}{\sigma}(\delta_{ki} + s_i \eta);$$

where $\delta_{ki}$ is the Kronecker delta ($\delta_{ki} = 0$ for $i \neq k$, and $\delta_{kk} = 1$), $s_i$ is the within-group expenditure share of constrained good $i$, $p_i q_i / p'q$, and $\eta$ is the constrained good's expenditure share of all expenditure.

In practice, it is important to differentiate between exporters for each good $i$. We assume that all exporters of good $i$ produce a perfect substitute product. Then the inverse elasticity of domestic price $k$ with respect to exports of good $i$ from exporter $h$ is written

$$\text{(A-10)} \qquad \frac{q_i^h \partial p_k}{p_k \partial q_i^h} = -\frac{1}{\sigma}(\delta_{ki} + s_i \eta)\theta_i^h$$

where $\theta_i^h$ is the share of good $i$ sold by exporter $h$, $q_i^h/q_i$. In what follows, we dispense with explicitly indexing the exporter and present expressions for a "typical" exporter.

### *The Shadow Price of Quotas*

The shadow price of quota $k$ in the case where export licenses are included in the tax base for ad valorem import tariffs is based on the arbitrage equation

$$\text{(A-11)} \qquad p_k = (p_k^* + \rho_k)(1 + \tau_k)$$

where $p_k^*$ is the typical exporter's supply price for good $k$ and $\rho_k$ is the typical exporter's license price for good $k$.

The shadow price of quotas for the United States is found by differentiating the U.S. balance-of-trade function with respect to the typical exporter's sales of good $i$:

$$\text{(A-12)} \qquad -B_i = -\sum_k q_k \frac{\partial p_k}{\partial q_i} + \tau_i(p_i^* + \rho_i) + \sum_k q_k \tau_k \frac{\partial \rho_k}{\partial q_i} - vp_i$$

where $v$ is equal to the trade-weighted average tariff on non-MFA-constrained imports. The $vp_i$ term is the cross-effect of quota change $i$ on tariff revenue collected on non-MFA-constrained imports, using the results of Anderson and Neary (1992b) for the separable case. We simplify on the right-hand side of equation A-12 by substituting aggregate exports using the arbitrage equation A-11 where possible. First, differentiating equation A-11 and using the assumption of constant supply price, $\partial\rho_k/\partial q_i$ is equal to $(\partial p_k/\partial q_i)/(1+\tau_k)$. Then

$$-B_i = \sum_k q_k \frac{\partial p_k}{\partial q_i} + \tau_i(p_i^* + \rho_i) + \sum_k q_k \frac{\tau_k}{1+\tau_k} \frac{\partial p_k}{\partial q_i} - vp_i.$$

Next, combining the first and third terms and using equation A-11,

$$-B_i = \frac{\tau_i}{1+\tau_i} p_i - \sum_k q_k \frac{1}{1+\tau_k} \frac{\partial p_k}{\partial q_i} - vp_i.$$

Finally, using symmetry of the inverse-compensated-demand system,

$$\text{(A-12}'\text{)} \qquad -B_i = \frac{\tau_i}{1+\tau_i} p_i - \sum_k \frac{1}{1+\tau_k} \frac{q_k \partial p_i}{p_i \partial q_k} p_i - vp_i.$$

(In the middle term of the last line the summation runs over all exporters.)

Substituting from equation A-9 into the next-to-last expression in equation A-12′, rearranging terms, and using the arbitrage equation,

$$\text{(A-13)} \qquad -B_i = \frac{(1/\sigma) + \tau_i}{1+\tau_i} p_i + \frac{\eta}{\sigma} \phi p_i - vp_i$$

where

$$\phi = \sum s_k \frac{1}{1+\tau_k} = \frac{1}{1+\tau}.$$

The term $\phi$ is the trade-weighted average tariff factor deflator.

For the typical exporter,

$$\text{(A-14)} \qquad -B_i^* = \rho_i + \sum_k q_k \frac{1}{1+\tau_k} \frac{\partial p_k}{\partial q_i}.$$

There is no cross effect with tariffs because of the constant-supply-price assumption. The symmetry of the inverse-demand system and the simple elasticity form equation A-9 reduce equation A-14 to

$$\text{(A-15)} \qquad -B_i^* = \rho_i - \frac{\theta_i}{\sigma} \frac{1}{1+\tau_i} p_i - \left(\frac{\eta}{\sigma} \sum_i \theta_k s_k \frac{1}{1+\tau_k}\right) p_i.$$

The last bracketed term is close to zero because MFA expenditure is rather small as a share of the total budget ($\eta$ is small) and is being multiplied by another share, $\theta$, which is also small—even Hong Kong has less than a quarter of the market. In practice, we ignore this term to ease the burden of calculation and data collection.

## References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Anderson, James E. 1987. "Quotas as Options: Optimality and Quota License Pricing under Uncertainty." *Journal of International Economics* 23(1–2):21–39.

———. 1988. *The Relative Inefficiency of Quotas*. Cambridge: MIT Press.

———. 1991. "The Coefficient of Trade Utilisation: The Cheese Case." In Robert E. Baldwin, ed., *Empirical Studies of Commercial Policy*. University of Chicago Press.

Anderson, James E., and J. Peter Neary. 1990. "The Coefficient of Trade Utilization: Back to the Baldwin Envelope." In Ronald W. Jones and Anne O. Krueger, eds., *The Political Economy of International Trade: Essays in Honor of Robert E. Baldwin*. Oxford: Basil Blackwell.

———. 1992a. "A New Approach to Evaluating Trade Policy." WPS 1022. World Bank, International Economics Department, Washington, D.C. Processed.

———. 1992b. "Trade Reform with Quotas, Partial Rent Retention and Tariffs." *Econometrica* 60(1):57–76.

———. 1994. "Measuring the Restrictiveness of Trade Policy." *The World Bank Economic Review* 8(2):151–69.

Hamilton, Carl B. 1988. "Restrictiveness and International Transmission of the 'New Protectionism.'" In Robert E. Baldwin, Carl B. Hamilton, and Andre Sapir, eds., *Issues in US-EC Trade Relations*. University of Chicago Press.

Hamilton, Carl B., and Chungsoo Kim. 1990. "Republic of Korea: Rapid Growth in Spite of Protectionism Abroad." In Carl B. Hamilton, ed., *Textiles Trade and the Developing Countries*. Washington, D.C.: World Bank.

Krishna, Kala, Will Martin, and Ling-Hui Tan. 1992. "Imputing License Prices: Limitations of a Cost-based Approach." Pennsylvania State University, Department of Economics, College Park, Pa. Processed.

Neary, J. Peter, and Kevin W. S. Roberts. 1980. "The Theory of Household Behaviour under Rationing." *European Economic Review* 13(1):25–42.

Trela, Irene, and John Whalley. 1990. "Unraveling the Threads of the MFA." In Carl B. Hamilton, ed., *Textiles Trade and the Developing Countries*. Washington, D.C.: World Bank.

# Labor Supply and Targeting in Poverty Alleviation Programs

Ravi Kanbur, Michael Keen, and Matti Tuomala

*The introduction of variable labor supply raises some fundamental issues in analyzing the targeting of poverty alleviation programs in developing countries. It forces a reconsideration of the standard objective function, which is based on income or expenditure and so makes no allowance for the effort made in earning that income. We show that alternative views on the appropriate valuation of effort have very different implications for commodity-based targeting rules. We also establish a benchmark for marginal effective tax rates (inclusive of benefit withdrawal) in income-tested schemes and show that indicator targeting rules may also have to be modified significantly when labor supply responses are recognized.*

For many governments of developing countries, finer targeting of programs to alleviate poverty appears an attractive option in an era of greatly constrained expenditure budgets. It seems as though policymakers could achieve greater poverty reduction with fewer resources if only they would resort to the magic of targeting. But fine targeting is not without its costs. It is now appreciated that the administrative costs of ensuring that benefits from a program reach the target group can be high (see Besley and Kanbur 1993). One response is to target by subsidizing commodities largely consumed by the poor or on the basis of other observable indicators—such as age, gender, region, or crop group—that are correlated with deprivation. There is now a literature on how such indicators might be used for optimal targeting (see, for instance, Akerlof 1978, Atkinson 1992, Besley and Kanbur 1988, and Kanbur 1987).

An aspect of the costs of fine targeting that has not been as well appreciated in the development literature as it should be is the effect on incentives. Consider, for example, the moves in Sri Lanka to target the rice ration subsidy in the wake of the economic reforms of the late 1970s (see Anand and Kanbur 1991). The system was transformed from one with a universal benefit to one in which the benefit was restricted (in principle) to those with incomes below a critical level

Ravi Kanbur is with the Western Africa Department at the World Bank, Michael Keen is with the Department of Economics at the University of Essex, and Matti Tuomala is with the Department of Economics at the University of Jyväskylä. The authors are grateful to Steve Coate, Jonathan Morduch, Kim Nead, and Dominique van de Walle for helpful comments.

(although in practice, as discussed by Sahn and Alderman 1992, enforcement of the income test was highly imperfect). Such income testing targets benefits on the poor, but it is administratively costly and, in particular, has incentive effects that diminish its appeal. Consider the position of a household that increases its income, through its own efforts, to such an extent that it crosses the critical threshold and so (with perfect enforcement) loses eligibility for the benefit. The benefit's relation to income clearly reduces the incentive for the household to increase its income. Such effects are a familiar part of the policy debate in industrial countries (a representative example being Dilnot and Stark 1989), but debate over such matters has not been prominent in developing countries. The potential practical significance of incentive effects is emphasized, however, by the empirical work of Sahn and Alderman (1992). They find that the Sri Lankan rice subsidy was associated with a substantial reduction in the labor supply of recipients, with potentially important implications for the evaluation of the scheme.

The purpose of this article is therefore to take up the broad issue of targeting and incentives in developing countries and, in particular, to explore the implications of variable labor supply for the design of poverty alleviation programs. It will be seen that once the potential incentive effects of such programs are recognized, previous discussions of optimal targeting require revision. Simple rules of thumb—for example, "spend more on the group with greater measured poverty"—have to be modified to take into account such features as differing labor supply elasticities. Similarly, commodity-based targeting rules need to be modified.

The general issues have been extensively addressed in the optimal tax literature, with tools and results that can be borrowed for the analysis of the incentive effects of targeting in developing countries. The approach pursued, however, departs from the usual optimal tax tradition in two ways. First, we take the objective of policy to be the minimization of a poverty index rather than the maximization of a social welfare function. This approach is not without its critics (see, for instance, Stern 1987), but in a technical sense at least it is relatively straightforward. Loosely speaking, a poverty index defined on utility—attaching zero weight to all households above some threshold—is merely a special form of social welfare function.

Our second point of departure is more fundamental. The motivation for this departure begins with the observation that the poverty indexes on which much policy discussion focuses are, in practice, almost invariably defined in terms not of utility but of income. In the presence of incentive effects, these criteria are very different things (and can move in different directions). Indexes that focus on income attach no significance to the effort put into earning income or, put another way, attach no weight to the leisure of the poor. And indeed it is clear that much policy debate is cast in precisely these terms: the focus is on the income of the poor, not on how hard they work to get it. This is not to deny that, from the Factory Acts to the Maastricht Treaty, policymakers have been

concerned about avoiding excessive work hours. In 1904, for instance, Churchill was arguing eloquently that "[working people] demand time to look about them, time to see their homes by daylight, time to see their children, time to think and read and cultivate their gardens—time, in short, to live" (quoted in Gilbert 1991: 196). The point is rather one of emphasis.

A large part of our purpose here is to explore the implications of a nonwelfarist approach to policy analysis. We do not necessarily advocate the evaluation of policy by its effects on some income-based poverty index, but it may nevertheless be useful to explore the implications of doing so, for two reasons. First, to the extent that—right or wrong—policy is often evaluated, at least in part, by the use of such indexes, it is helpful to know what kind of policy would be implied by the explicit pursuit of such a minimand. Second, following the work of Sen (1985), there has been growing interest in nonwelfarist approaches to policy analysis. Yet there has been relatively little formal work along these lines (an exception is Ulph 1991). The work reported here may contribute to the development of this research agenda. It should immediately be stressed, however, that we do not set out to capture the richness of Sen's capabilities approach. For although there is only one kind of welfarism, there are potentially very many kinds of nonwelfarism. Here we report on and extend recent work on one perhaps crude form of nonwelfarism, one that has the merits of capturing the common preoccupation with income-based measures of poverty and, moreover, of being readily tractable.

Section I tackles the underlying conceptual problem: the measurement of poverty when labor supply is variable. The rest of the article applies this nonwelfarist approach to three aspects of targeting, focusing in each case on the role played by labor supply responses. Section II considers commodity-based targeting, section III analyzes income-based targeting, and section IV discusses the use of other observable characteristics for indicator targeting. Section V offers conclusions.

## I. Labor Supply and the Measurement of Poverty

The measurement of poverty is of course a vast topic in itself, and we make no attempt to survey it here (see Atkinson 1987 and, for a general survey, Ravallion 1993). But before addressing the implications of variable labor supply for the design of poverty alleviation programs, we need to specify the way in which labor supply affects the perceived extent of poverty, this being the presumed minimand for the policy exercises in subsequent sections. The first task is thus to consider the ways in which variable labor supply might be incorporated into the measurement of poverty. These issues have been little discussed and, as will be seen, remain somewhat perplexing.

The standard approach to the measurement of poverty proceeds by comparing $y^h$, the income some individual $h$ has available to spend, to a poverty

line $z$.[1] Because $h$ is poor if $y^h \leq z$, aggregate poverty might be measured by an index of the form

$$P = \int_0^\infty D(z, y) f(y)\, dy \tag{1}$$

where $D(z, y)$ can be thought of as the deprivation of an individual with income $y$, and $f(y)$ denotes the density of $y$. This is a fairly general form of poverty index, encompassing a range of widely used measures as special cases. (The only substantive restriction is additive separability, which precludes Gini-based indexes such as that of Sen 1976.) If $D(z, y) = \max(z - y, 0)$, for example, then $P$ is the aggregate poverty gap. For our purposes, we need not assume any particular form for $D(\cdot)$ beyond making the natural assumptions that deprivation is positive only for the poor, so that $D(z, y)$ is strictly positive if $y < z$ and zero otherwise, and decreases as income rises toward the poverty line (so that $D_y < 0$ for $y < z$, with the subscript indicating differentiation).

In the absence of labor supply responses, and assuming there to be only a single consumption good, there are two very different ways in which this approach to measuring poverty might be justified. The first is welfarist in the sense that the primitive concern in identifying and quantifying poverty is with individuals' realized levels of welfare. In this view, a household is poor if and only if it fails to achieve some poverty line utility level, $u_z$. With only a single consumption good, $h$'s utility is simply $u(x^h)$, where $u(\cdot)$ is the direct utility function and $x$ denotes consumption of the single good. Because $x^h = y^h$ from $h$'s budget constraint, the condition $u(x^h) \leq u_z$ is equivalent to $y^h \leq z$, where $z = u^{-1}(u_z)$. The poverty index $P$ simply puts a metric on the shortfall of utilities from $u_z$.

The second justification of equation 1 makes no appeal to notions of utility and is in that sense nonwelfarist. It views poverty as the inability to acquire an amount $z$ of the consumption good. The primitive concern is with the potential to consume rather than the well-being derived from doing so.

In the simplest case—which is the case implicitly assumed in much of the literature—the welfarist and nonwelfarist approaches are thus indistinguishable in terms of poverty measurement. The equivalence collapses, however, when labor supply responses are admitted, because the assumption of one good is relaxed. Households acquire the consumption good, at least in part, by forgoing another good, leisure. Some way must then be found to compare deprivation across alternative bundles of consumption $x$ (now thought of as an $N$-vector) and labor supply $L$. The welfarist will use individuals' own preferences to make the comparison; a nonwelfarist may not.

Note that, in a sense, there is nothing special about labor supply here. The same issue—that of deciding how to evaluate the deprivation associated with distinct bundles—would arise if labor supply were fixed but there were two consumption goods. Rather the point is that labor supply makes the issue

1. The terms "individual" and "household" are used synonymously in what follows, the issues raised by the distinction between the two being somewhat removed from the central concerns here.

unavoidable and has a distinctive feature that adds a further layer of complexity. Although it is usually reasonable to assume that consumer and producer prices for consumption goods do not vary across households, wage rates clearly do. Indeed, the tradition of the literature on optimal income tax is to view variation in the wage rate—ability, broadly interpreted—as essentially the only way in which households differ.

To bring out these points concretely, consider first a welfarist approach to the measurement of poverty in the presence of labor supply responses. Starting with a poverty line utility level $u_z$, define the indirect utility function $V(q, w, B)$, giving the maximum utility that can be attained at consumer prices $q = (q_i)$ for the $N$ goods at wage rate $w$ and with lump-sum income $B$. We define the lump-sum income to be exclusive of the value of the individual's endowment of time and, perhaps, goods. Assuming that consumer prices are common across households, the utility achieved by $h$ is thus $V(q, w^h, B^h)$; so $h$ is poor if

$$V(q, w^h, B^h) \leq u_z. \tag{2}$$

To move from utility to income space, follow King (1983) in defining the equivalent income function $y_E(q, q^R, w^h, w^R, B^h)$ by

$$V(q^R, w^R, y_E) = V(q, w^h, B^h). \tag{3}$$

That is, $y_E$ is the lump-sum income at which $h$ would be as well off when facing reference consumer prices $q^R$ and a reference wage $w^R$ as in the situation being evaluated (the latter being described by the consumer prices $q$, wage rate $w^h$, and lump-sum income $B^h$ that $h$ actually faces).[2] The condition (equation 2) for $h$ to be poor is then equivalent to $y_E \leq z$, where $z$ is now defined by $V(q^R, w^R, z) = u_z$, and poverty is naturally measured by

$$P = \int_0^{\hat{w}} D(z, y_E)\, g(w)\, dw \tag{4}$$

where $\hat{w}$ is the poverty wage defined by $z = y_E(q, q^R, \hat{w}, w^R, 0)$, and $g(w)$ is the density of wages. (To avoid a double integral we assume, here and henceforth, that $B^h = 0, \forall h$.) In principle, there is thus no difficulty in developing welfarist measures of poverty in the presence of labor supply responses and household-specific wage rates. Note, though, the element of arbitrariness in the choice of reference prices and, in particular, the reference wage for the evaluation of equivalent incomes.[3] Different choices of reference prices may lead to different rankings of poverty alleviation strategies.

An alternative to developing a welfarist measure of poverty would be to generalize the nonwelfarist approach. As noted in the introduction, there are

2. Use of the equivalent income function is not unproblematic: Blackorby and Donaldson (1988) show that it is not in general concave (in the underlying consumption bundle). In the present context, this is liable to mean, for instance, that transferring commodities from a poor person to a richer person could actually reduce the aggregate poverty gap measured in terms of $y_E$.

3. To be precise, what really matters is the vector of relative reference prices, $q^R/w^R$. Because indirect utility is homogeneous of degree zero in prices and income, it follows from equation 3 that equivalent income measured in units of leisure (that is, $y_E/w^R$) depends on reference prices only through $q^R/w^R$.

many conceivable kinds of nonwelfarism. For example, one might focus on the income that could be earned by working an acceptable number of hours. The implications of this approach for the empirical measurement of poverty have indeed been explored by Garfinkel and Haveman (1977) and by Haveman and Buron (1993). It may well be that their approach—perhaps closer to the capabilities notion—yields qualitatively very different conclusions from that pursued here. Kanbur and Keen (1989) show that the two approaches—of using what they call standard and received income—do indeed have distinct implications for the design of a linear income tax. We pursue only one kind of nonwelfarism here, not because we advocate it, but for brevity and because it seems to capture much of the common tone of policy discussion.

In generalizing the nonwelfarist approach, one starting point is the specification of a particular bundle of consumption and labor supply to act as the reference for evaluating actual bundles. This target $(N + 1)$ – vector $(x^*, L^*)$ is generated not by any reference to utility but by prior views as to what consumption standards households need to attain. To avert deprivation, for instance, households ought to be able to attain a reasonable nutritional intake without an excessive amount of work effort. The question then becomes how to measure the distance between an individual's actual consumption vector $(x, L)$ and the target vector. There are an infinite number of possible metrics.

For simplicity, we consider here only deprivation measures of the form $D[z, y(q, w)]$, where

$$z^h = s_x \cdot x^* - s_L^h L^* \tag{5}$$

and

$$y(q, w^h) = s_x \cdot x(q, w^h) - s_L^h L(q, w^h) \tag{6}$$

where $x(\cdot)$ and $L(\cdot)$ denote the Marshallian commodity demand and labor supply functions. In equation 5, $z^h$ is a poverty line defined as the value, at some shadow prices $S^h = (s_x, s_L^h) > 0$, of the resources needed to attain the target vector. In equation 6, $y(q, w^h)$ is the shadow value of the net resources actually enjoyed by $h$. Assuming that $y$ is strictly increasing in $w$ and that $z^h$ does not increase too rapidly with $w^h$ (typically it would fall), there exists a unique poverty line wage $w^*$ at which $y(q, w^*) = z$, and poverty can be measured as

$$P = \int_0^{w^*} D[z, y(q, w)]\, g(w)\, dw. \tag{7}$$

Consider, for concreteness, the case in which $D(\cdot)$ depends on the poverty gap $z - y(w)$. Deprivation, and hence poverty, can then be assessed by asking, by how much do the resources that $h$ needs to attain the target vector $(x^*, L^*)$ exceed those actually made available to $h$?

The question, then, is how to value these resources—the specification, that is, of the shadow prices $S^h$. There are an infinite number of possible choices, and few natural axioms to invoke. Both technological and ethical considerations arise. Emphasizing the former, a natural approach is to value resources at pro-

ducer prices. This can be done by taking $s_x = p = q - t$, where $p$ denotes producer prices, $t$ denotes the vector of commodity taxes, and (assuming that earned income is untaxed) $s_L^h = w^h$. The value judgments underlying the non-welfarist approach, however, may point toward other shadow prices. It may be, for instance, that some goods are felt to be irrelevant to the achievement of minimal living standards, in which case their appropriate shadow price is zero.

Of more particular importance to our concerns here, policymakers often seem to attach positive weight to the capacity to consume but zero weight to the enjoyment of leisure. Put crudely, policymakers may not care how hard people have to work so long as they are able to sustain a decent level of consumption. Taken to the extreme, such a view corresponds to $s_L^h = 0$. An alternative and very different ethical position is that work effort in excess of the target $L^*$—or, underenjoyment of leisure in relation to some target—should not only be valued positively, but, at a minimum for the least able, should be valued at a shadow wage in excess of the actual wage $w^h$. Taking $s_L^h = w^h$ implies that a given shortfall in leisure hours in relation to the target translates into less deprivation for a low-paid individual than for a high-paid one.

These issues will not be resolved here, but we focus on their implications for the incorporation of labor supply responses into the analysis of poverty alleviation. We simply assume, when defining poverty as in equation 7, that $s_x = p$. That is, we assume that consumption goods are valued at producer prices but for the moment the value placed on leisure, $s_L^h$, is unrestricted. We assume, however, that $s_L^h$ is independent of consumer prices, $q$. This seems a reasonable simplification for our purposes, there being no instantly compelling reason to suppose that the extent of deprivation in excessively hard work depends on the prices of the goods it buys.

## II. Targeting by Commodities

Which kinds of goods should be subsidized by a government seeking to alleviate poverty, and which should be taxed? Specifically, suppose there are two commodities, 1 and 2, that the government can subsidize or tax as it pleases subject only to an overall budget constraint. Labor supply is variable—the implications of this being our central concern—but labor income cannot be taxed. Starting from a position in which neither good is subsidized (or taxed), what is the effect on aggregate poverty, $P$, of introducing a small subsidy on good 2 financed by a tax on good 1?

Besley and Kanbur (1988) address this question in a welfarist context. They effectively assume labor supply responses to be zero, but it is straightforward to show that relaxing this assumption does not affect their central result. Defining poverty in terms of equivalent income (taking the reference wage for each household to be its actual wage), the effect on $P$ of introducing a small tax on good 1 (and hence a small subsidy on good 2) can be shown to be

$$(8) \qquad \frac{\partial P}{\partial t_1} = (\bar{x}_1) \int_0^{\tilde{w}} \left(\frac{x_2}{\bar{x}_2} - \frac{x_1}{\bar{x}_1}\right) D_y\,[z, y_E\,(q, p, w, w)]\, g(w)\, dw$$

where $x_i$ denotes consumption of good $i$ and $\bar{x}_i$ its mean in the population.[4]

Recalling that $D_y < 0$, the interpretation of equation 8 is straightforward: good 2 should be subsidized by a tax on good 1 if and only if consumption of good 2 is more heavily concentrated among the poor than is consumption of good 1 (weighted, by the terms $D_y$, to attach most importance to the most deprived households). Note that the convenient absence of price elasticities reflects the assumption that the starting point is one with no taxes or subsidies. When considering large reforms, considerations of excess burden will also arise, bringing into play elasticities of demand for consumer goods and hence labor supply responses.

The condition expressed in equation 8 is essentially the same form as the result of Besley and Kanbur (1988, equation 26). The only difference is that Besley and Kanbur work with the poverty index of Foster, Greer, and Thorbecke (1984) rather than with the more general form used here. Variable labor supply thus makes no difference to the welfarist analysis. The impact on a household's welfare of a change in the consumer price of good $i$ is simply proportional to its consumption of $i$. The induced effect on the pattern of consumption—and, by the same token, on labor supply—drops out by the envelope property. That is, because the initial level of labor supply is chosen by the individual so as to maximize utility, a small change in that level—as might be induced by the commodity price change associated with the tax/subsidy scheme—will have no (first-order) effect on welfare and hence also no effect on poverty defined in welfarist terms.

Consider now the same problem from the nonwelfarist perspective. It is shown in the appendix that the effect on a poverty index of the form in equation 7 of the revenue-neutral introduction of a small subsidy on good 2 is given by

$$(9) \qquad \frac{\partial P}{\partial t_1} = (\bar{x}_1) \int_0^{w^*} (A_2 - A_1)\, D_y[z^*, y(q, w)]\, g(w)\, dw$$

where

$$(10) \qquad A_i = \frac{x_i + (s_L^h - w^h)\,(\partial L / \partial q_i)}{\bar{x}_i}.$$

Comparing equations 9 and 10 with equation 8, the sole consequence of adopting the nonwelfarist approach[5] to the targeting of subsidies is thus to introduce labor supply considerations in the form of the terms $(s_L^h - w^h)\, \partial L / \partial q_i$.

4. In equation 8, $\tilde{w}$ is a poverty line wage defined by $y_E\,(q, q^R, \tilde{w}, \tilde{w}) = z$. For brevity, the derivation of equation 8 is omitted. It is similar to that of equation 9, which is sketched in the appendix. The critical step is to note (in place of equations A-5 and A-6) that $\partial y_E / \partial q_i = -x_i$ at $q^R = q$ and $w^R = w^h$. The simplicity of equation 8 would not be obtained if the reference wage were specified to be other than $w^h$.

5. Here and elsewhere, "the" nonwelfarist approach refers to the particular variant of nonwelfarism described in the previous section.

There are two cases of interest in which these terms vanish. The first, trivially, is that in which labor supply decisions are unaffected by the consumer prices of the goods being studied ($\partial L/\partial q_i = 0$). In practice, it may often be tempting to assume that this is indeed the case. As an antidote to routinely doing so, however, it is worth recalling that (in the absence of lump-sum income) to assume labor supply to be independent of all consumer prices is to assume it also to be independent of the wage rate. This follows from homogeneity of degree zero of the labor supply function $L(q, w)$. The second case in which these labor supply terms disappear is that in which the leisure component of an individual's deprivation is valued at the individual's wage rate ($s_L^h = w^h$). The intuition for this is that with $s_x = q$ and $s_L^h = w^h$, the impact on deprivation of behavioral responses to the tax reform is being evaluated at the prices actually faced by the consumer; just as those responses can have no effect on the consumer's budget constraint, so too they can have no effect on measured deprivation.

But there are, of course, other possible choices of $s_L^h$. One is to attach no weight at all to labor supply ($s_L^h = 0$). Then equation 10 becomes

$$A_i = \frac{b_i - \epsilon_{Li}}{\bar{b}_i} \tag{11}$$

where $b_i = q_i x_i / wL$ denotes the budget share of good $i$, $\bar{b}_i$ its mean, and $\epsilon_{Li}$ ($= \partial \ln L / \partial \ln q_i$) the elasticity of labor supply with respect to $q_i$. Other things being equal, the case for subsidizing good $i$ is thus weaker the more positive $\epsilon_{Li}$ is, that is, the more such a subsidy would tend to reduce labor supply. This is reminiscent of optimal tax arguments and points to relatively heavy taxation of relatively strong complements with leisure[6]—but the underlying reasoning is very different. With $s_L^h = 0$, the sole object of policy is to push poor individuals' consumption as far toward the target consumption vector ($x_1^*$, $x_2^*$) as possible. There are, broadly, two ways of doing this. The first is to deploy subsidies that enable households to afford greater quantities at any given income. The other is to encourage households to increase their earnings and so to purchase more at any given price. It is on this second front that complementarities between goods and leisure enter the picture. Expanding the consumption of (relative) complements with leisure may be more expensive than expanding the consumption of substitutes. The effects of subsidizing the former are liable to be at least partly offset by an induced reduction in disposable income, whereas subsidizing the latter generates a reinforcing expansion of income.

For reasons discussed above, however, it might be preferable to attach a large positive weight to leisure in measuring deprivation. The implications of doing so are not merely qualitative; they can reverse the conclusions of the previous

6. Some emphasis should be put on the word "relatively." It is not necessary for the arguments here that there exist any good $j$ that is complementary with leisure in the sense that the compensated demand for $j$ falls as $w$ rises; indeed, there may exist no such good. It is the degree of complementarity that is important. It is convenient, for clarity, to speak of taxing or subsidizing complements with leisure. The more delicate and exact formulation of the argument is straightforward, but cumbersome.

paragraph. Suppose, for instance, that $s_L^h$ is at or above the poverty line wage $\hat{w}$. Other things being equal, the goods that ideally should be subsidized to alleviate poverty are thus precisely those for which $\partial L/\partial q_i > 0$. Instincts honed on the optimal tax literature are confounded because complements with leisure should be subsidized rather than taxed. The reason, however, is straightforward. With $s_L^h > 0$, one way of reducing deprivation is by reducing the work effort of the poor. Less effort means less consumption and so a greater shortfall from $x^*$; but with $s_x$ equal to the prices faced by the consumer and $s_L^h > w^h$, the deprivation measure attaches more weight to the hour of leisure gained than it does to the consumption consequently forgone.

Consider, for instance, the appropriate treatment of food. In the developing-country context it could plausibly be argued that greater food intake might enhance the capacity to work. Food and leisure would then be thought of as substitutes: a reduction in the consumer price of food will tend to increase labor supply (so that $\epsilon_{L,\text{food}} < 0$). With $s_L^h = 0$, the implication of the analysis above is that (other things being equal) food should be subsidized, reflecting the secondary benefit of expanded income that such a subsidy induces. If, however, $s_L^h$ is above the poverty line wage, the implication is that food should be taxed. This may at first seem strange, but it has a straightforward explanation. Although taxing food in itself increases deprivation, this may be more than offset by the expanded consumption of both leisure and commodity 2 in the background, the latter now being subsidized by receipts from the food tax.

Pursuing the nonwelfarist approach, rules of thumb for commodity targeting are thus highly sensitive to the weight attached to deprivation of leisure. The issue is troublesome. Having raised it, however, we now put it aside by assuming in the next two sections that $s_L^h = 0$. Because there will also be only one consumption good—commodity subsidies not being at issue—the nonwelfarist deprivation measure in what follows is simply the shortfall of aggregate consumption (equivalently, of net income) from some poverty line.

## III. Targeting by Income

In the absence of incentive effects (and with sufficient resources available for poverty relief) the design of income-based targeting is a trivial exercise. After the poverty line is established, those individuals who are initially below it are given exactly that transfer needed to bring them just above it.[7] Such a scheme involves no leakages. If there are no labor supply or other effects in transferring or raising these resources, and if the informational and administrative requirements can be met without cost, this method gives perfect targeting. But once incentive effects are admitted, the difficulties noted in the introduction arise. Because perfect targeting implies an effective marginal tax rate of 100 percent on those below

7. Bourguignon and Fields (1990) examine the optimal poverty alleviation strategy (in the absence of behavioral responses) when the available budget is insufficient to eliminate poverty.

the poverty line, the poor have no incentive to earn income. Their rational labor supply decisions would then be likely to greatly increase the revenue costs of alleviating their poverty. Incentive effects thus rule out marginal rates of 100 percent on the poor. The questions of precisely how high or low those rates should be, and of how they should vary with income, then become considerably more complex.

There is a large literature, initiated by Mirrlees (1971), that addresses the optimal design of nonlinear income taxes in a welfarist setting. In this work, the issue of incentives for supplying labor is tackled directly by modeling individuals as choosing between work and leisure given the tax-transfer schedule they face. There are assumed to be a large number of individuals, differing only in the pretax wage they can earn. (We relax this homogeneity assumption in the next section.) The government then chooses a schedule that maximizes a social welfare function defined on individuals' welfare, that is, on the utility they derive from their consumption-leisure bundles. As noted in the introduction, however, there is a striking and fundamental dissonance between this welfarist approach and the tone of much policy debate. It is the consequences of reform for the incomes of the poor—the money in their pockets, not something akin to money metric measures of their welfare—that are commonly discussed and analyzed. Kanbur, Keen, and Tuomala (forthcoming) therefore examine the implications of an alternative approach to the design of nonlinear income tax schemes. Besley and Coate (1992) adopt a similar approach in analyzing the case for workfare schemes. In Kanbur, Keen, and Tuomala's approach, the objective of policy is to minimize an income-based poverty index rather than to maximize social welfare. This section reviews their conclusions.

We begin by recalling the main lessons from the welfarist literature on optimal nonlinear income taxation (as reviewed, for instance, in Tuomala 1990). Three general qualitative conclusions emerge:

- The marginal tax rate should everywhere be non-negative.
- The marginal tax rate on the lowest earner should be zero so long as everyone supplies some labor at the optimum.
- The marginal tax rate on the highest earner should be zero so long as wages in the population are bounded above.

The first result is more striking than is commonly recognized. Although it may well be optimal for the average tax rate on the least well off to be negative, it cannot be desirable to subsidize their earnings at the margin. The limitations of the second and third results concerning the endpoints are well known: simulations suggest that zero may be a bad approximation to optimal marginal tax rates in the tails of the distribution, and it can be shown that if it is optimal for some not to work, then the optimal marginal tax rate at the bottom of the income distribution is strictly positive (Tuomala 1990). Nevertheless, these results continue to color professional thinking on issues of rate structure. The lower endpoint result, in particular, has been taken as suggestive in arguing

against very high effective marginal rates on the poor (as, for instance, in Kay and King 1986).

Do these conclusions continue to apply when the objective of policy is not the maximization of social welfare but the minimization of income poverty? The third result certainly does (see Kanbur, Keen, and Tuomala forthcoming for proofs of this and the claims below). As expected, the third result applies because in the context of poverty alleviation the only reason to care about the highest earner—indeed about any of the nonpoor—is as a source of revenue. It is well known that in these circumstances the marginal tax rate on the highest earner should be zero; if it were strictly positive, additional revenue could be extracted by slightly lowering it and thereby inducing the highest earner to earn additional taxable income.

The first and second results, in contrast, are overturned if the objective is to minimize income poverty. If it is optimal for the lowest-ability earner to work, the marginal tax rate at the lower end of the distribution should be strictly negative; that is, a marginal earnings subsidy should be paid to the very poorest. To see why this is optimal from the nonwelfarist perspective even though it cannot possibly be optimal from a welfarist one, consider an initial position in which the individual with the lowest ability works and faces a strictly negative marginal tax rate. Imagine now increasing the marginal tax rate faced by this individual while leaving the average rate at the individual's initial gross income unchanged. The effects of this rotation of the poorest worker's budget constraint through the initial consumption-leisure bundle are that the individual's welfare rises (because if the initial consumption-leisure bundle remains feasible, any change in the individual's behavior must signify an increase in welfare); the individual's net income falls (because the only incentive effect is a substitution toward leisure induced by the higher marginal tax rate); and the government's revenue increases (because the subsidy is paid at a lower rate on a narrower base). From the welfarist perspective, the combination of the utility gain to the individual and the revenue gain to the government makes this reform unambiguously desirable. From the nonwelfarist perspective, however, opposing effects are at work. The revenue gain is desirable, but the net income loss to the poorest worker is not. Minimization of an income-based poverty index will require striking a balance between the two effects, which will make a marginal subsidy on the very poorest optimal.

The possibility of an optimally negative marginal tax rate is confined, however, to the poorest of the poor. For those who find themselves exactly at the poverty line, the optimal marginal rate can be shown to be strictly positive.

These qualitative implications of the nonwelfarist approach thus point to a pattern of marginal tax rates below the poverty line that is both complex and potentially very different from that suggested by the welfarist tradition. But how far do low or even negative marginal tax rates on the very poorest individuals extend into the range of incomes? And how is the poverty-minimizing rate

structure affected by the precise location of the poverty line $z$ and by the form of the deprivation function $D(\cdot)$?

Table 1 reports results of simulations intended to address these concerns. The results assume Cobb-Douglas preferences

$$u(x, L) = (1 - \delta)\ln(x) + \delta\ln(1 - L) \qquad \delta \in (0, 1) \tag{12}$$

(the time endowment being normalized at unity) with $\delta = 1/2$, $\ln(w)$ normally distributed (with mean $-1$ and standard deviation 0.39), and that the revenue requirement is about 10 percent of gross income. These are the standard assumptions in simulations of this sort. The novelty is in the form of the objective function, for which we take a poverty index of the form developed by Foster, Greer, and Thorbecke (1984):

$$P^{\alpha} = \int_0^{w^*} \left(\frac{x(w) - z}{z}\right)^{\alpha} g(w)\, dw \qquad \alpha > 1 \tag{13}$$

Table 1. *Simulation Results for Optimal Average and Marginal Tax Rates at Various Percentiles of the Wage Distribution*

| *Poverty line and percentile of the wage distribution* | | *Average tax rate* | *Marginal tax rate* |
|---|---|---|---|
| *a. Low poverty line* | | | |
| Low[a] | 0.06 | −100 | 69 |
| Poverty line | 0.31 | −3 | 62 |
| | 0.50 | 12 | 53 |
| | 0.90 | 29 | 35 |
| High | 0.99 | 29 | 23 |
| *b. Middle poverty line* | | | |
| Low[a] | 0.02 | −100 | 63 |
| Poverty line | 0.43 | 0 | 54 |
| | 0.50 | 9 | 53 |
| | 0.90 | 27 | 34 |
| High | 0.99 | 27 | 17 |
| *c. High poverty line* | | | |
| Low[a] | 0.003 | −87 | 56 |
| | 0.50 | 8 | 54 |
| Poverty line | 0.56 | 16 | 48 |
| | 0.90 | 26 | 34 |
| High | 0.99 | 26 | 17 |
| *d. Maximin*[b] | | | |
| Low[a] | 0.16 | −100 | 73 |
| | 0.50 | 17 | 53 |
| | 0.90 | 32 | 35 |
| High | 0.99 | 32 | 26 |

*Note*: In all four groups, the ratio of aggregate consumption to aggregate output is 0.9. In groups a, b, and c, the parameter for aversion to inequality among the poor, $\alpha$, is 2. In groups a, b, and c, the minimum level of consumption, $x(n_0)$, is 0.06; in d it is 0.07.

a. The percentile of the wage distribution below which individuals choose not to work.

b. Assumes infinite aversion to inequality among the poor ($\alpha = \infty$).

*Source*: Authors' calculations.

where $x(w)$ denotes the consumption of an individual with wage $w$. The parameter $\alpha$ in equation 13 provides a convenient parameterization of alternative degrees of aversion to inequality among the poor.

One immediate implication of this specification should be noted. With Cobb-Douglas preferences (so that the marginal rate of substitution between consumption and work is strictly positive at zero hours) and a lognormal wage distribution (so that the lower bound of $w$ is zero), there are some who will work only if the marginal tax rate at the bottom of the distribution is infinitely negative. In both the welfarist context and that of income poverty minimization, it would be optimal to have some of the population idle. As noted above, in the welfarist case the optimal marginal rate at the bottom of the income distribution is then strictly positive. However, for the case in which the objective is to minimize income poverty and some households are idle at the optimum, we have been unable to derive any general result on the sign of the optimal marginal rate at the lower endpoint. The simulations provide some indication of the extent to which the argument for nonpositive marginal rates at the lower end (when the poorest work) continues to exert some force when instead the wage distribution is not bounded away from zero. Table 1 gives optimal average and marginal tax rates at various percentiles of the wage distribution, starting at the bottom and including the point at which the assumed poverty line is to be found. The first three groups (a, b, and c) all take $\alpha = 2$ and differ in taking successively higher poverty lines. The last (d) looks at the maximin case, which corresponds to $\alpha = \infty$.

Several features stand out in the table. First, the marginal rate on the lowest gross income—which, as just noted, we are unable to sign in principle—emerges as very strongly positive: not only is it not negative, it is not even low. Second, marginal tax rates decline monotonically from the poorest to the richest individual, implying that the dictates of effective targeting can run exactly counter to the popular notion that equity concerns require the marginal tax rate to increase with income. Such declining marginal tax rates run counter to the conclusion sometimes drawn from the welfarist literature that the administrative advantages of linear taxation can be bought at relatively little loss in terms of policy effectiveness.

The third feature of the table is that (comparing a and c) increases in the poverty line reduce optimal marginal rates at and below the poverty line. The intuition for this seems to be that the case for low marginal rates intended to encourage those at or near the poverty line to move out of poverty, becomes stronger as the poverty line moves into denser parts of the distribution. Fourth, comparing the maximin case with the others, increases in the extent of aversion to inequality among the poor tend to increase the marginal rates that they optimally face. Other simulations (not reported here) suggest that moderate variation in the revenue requirement affects the general level of marginal tax rates (which tend to increase with the revenue required) but not the qualitative pattern of their variation with income. This is perhaps as would be expected

because the greater the concern with alleviating poverty, the more attractive schemes that approach minimum income guarantees are likely to be: the emphasis is then on raising the consumption of the very poorest, and financing the transfer this requires calls for relatively high marginal tax rates in the lower part of the distribution in order to impose sufficiently high average tax rates further up the distribution.

But perhaps the most important feature of the results is the finding of marginal tax rates on the poor that are invariably rather high (bearing in mind the fairly minimal revenue requirement). In most cases marginal rates on the bulk of the poor exceed 60 percent, and in all cases they exceed 50 percent. The case for low marginal tax rates to encourage the poor to help themselves thus is less discernible in the simulations than expected. Even with the relatively elastic labor supply responses implicit in Cobb-Douglas preferences (the elasticity of substitution between consumption and leisure being unity), a stronger mark is left by the case for high marginal rates associated with the unattainable ideal of perfect targeting described at the start of this section. Simulations for the case in which the elasticity of substitution is 0.5 (reported in Kanbur, Keen, and Tuomala forthcoming) confirm this impression.

The optimal marginal tax rates that emerge from these simulations are not necessarily higher in the nonwelfarist case than in the welfarist one. Indeed, it is not clear that a coherent comparison between the two approaches can be made because the latter, but not the former, depends on the cardinal representation of preferences. The safest conclusion—albeit a provisional one, because our simulations are inevitably only special cases—seems to be that a concern with income poverty does not in itself provide a strong case for marginal tax rates on the bulk of the poor that are substantially lower than expected from the perspective of the welfarist tradition. The reason for this, it seems, is that shifting from the welfarist to the nonwelfarist perspective introduces two considerations that point in opposite directions. First, the case for lower marginal tax rates on the poor is strengthened by the prospect of inducing them to raise their own incomes. The nonwelfarist view attaches no weight to the leisure that the poor forgo; this underlies the result that a marginal earnings subsidy on the very poorest is optimal when that individual works. Second, the case for lower marginal tax rates on the poor is weakened by the need to support the incomes of the poor, rather than their welfare, which could be "bought" by allowing them a relatively high amount of leisure: supporting the incomes of the poor calls for relatively high marginal tax rates in the lower part of the income distribution, and the revenue needed for this support requires that sufficiently high average tax rates be imposed on higher incomes. The simulations suggest that these two opposing effects broadly offset one another.

## IV. Targeting by Indicators

In the analysis so far, individuals have been assumed to differ only in their unobserved ability. It is now widely recognized, however, that there are poten-

tially severe incentive and other costs of administering income-related transfers. One way of overcoming these costs, particularly in developing countries, is to differentiate the population by easily observable indicators that are correlated with the unobservable characteristic of interest. An individual's labor market status or demographic attributes, for instance, may convey information on underlying ability. Transfers can usefully be made contingent upon such characteristics. The theory of the optimal use of such information was first considered by Akerlof (1978) and developed by, among others, Kanbur (1987), Besley and Kanbur (1988), and Ravallion (1987). But most of the simple rules of thumb for targeting that have been developed simply assume away labor supply effects. An exception is Kanbur and Keen (1989), who develop a relatively simple framework that gives some feel for the optimal use of nonincome information in the presence of incentive effects. This section reports on that work.

Suppose the population can be divided into two mutually exclusive and exhaustive groups, A and B. The underlying contingencies are assumed to be absolute, so households are unable to switch between groups. The contingency is costlessly verifiable, but we assume—to keep matters simple—that only linear income taxation is feasible. What makes the problem interesting is that distinct schedules may be applied to the two groups: they may be faced, that is, with different poll subsidies $G_K$ and with different marginal tax rates $t_K$ (for $K = A$, $B$). This ability to treat the two groups differently is only valuable, of course, if they differ in some way that is relevant for poverty alleviation. We allow them to differ in two respects. First, the within-group wage distributions $g_K(w)$ may differ. Thus one group may, for instance, be systematically poorer than the other. Second, they may differ in the responsiveness of their labor supply behavior.

Specifically, we assume that although all individuals have Cobb-Douglas preferences, as in equation 12, the parameter $\delta$ may differ across the two groups. Imposing the further restriction, for definiteness, that poverty is to be assessed in terms of the Foster-Greer-Thorbecke index, the objective of policy is thus taken to be the minimization (subject to the government's budget constraint) of

$$P^{\alpha} = \Theta\, P_A^{\alpha} + (1 - \Theta)\, P_B^{\alpha} \tag{14}$$

where $P_K^{\alpha}$ is defined as in equation 13, $\theta$ is the proportion of the population in group A, and the net income of a type $K$ household with pretax wage $w$ is

$$x(w; K) = (1 - \delta_K)\,[(1 - t_K)\, w + G_K]. \tag{15}$$

The two groups are assumed to have the same poverty line, $z$. This precludes a range of (troubling) issues concerning the relation between needs and optimally targeted benefits. Depending on the form of the deprivation function, $D(\cdot)$, it may be, for example, that the level of support optimally targeted to a group varies inversely with its neediness, as measured by $z$. The intuition is that the very needy may simply be too expensive to help (see Keen 1992).

Taking the tax rates $t_K$ as given, under what circumstances would aggregate poverty $P^\alpha$ be reduced by cutting the poll subsidy given to one group in order to finance an increase in that paid to the other? A retargeting of support of this kind away from group B and toward group A can be shown to reduce aggregate poverty if and only if[8]

$$\sigma(\delta_A, t_A)\, P_A^{\alpha-1} > \sigma(\sigma_B, t_B)\, P_B^{\alpha-1} \tag{16}$$

where

$$\sigma(\delta, t) = \frac{(1-\delta)(1-t)}{1-t(1-\delta)}. \tag{17}$$

To develop the intuition behind inequality 16, consider first the role of the $P_K^{\alpha-1}$ terms. These emphasize the simple but important point that the reduction of aggregate poverty measured in some particular way is typically not best pursued by redirecting resources toward whichever group is poorest in terms of that same measure. What matters is the marginal effect on the measure of interest. The structure of the $P^\alpha$ index happens to be such that the implied rule takes an especially simple form. Assuming away incentive effects for the moment, so that $\delta_A = \delta_B = 0$, minimization of the aggregate index for some specific choice of $\alpha$ requires looking first at the within-group indexes for $\alpha - 1$. Suppose, for instance, that we have chosen $\alpha = 1$. This means that our objective is simply to minimize the aggregate poverty gap, or, equivalently, to maximize the net income of the poor. Imagine now that we have some fixed sum to spend on increasing the poll subsidy $G_K$, to one group or the other (and assume for simplicity that $\theta = 1/2$). Which group should we favor? The disadvantage of having to spend this money as a poll subsidy is that some of it will be wasted on the nonpoor; giving it only to group K, the proportion of our fixed sum that will reach the poor is just the proportion of that group that is in poverty. To achieve the largest possible increase in the total income of the poor, we should therefore allocate the funds to whichever group has the larger number of poor individuals—that is, to whichever group has the higher $P_K^0$.

Incentive effects enter the story through the $\sigma(\cdot)$ terms in inequality 16, with retargeting toward group A more likely to be desirable, other things being equal, the higher $\sigma(\delta_A, t_A)$ is and the lower $\sigma(\delta_B, t_B)$ is. It is easily seen from equation 17 that $\sigma(\delta, t)$ is decreasing in both $\delta$ and $t$. Thus group A is more likely to be favored the less responsive its labor supply behavior is and the lower the marginal tax rate is that it initially faces. The intuition is straightforward. When $\delta_A$ is relatively low, the income effect of increasing the poll subsidy to group A—which points toward a reduction in hours and hence in net income, dampening the beneficial impact on poverty—is relatively weak. Conversely, a high $\delta_B$ indicates a relatively powerful income effect acting to mitigate the impact of reduc-

8. Proofs of the claims that follow are in Kanbur and Keen (1989).

ing the poll subsidy to group B. And when $t_A$ is relatively low, so too is the revenue cost of the reduction in hours worked—and hence taxes paid—by members of group A as a result of their higher lump-sum income. Conversely, a high $t_B$ is helpful in recouping revenue from the increased labor supply of group B.

The tension to which inequality 16 points emerges especially clearly if the initial position is one in which $t_A = t_B = 0$. Retargeting toward A is then desirable if and only if

$$(1 - \delta_A) P_A^{\alpha-1} > (1 - \delta_B) P_B^{\alpha-1}. \tag{18}$$

We have already discussed why the natural inclination to favor the group with the higher incidence of poverty has to be modified to favor the group with higher $P^{\alpha-1}$. But incentive effects can more than offset this consideration. It may be optimal to cut the poll subsidy paid to the group with the higher $P^{\alpha-1}$ if its labor supply behavior is sufficiently more sensitive (that is, if $\delta$ for that group is sufficiently high).

The targeting rule expressed in inequality 16 is valid for arbitrary marginal tax rates $t_K$. When these, too, can be chosen by the government, from the associated first-order conditions,[9] poverty minimization requires that

$$\frac{P_A^{\alpha-1} - P_A^{\alpha}}{P_B^{\alpha-1} - P_B^{\alpha}} = \frac{\bar{x}_A}{\bar{x}_B} \tag{19}$$

where $\bar{x}_K$ denotes the mean net income of group K. The difference between the $P^{\alpha}$ and $P^{\alpha-1}$ indexes must thus stand in the same ratio across groups as do their mean net incomes. For the case in which $\alpha = 1$, this reduces to the simple condition that

$$\Gamma_A = \Gamma_B \tag{20}$$

where $\Gamma_K$ denotes the share of the poor in group K of the total net income of that group. The significance of the rules expressed in equations 19 and 20 is less in the additional insight they convey—which adds little to what has gone before—than in their applicability. They show how simplifying assumptions can be used to incorporate labor supply responses, in a relatively straightforward way, into the use of indicators for targeting.

## V. Conclusion

Labor supply introduces some new considerations into the design of poverty alleviation programs. First and foremost, it forces us to reconsider the standard objective function according to which these programs are evaluated: the minimization of poverty as measured by the shortfall of income or expenditure from a critical value. This objective leaves out of consideration the effort that individuals make in earning their incomes. How is this effort, or rather the leisure that is lost in making it, to be valued? Valuing it at the market wage—which is the

9. Details of the proof can be found in Kanbur and Keen (1989).

welfarist approach, because this is how individuals would value it themselves—has the unappealing feature that the effort of less able individuals is valued less. Many poor men and women perform backbreaking labor to earn a meager living, and that surely should be recognized. The conceptual issues are not easily resolved, and here we have done no more than make a start. We feel that nonwelfarist perspectives have special interest where labor supply is concerned, but in this article we have restricted ourselves to examining the consequences of just one particular—and particularly convenient—approach within the broad class of nonwelfarist approaches.

A good example of how the new perspective can alter basic results in the targeting literature is provided in the section on targeting by commodities. Besley and Kanbur (1988) establish, under certain conditions, the validity of the simple rule of thumb that commodity subsidies should focus on those commodities whose consumption by the poor is a large fraction of total consumption. This is done in a welfaristic framework. However, if labor supply is elastic, then, under the nonwelfarist approach considered here, the rule is modified depending on the weight given to disutility of effort in evaluating poverty. If no weight is given at all, then the case for subsidizing good $i$ is weaker the more such a subsidy would tend to reduce labor supply, that is, the greater the complementarity between $i$ and leisure. But this result is reversed by attaching a sufficiently high weight to the disutility of effort. It is then no longer acceptable to provide consumption at the poverty line by inducing individuals to work excessively, and complements with leisure should therefore be subsidized rather than taxed.

The rest of the article followed through the consequences of assuming that no weight is given to leisure in the social welfare function. For income-based targeting (and for conventional parameter values), the optimal marginal withdrawal of benefits as income increases is around 50 to 60 percent. This should provide a benchmark for the evaluation of income-tested schemes. Marginal withdrawal rates far above this may look good from the simplest targeting perspective, but the incentive effects are liable to dominate any targeting gains. Finally, we considered modifications to rules of thumb in non-income-based targeting. We showed that for any indicator that divides the population into mutually exclusive groups for targeting purposes, positive correlation between labor supply elasticity and poverty incidence across the groups reduces the usefulness of the indicator. Thus, relying only on poverty incidence can give a false sense of the value of an indicator for targeting purposes.

This article is only a start in the direction of introducing labor supply considerations into the targeting of poverty alleviation programs in developing countries. We end by noting that the issues raised here extend well beyond the specific case of labor supply and income poverty. They apply to any measure of the standard of living (such as nutrition) when individuals have choices to make between alternative forms of consumption and differ in their ability to transform one type of consumption into another.

## Appendix. The Derivation of Equation 9

We describe here the derivation of equation 9. Using the government's budget constraint

$$R = t_1 \int_0^\infty x_1(q, w)\, g(w)\, dw + t_2 \int_0^\infty x_2(q, w)\, g(w)\, dw \tag{A-1}$$

to define $t_2$ as a function of $t_1$ (for fixed $R$), the effect on poverty—defined as in equation 7—of slightly increasing the tax on good 1 in order to lower that on good 2 is given by

$$\frac{dP}{dt_1} = \frac{\partial P}{\partial t_1} + \frac{\partial P}{\partial t_2}\frac{dt_2}{dt_1} \mid R. \tag{A-2}$$

Differentiating A-1 at $t_1 = t_2 = 0$ gives

$$\frac{dt_2}{dt_1} \mid R = -\frac{\bar{x}_1}{\bar{x}_2} \tag{A-3}$$

and from equation 7, and assuming that $D(z, z) = 0$,

$$\frac{\partial P}{\partial t_i} = Dy \frac{\partial y}{\partial q_i} g(w)\, dw. \tag{A-4}$$

Assuming constant producer prices, $p$, equation 6 implies that

$$\frac{\partial y}{\partial q_i} = s_x \cdot \frac{\partial x}{\partial q_i} - s_L^p \frac{\partial L}{\partial q_i} \tag{A-5}$$

$$= -\left[x_i + (s_L^p - w)\frac{\partial L}{\partial q_i}\right]. \tag{A-6}$$

Equation A-6 follows from the choice $s_x = p$, the assumption that both taxes are initially zero (so that $p = q$), and differentiation of the individual's budget constraint, $q \cdot x - wL = 0$. Substituting equations A-3 to A-6 in equation A-1 gives equation 9.

## References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Akerlof, G. A. 1978. "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review* 68:8–19.

Anand, Sudhir, and Ravi Kanbur. 1991. "Public Policy and Basic Needs Provision: Intervention and Achievement in Sri Lanka." In Jean Dreze and Amartya Sen, eds., *The Political Economy of Hunger*. Vol. 3. Oxford: Clarendon Press.

Atkinson, A. B. 1987. "On the Measurement of Poverty." *Econometrica* 55(July):749–64.

———. 1992. "On Targeting Social Security: The Experience of Western Countries with Family Benefits." London School of Economics. Mimeo.

Besley, Timothy. 1993. "The Principles of Targeting." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor.* World Bank Regional and Sectoral Studies. Washington, D.C.

Besley, Timothy, and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programmes." *American Economic Review* 82:249–61.

Besley, Timothy, and Ravi Kanbur. 1988. "Food Subsidies and Poverty Alleviation." *Economic Journal* 92(September):701–19.

Blackorby, Charles, and David Donaldson. 1988. "Money Metric Utility: A Harmless Normalization?" *Journal of Economic Theory* 46:120–29.

Bourguignon, François, and G. S. Fields. 1990. "Poverty Measures and Anti-Poverty Policy." *Recherches Economiques de Louvain* 56:409–27.

Dilnot, Andrew, and Graham Stark. 1989. "The Poverty Trap, Tax Cuts, and the Reform of Social Security." In Andrew Dilnot and Ian Walker, eds., *The Economics of Social Security.* Oxford: Clarendon Press.

Foster, James, Joel Greer, and Erik Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52:761–66.

Garfinkel, Irvin, and Robert Haveman. 1977. "Earnings Capacity, Economic Status, and Poverty." *Journal of Human Resources* 12:49–70.

Gilbert, Martin. 1991. *Churchill: A Life.* London: Minerva.

Haveman, Robert, and L. F. Buron. 1993. "Escaping Poverty through Work: The Problem of Low Earnings in the United States, 1981–90." *Review of Income and Wealth* Series 39:141–57.

Kanbur, Ravi. 1987. "Transfers, Targeting, and Poverty." *Economic Policy* 4:112–36, 141–47.

Kanbur, Ravi, and M. J. Keen. 1989. "Poverty, Incentives, and Linear Income Taxation." In Andrew Dilnot and Ian Walker, eds., *The Economics of Social Security.* Oxford: Clarendon Press.

Kanbur, Ravi, M. J. Keen, and Matti Tuomala. Forthcoming. "Optimal Non-linear Income Taxation for the Alleviation of Income Poverty." *European Economic Review.*

Kay, J. A., and M. A. King. 1986. *The British Tax System,* 4th ed. Oxford: Oxford University Press.

Keen, M. J. 1992. "Needs and Targeting." *Economic Journal* 102(January):67–79.

King, M. A. 1983. "Welfare Analysis of Tax Reforms Using Household Data." *Journal of Public Economics* 21(July):183–215.

Mirrlees, J. A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38:175–208.

Ravallion, Martin. 1987. "Land-Contingent Poverty Alleviation Schemes." *World Development* 17(August):1223–33.

———. 1993. *Poverty Comparisons.* New York: Harwood Academic Press.

Sahn, David, and Harold Alderman. 1992. "The Effect of Food Subsidies on Labor Supply in Sri Lanka." Presented at World Bank conference on Public Expenditures and the Poor: Incidence and Targeting, Washington, D.C. Processed.

Sen, A. K. 1976. "Poverty: An Ordinal Approach to Measurement." *Econometrica* 44:219–31.

———. 1985. *Commodities and Capabilities.* Amsterdam: North-Holland.

Stern, N. H. 1987. "'Comment' on Kanbur (1987)." *Economic Policy* 4:136–41.

Tuomala, Matti. 1990. *Optimal Income Tax and Redistribution.* Oxford: Clarendon Press.

Ulph, David. 1991. "Optimal Income Taxation: Resourcism vs Welfarism." University of Bristol, Department of Economics, United Kingdom. Processed.

# Dual Exchange Rates in Europe and Latin America

Nancy P. Marion

*This article uncovers some important empirical regularities surrounding the operation of formal dual exchange rates in Europe and Latin America in the 1970s and 1980s. Although there are parallels between the European and Latin American experiences, there are also interesting differences in terms of the size and nature of the distortion created by two official exchange rates; the response of the distortion to foreign interest rates, real commercial exchange rates, and domestic budget deficits; and the motives for adopting this exchange rate regime. Empirical work on dual exchange rate regimes is made difficult by the transitory nature of these regimes and by frequent changes in institutional practices.*

Countries adopt dual exchange rates in an attempt to insulate their economies from the effects of international capital movements. In the early 1970s, some European countries with fixed exchange rates faced large, speculative capital flows that brought about unwanted fluctuations in international reserves. The European authorities were reluctant to switch to flexible exchange rates, however, fearing large exchange rate movements with uncertain effects on trade. Dual exchange rates were adopted as a temporary middle ground between the extremes of fixed and flexible rates. Dual exchange rates involved the formal establishment of separate exchange markets, with separate exchange rates, for current account and capital account transactions. In the 1970s and 1980s, many developing countries used dual exchange rates to reduce the pressure from capital flight on their fixed exchange rate, thus avoiding the adverse inflationary consequences of an across-the-board devaluation. The arrangement was used as an interim measure prior to devaluing the currencies.

Dual exchange markets can be set up in a variety of ways. The authorities may establish a fixed commercial exchange rate for current account transactions and a floating financial rate for capital account transactions. Belgium and France operated their exchange markets in this way. Alternatively, the authorities may

Nancy P. Marion is with the Department of Economics at Dartmouth College. This article is an outgrowth of the World Bank's project on Macroeconomic Aspects of Multiple and Black Exchange Markets. The author thanks project organizers Miguel Kiguel, Saul Lizondo, and Steve O'Connell; Robert Flood for helpful discussions; Joshua Aizenman, Pablo Guidotti, Mike Knetter, Carsten Kowalczyk, Andrew Oswald, and Alex Zanello for comments on an earlier draft; and John Dean, Murtaza Moochhala, and Rodolfo Luzio-Antezana for research assistance.

allow some current account transactions to take place at a fixed or managed rate, with the remaining current account transactions and all capital account transactions taking place at a floating rate. Argentina and Mexico, for example, used this version. Although it is common to fix the commercial exchange rate and let the financial exchange rate float freely, other intervention strategies have also been followed. Italy allowed both the commercial and the financial exchange rates to float freely. El Salvador and Jamaica had periods when both rates were fixed. Not all countries officially sanction both exchange markets. Before adopting two official exchange rates, the Dominican Republic adopted a more informal approach, permitting some commercial transactions to take place in an officially sanctioned market and relegating other transactions, mostly financial, to an unofficial parallel market. Some countries, such as Venezuela, establish more than two official exchange rates, thus combining aspects of dual exchange markets and a system of multiple exchange rates.

The European experience has inspired a large theoretical literature but scant empirical work. In addition, little is known about the more recent dual exchange rate episodes in developing countries. The purpose of this article is to uncover some empirical regularities surrounding the experience with dual exchange markets in Europe and Latin America.[1] The analysis highlights some important similarities and differences between the European and Latin American episodes that enhance understanding of dual exchange markets in practice.

Section I sets out a theoretical framework that describes the distortion created by dual exchange rates and examines the magnitude of this distortion in several countries. Section II identifies some of the economic factors that influence the distortion and looks at the empirical evidence. Section III confronts some serious issues that may compromise empirical work on dual exchange markets, and section IV provides concluding remarks.

## I. A Theoretical Framework

The defining feature of a dual exchange market is that current account and capital account transactions are channeled into separate exchange markets—a commercial exchange market for current account transactions and a financial exchange market for capital account transactions (Fleming 1971; Lanyi 1975). Foreign exchange may stand at a premium or a discount in the financial exchange market compared with its price in the commercial exchange market. The classic dual rate system (DRS) introduces a distortion into asset portfolios. The distortion is created by the spread between the two exchange rates and its evolution over time.

1. Developing countries outside Latin America have also experimented with dual exchange rates. For example, Egypt, Iran, Nigeria, South Africa, Sudan, Syria, Uganda, Zaire, and the francophone African countries, among others, have tried dual exchange rates at various times.

### *The Distortion Created by Dual Rates*

To illustrate the distortion created by dual rates, consider the case in which a resident of a country that operates a DRS wishes to hold a foreign currency–denominated asset for one period. The resident must purchase capital-account-eligible foreign exchange at the financial exchange rate, the applicable rate for financial transactions. In the next period, the resident repatriates the principal at the next period's financial exchange rate and repatriates the interest income at the next period's commercial exchange rate, the rate applicable for current account transactions. This period's expected nominal return from holding the foreign currency–denominated security for one period, $N$, is therefore

$$N = i^* \left(\frac{1}{z}\right)(1 + \delta) + f \tag{1}$$

where $i^*$ is the interest rate on the foreign currency–denominated asset; $z$ is the spread, which is the ratio of the home country price of foreign currency in the spot financial market to that in the spot commercial market; $\delta$ is the expected rate of depreciation of the commercial exchange rate; and $f$ is the expected rate of depreciation of the financial exchange rate.

Equation 1 shows that the dual exchange market distorts the effective rate of return by means of the spread and the expected rates of change of the financial and commercial exchange rates. By distorting the effective rate of return, the DRS influences international capital flows. By implication, it also influences the individual's intertemporal allocation of consumption (and of output, if investment is allowed for). The DRS may provide insulation from foreign interest rate shocks if such shocks generate offsetting movements in the spread or in the expected rates of change in exchange rates.

The nature of the distortion is altered by the perceived temporariness of the dual exchange market. Dual exchange rates are usually adopted to ease the transition from a unified exchange rate peg to another kind of exchange rate regime or a different unified peg. For example, France and Italy adopted dual exchange rates for thirty-one months and fifteen months, respectively, as a transition from the unified fixed exchange rate regime under Bretton Woods to the more flexible arrangements of the post–Bretton Woods era. In the 1980s, Bolivia, Costa Rica, and Jamaica all used the arrangement for less than thirty-six months as a transition from one unified fixed exchange rate regime to another.

The domestic rate of return described by equation 1 was derived assuming that the dual exchange market would be in place when repatriation occurs. If, instead, the investor believes the dual market is temporary, the calculation of the return must take into account not only the possible exchange rate regimes to follow but also the probability of each being in effect at the time of repatriation. Clearly, the perceived temporariness of the dual exchange market combined with uncertainty about the regime to follow can alter the nature of the distortion, especially as the date of reunification approaches and probability estimates are revised (Flood and Marion 1982).

Table 1. *Descriptive Features of the Latin American DRS in the 1980s*

| *Feature* | *Argentina* 1981 | *Argentina* 1982 | *Bolivia* | *Costa Rica* | *Dominican Republic* | *El Salvador* | *Guatemala* | *Jamaica* | *Mexico* |
|---|---|---|---|---|---|---|---|---|---|
| Adoption date | 6/22/81 | 7/6/82 | 3/24/82 | 3/9/81 | 8/24/82 | 8/9/82 | 11/16/84 | 1/10/83 | 8/5/82 |
| Reunification date | 12/24/81 | 10/30/82 | 11/5/82 | 11/11/83 | 1/23/85 | 1/22/86 | 6/23/88 | 11/24/83 | 2/88 |
| Duration (months) | 6 | 4 | 7 | 32 | 29 | 41 | 43 | 10 | 66 |
| Pre-DRS regime | Crawl | Controlled float | Dollar peg | Dollar peg | Dollar peg | Dollar peg | Dollar peg | Dollar peg | Controlled float |
| Post-DRS regime | Controlled float | Controlled float | Dollar peg | Dollar peg | Controlled float | Dollar peg | Dollar peg | Dollar peg (1 month) controlled float | Controlled float |
| Price on secondary foreign exchange market | Premium | Premium | Premium | Premium (brief discount) | Premium | Premium | Premium | Premium | Premium (brief discount) |
| Devaluation of home currency in principal market | | | | | | | | | |
| At adoption | Yes | Yes | No | Yes | No | No | No | No | Yes |
| During DRS | Yes | Yes | No | Yes | No | No | Yes | No | Yes |
| At reunification | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Use of additional exchange rates or mixed rates | Mixed | Mixed | Mixed | Additional and mixed | Additional | Mixed | Additional and mixed | Additional | Additional |
| Use of additional controls on capital | Yes | Yes | Yes | Yes | Yes | Yes | Few | Yes | Few |
| Reclassification of transactions during DRS | No | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Size of spread in month preceding reunification (percent) | 38 | 20 | 384 | 3.6 | 206 | 93 | 3.6 | 66 | 2 |
| Management of principal rate | Controlled float (mini devaluation) | Controlled float (mini devaluation) | Peg | Peg, then float | Peg | Peg | Peg | Peg | Controlled float (mini devaluation) |
| Management of secondary rate | Free | Free, peg | Free | Managed, peg | Controlled float | Managed, peg | Free, controlled float | Free, peg | Free, crawl |

*Note*: DRS denotes dual rate system, or a dual exchange rate market. In the Mexican and Guatemalan cases, reunification is de facto, not de jure.
*Source*: IMF (various years a); IMF (various issues b); International Currency Analysis, Inc. (various issues).

Although theoretical work on dual exchange markets often assumes a complete segmentation of the two exchange markets, in practice this is not possible (Bhandari and Decaluwe 1987; Gros 1988). When the spread widens, agents have an increased incentive to channel transactions through the market with the more attractive rate. The authorities have sometimes tried to discourage fraudulent leakages across markets by reclassifying transactions in such a way as to narrow the spread. Moreover, current account transactions that are difficult to monitor, such as tourist expenditures and remittances, are often assigned to the financial exchange market from the start. In most developing countries, reclassification of transactions occurs frequently. These cross-market leakages change the nature of the distortion facing asset holders because only a fraction of foreign assets is purchased or repatriated at the financial rate and only a fraction of interest income is repatriated at the commercial rate. The distortion still depends on the spread and rates of exchange rate change, but in a much more complicated way.

Cross-market leakages also distort relative prices of goods and services. If $P^*$ is the foreign price level, $SP^*$ is the domestic price of goods and services channeled through the commercial exchange market, and $XP^*$ is the domestic price of goods and services channeled through the financial exchange market, then the spread between the financial and commercial exchange rates measures the distortion in the relative price of goods bought and sold in the financial market.

### *A Look at the Data*

Among European countries, Belgium (actually the Belgium-Luxembourg Economic Union, or BLEU), France, and Italy were the major countries that used dual exchange rates in the early 1970s. (Belgium adopted its system in 1957 and kept it in place until 1990.) The United Kingdom and the Netherlands used a second exchange rate for a small group of capital account transactions, as did France before adopting its full-fledged dual rate system and again in the 1980s. Because of the limited availability of financial exchange rate data for these more minor episodes, this analysis focuses on the dual exchange market in Belgium between 1963 and 1988, in France between 1971 and 1974, and in Italy between 1973 and 1974.[2]

The list of Latin American countries that operated formal dual exchange rates is more extensive. In the 1970s, dual rates were used by Argentina, Chile,

2. Belgium operated a dual exchange market from 1957 to 1990. France used a dual exchange market from August 23, 1971, to March 21, 1974. France operated a devise titre, a second exchange rate applicable to resident purchases and sales of foreign securities, from August 11, 1969, to October 20, 1971, and again from May 21, 1981, to May 22, 1986. Italy operated a dual exchange market from January 22, 1973, to March 22, 1974. The Netherlands established an O-guilder market for nonresidents investing in guilder bonds between September 6, 1971, and February 1, 1974. The United Kingdom operated a separate investment currency exchange rate for certain capital account transactions conducted by residents in the United Kingdom from 1947 until October 23, 1979.

Ecuador, Jamaica, Nicaragua, Paraguay, Peru, and Uruguay. In the 1980s, formal dual exchange rates were also used by several countries. Table 1 provides details. Mexico operated such a system during much of the 1980s, and Argentina adopted dual rates for a period in 1981 and again in 1982. Bolivia, Costa Rica, the Dominican Republic, El Salvador, Guatemala, and Jamaica also used dual exchange rates for a time. A number of Latin American countries had informal dual exchange markets as well.

The spread between the two exchange rates represents the observable part of the distortion created by the dual exchange market. Some sense of the magnitude of the distortion is revealed by the data on spreads for a set of countries that have operated dual exchange markets. Table 2 summarizes key statistics on the spreads.

Figure 1 illustrates monthly movements in the spreads in Belgium, France, and Italy during the early 1970s. The figure reveals that the spreads between the

Table 2. *Key Statistics on Spreads in Formal and Informal Dual Exchange Markets*
(percent)

| *Country and time period* | *Mean* | *Maximum* | *Minimum* | *Standard deviation* |
|---|---|---|---|---|
| *European DRS* | | | | |
| Belgium, May 1971–March 1974 | 0.04 | 1.27 | −2.05 | 0.67 |
| France, August 1971–March 1974 | −0.46 | 5.03 | −5.54 | 2.35 |
| Italy, January 1973–March 1974 | 3.90 | 7.48 | 0.76 | 2.08 |
| *Latin America (formal DRS)* | | | | |
| Argentina, June–December 1981 | 44.33 | 57.21 | 32.08 | 8.21 |
| July–October 1982 | 48.50 | 79.72 | 19.92 | 26.29 |
| Bolivia, April–October 1982 | 289.79 | 524.01 | 84.70 | 167.34 |
| Costa Rica, March 1981–October 1983 | 18.78 | 76.51 | −4.59 | 17.13 |
| Dominican Republic, August 1982–December 1984[a] | 69.73 | 113.33 | 48.58 | 17.53 |
| El Salvador, August 1982–December 1985 | 62.99 | 93.2 | 43.00 | 12.20 |
| Guatemala, November 1984–June 1988[b] | 80.50 | 280.00 | 0.40 | 96.44 |
| Jamaica, January–October 1983 | 57.90 | 66.15 | 52.12 | 5.87 |
| Mexico, August 1982–February 1988[c] | 15.14 | 108.79 | −1.26 | 18.54 |
| *Latin America (informal/black DRS)* | | | | |
| Peru, January 1980–July 1986 | 12.57 | 63.08 | −2.15 | 18.70 |
| Dominican Republic, January 1980–July 1982 | 38.00 | 50.00 | 29.00 | 3.97 |

*Note*: In formal dual exchange markets, the spread is defined as $[(X - S)/S]*100$ where $X$ is the financial exchange rate and $S$ is the commercial rate, both expressed as the ratio of domestic currency to the U.S. dollar. In informal dual exchange markets, the spread is defined as $[(B - S)/S]*100$ where $B$ is the parallel (or black) exchange rate. Calculations are based on monthly data and are period averages for Belgium, Italy, and El Salvador and end-of-period observations for all other countries.

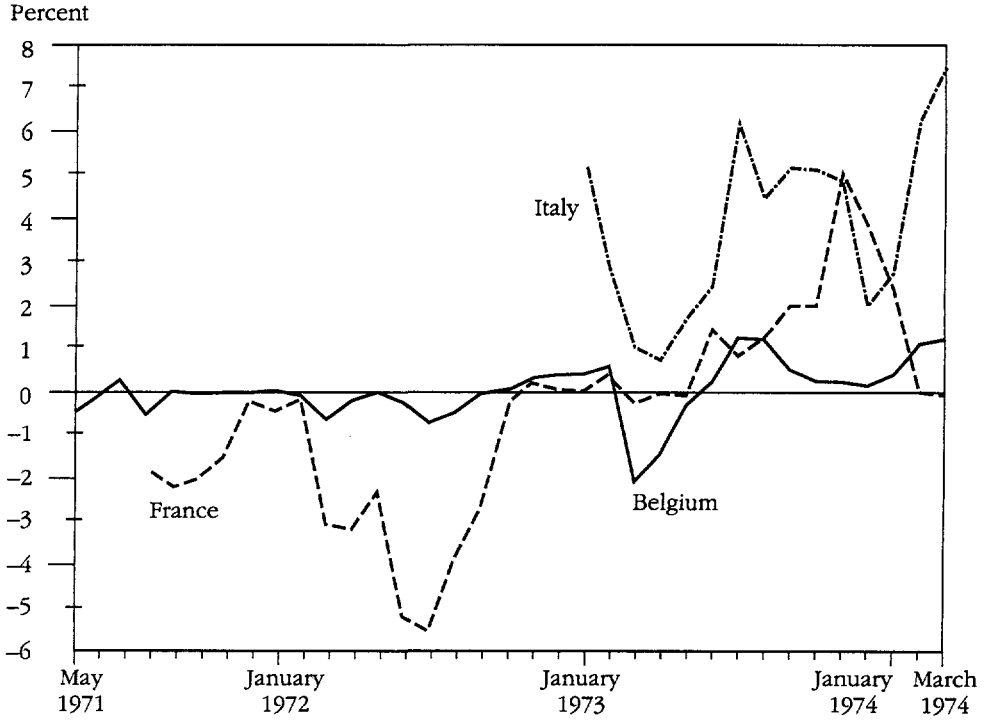
a. The DRS operated de jure over this period.

b. De facto reunification occurred in June 1988.

c. De facto reunification occurred in February 1988.

*Source*: IMF (various issues b); International Currency Analysis Inc. (various issues); unpublished country data.

Figure 1. *Spreads between the Official Financial and Commercial Exchange Rates in Belgium, France, and Italy, 1971–74*

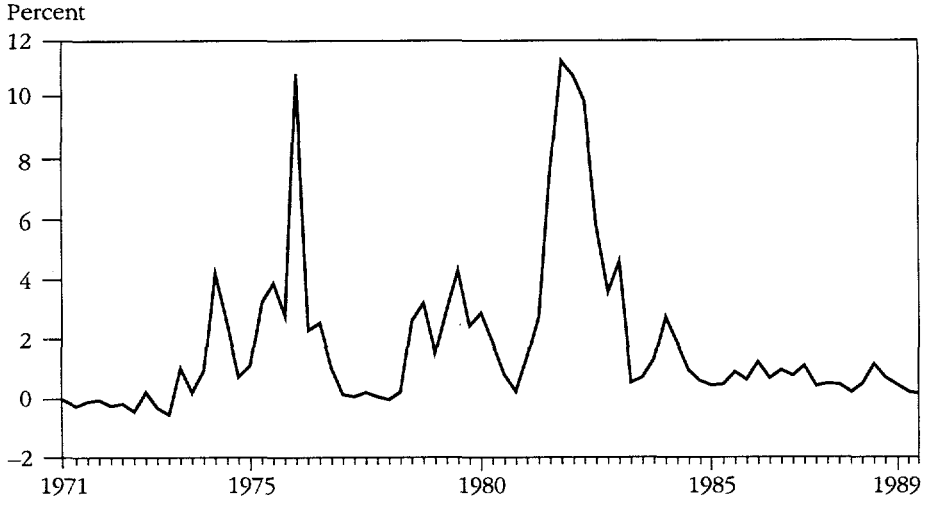



*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* Unpublished country data.

commercial and financial exchange rates were small for the three European countries. The Belgian spread was usually 1 percent or less. During the turbulent period surrounding the breakdown of Bretton Woods (May 1971 to March 1974), the mean spread was 0.04 percent, with a standard deviation of 0.67 percent. In France, the mean spread over the duration of its dual exchange market was −0.46 percent, with a standard deviation of 2.35 percent. Italy experienced somewhat larger spreads, on average, than its two European neighbors; spreads between 4 and 7 percent prevailed during eight of the fifteen months in which the Italian system operated. The mean spread was consequently a bit higher, at 3.9 percent, with a standard deviation of 2.08 percent. In all three countries, spreads in excess of 5 percent never persisted for more than two consecutive months in the early 1970s.

Figure 2, which shows the Belgian spread during a twenty-year period, reveals that the early period was not an aberration. Over the last two decades of Belgium's dual exchange market, the mean spread was 1.8 percent, with a standard deviation of 2.6 percent. The only big jumps in the spread occurred in

Figure 2. *Spread between the Official Financial and Commercial Exchange Rates in Belgium, 1971–89*



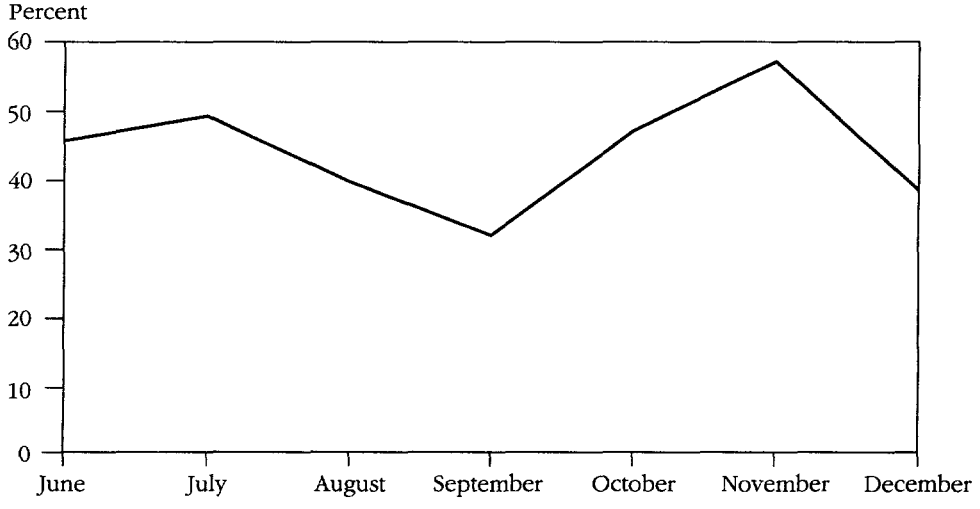

*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Quarterly data are used.
*Source:* IMF (various issues b).

the first quarter of 1976, when the spread exceeded 10 percent, and between the third quarter of 1981 and the second quarter of 1982, when spreads of 6 to 11 percent emerged.

Figures 3 to 11 present the spreads for Latin American countries operating formal dual markets in the 1980s. The data illustrate that dual exchange markets created larger and more persistent distortions in Latin America than in Europe. Whereas the spreads in the European episodes were quite small, usually in the 1 to 4 percent range, the Latin American spreads were large, generally in the 15 to 80 percent range. The figures show that in all the formal dual exchange market episodes, spreads in excess of 50 percent occasionally appeared, and in four out of the nine episodes spreads exceeded 100 percent at times. Guatemala experienced a period with spreads above 200 percent and Bolivia saw some spreads above 500 percent. Mexico and Costa Rica had relatively small spreads by Latin American standards, but even they had average spreads far in excess of the European spreads. Mexico's mean spread between 1982 and 1988 was 15 percent, and Costa Rica's was almost 19 percent. All Latin American countries experienced spreads in excess of 10 percent for sustained periods of six months or more. Figures 12 and 13 show that spreads in informal dual markets could be quite large and persistent as well.

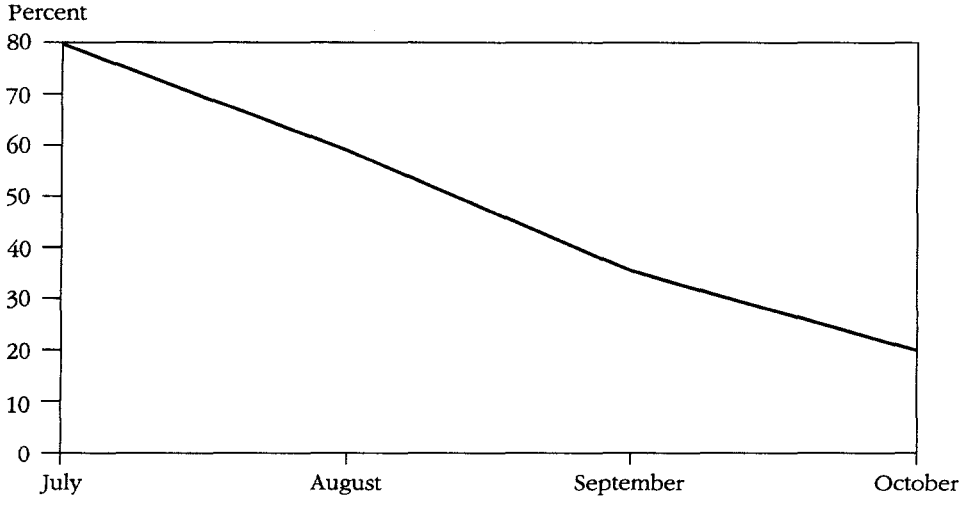
In summary, European spreads were small and spreads above 5 percent were short-lived, but Latin American spreads were large and persisted for extended

Figure 3. *Spread between the Official Financial and Commercial Exchange Rates in Argentina, 1981*



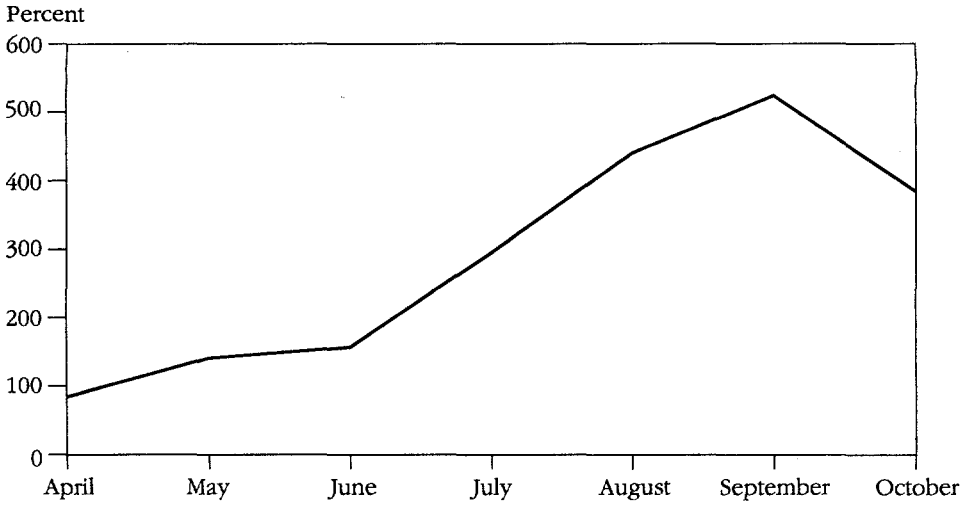

*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

Figure 4. *Spread between the Official Financial and Commercial Exchange Rates in Argentina, 1982*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

Figure 5. *Spread between the Official Financial and Commercial Exchange Rates in Bolivia, 1982*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

periods. The figures also reveal another difference across regions. In Europe, the price of foreign exchange could stand at a premium or discount in the financial exchange market compared with its price in the commercial exchange market. In Latin America, the price of foreign exchange invariably stood at a premium in the financial exchange market.

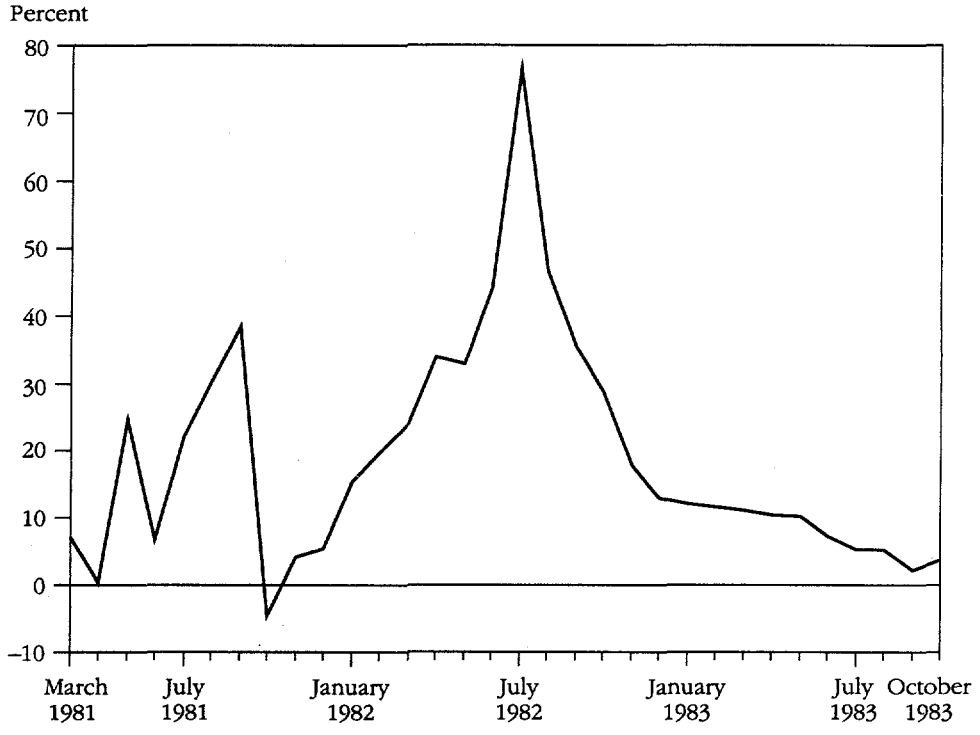
## II. Determinants of the Spread

The data raise the puzzling question of why spreads are so much larger and more persistent in Latin America than in Europe. It could be that the Latin American and European spreads are determined by the same set of economic variables, with the high-spread countries subject to larger disturbances in those underlying variables. Alternatively, the Latin American and European spreads could be determined by different factors, reflecting, in part, different institutional practices surrounding the operation of the dual exchange market. To shed some light on the puzzle, a simple model is used to identify key determinants of the spread and to derive a reduced-form equation for the spread. The equation is estimated using European and Latin American data to see whether the spreads in the two regions are determined by similar factors.

### *The Theory*

The simplest model for highlighting determinants of the spread between dual exchange rates is a stock-flow model that assumes a fixed or crawling commer-

Figure 6. *Spread between the Official Financial and Commercial Exchange Rates in Costa Rica, 1981–83*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

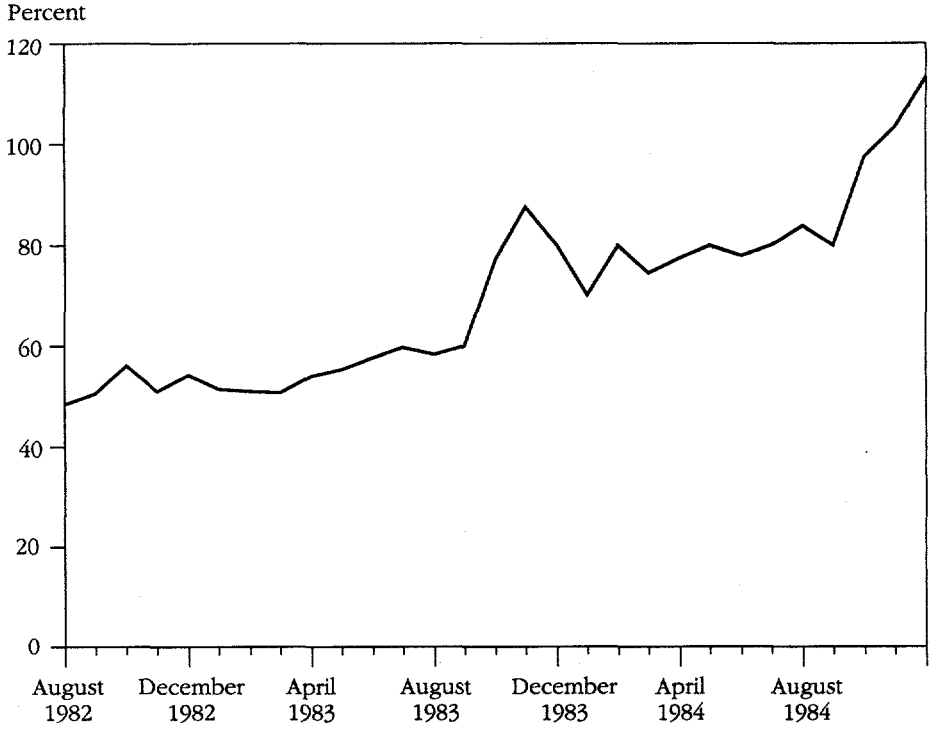
cial rate, a flexible financial exchange rate, domestic money and foreign bonds, purchasing power parity, and perfect foresight.[3] All commercial transactions are conducted at the commercial exchange rate and all financial transactions at the financial exchange rate. Complications arising from leakages are discussed later.

In the asset markets, the desired ratio of domestic money to foreign interest-bearing assets, $M/XF$, depends on the rate of return on foreign assets:

$$\frac{M}{XF} = L\left[i^*\left(\frac{1}{z}\right)(1+\delta)+f\right] \qquad L' < 0. \tag{2}$$

3. For examples of models with most or all of these features, see Flood (1978), Marion (1981), Flood and Marion (1982), Dornbusch and others (1983), Dornbusch (1986), Lizondo (1987), Pinto (1989), and Ghei and Kiguel (1992). Optimizing models can also be used to isolate economic determinants of the spread. See Adams and Greenwood (1985), Guidotti and Vegh (1988), Frenkel and Razin (1986), and Flood and Marion (1988).

Figure 7. *Spread between the Official Financial and Commercial Exchange Rates in the Dominican Republic, 1982–84*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

Domestic money earns no return. The return on foreign assets is calculated under the assumption that interest income, a current account transaction, is repatriated at the commercial exchange rate. The home currency price of foreign currency in the financial market is denoted by $X$, the stock of foreign assets by $F$, and the value of foreign assets in domestic currency by $XF$.

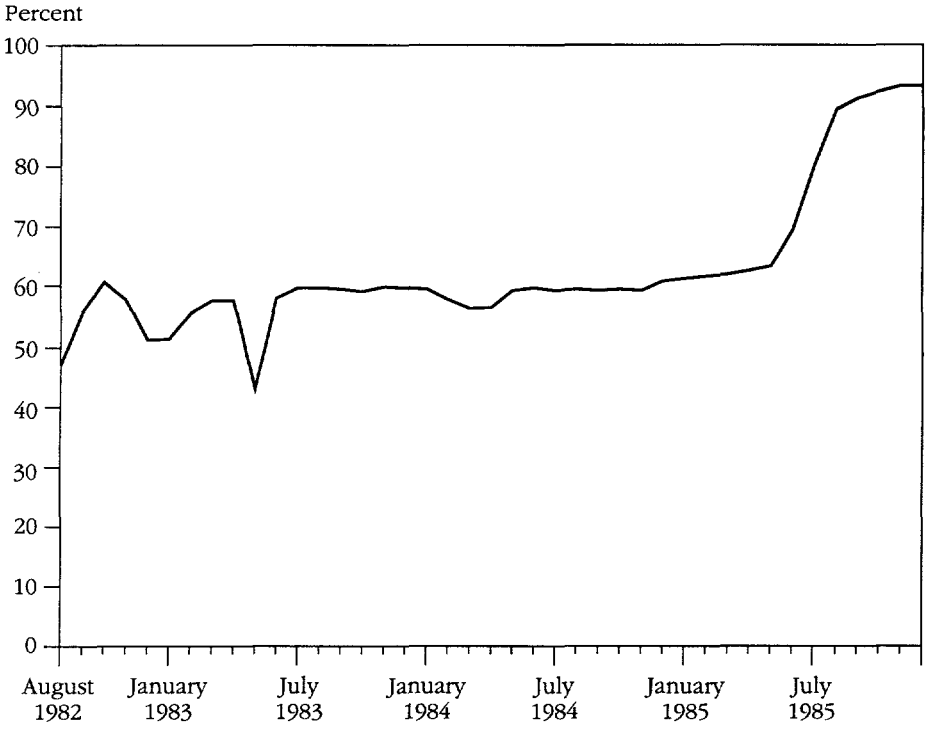
The rate of change of the spread, $\dot{z}/z$, is equal to the difference between the rate of depreciation of the financial exchange rate and the rate of crawl (managed depreciation) of the commercial exchange rate:

$$\frac{\dot{z}}{z} = f - \delta. \tag{3}$$

Substituting equation 3 into equation 2 and inverting gives the evolution of the spread over time:

$$\frac{\dot{z}}{z} = h\left(\frac{m}{zF}\right) - \left(\frac{i^*}{z}\right)(1 + \delta) - \delta, \qquad h = L^{-1}, h' < 0 \tag{4}$$

Figure 8. *Spread between the Official Financial and Commercial Exchange Rates in El Salvador, 1982–85*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

where $m = \frac{M}{S}$ and $z = \frac{X}{S}$

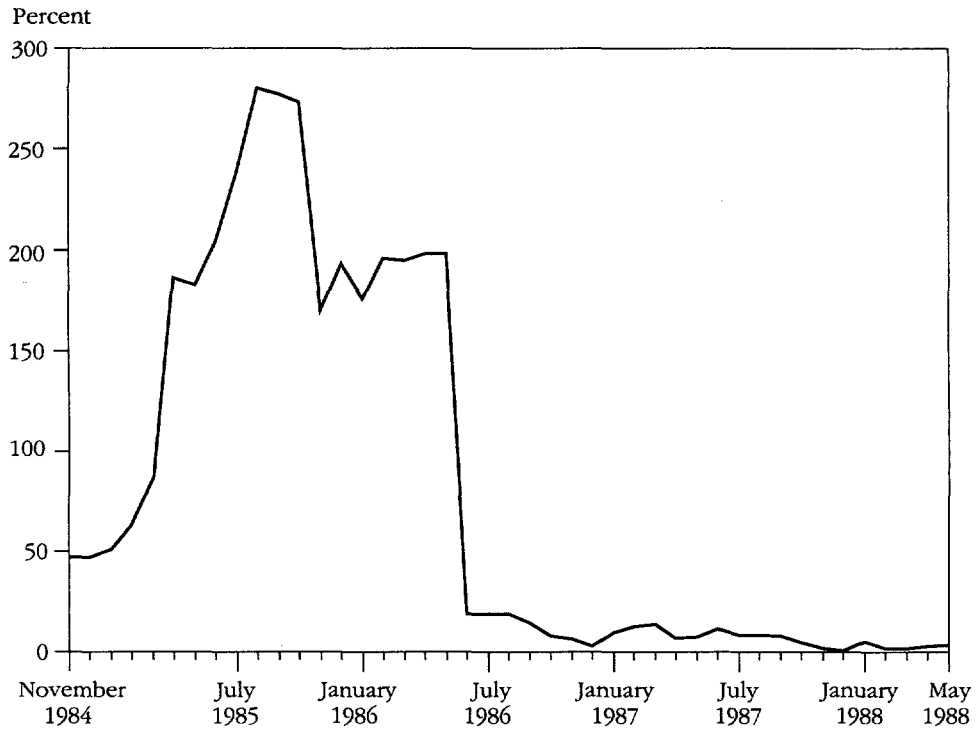
and $S$ is the home currency price of foreign currency in the commercial market. Equation 4 is the first dynamic equation of the model.

The second dynamic equation comes from equating expected real asset accumulation to planned real saving:

$$(5) \qquad \dot{m} + \left[h\left(\frac{m}{zF}\right) - \left(\frac{i^*}{z}\right)(1 + \delta) - \delta\right]zF = s(y, m + zF) \qquad s_1 > 0, s_2 < 0.$$

Households may increase their domestic currency wealth in several ways. They can add to their money holdings as the economy acquires reserves through current account surpluses or as the government undertakes domestic credit creation to finance budget deficits. They can also experience capital gains on their foreign asset holdings as a result of a depreciation of the financial exchange rate. They cannot increase their holdings of foreign assets, however. The flexible

Figure 9. *Spread between the Official Financial and Commercial Exchange Rates in Guatemala, 1984–88*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

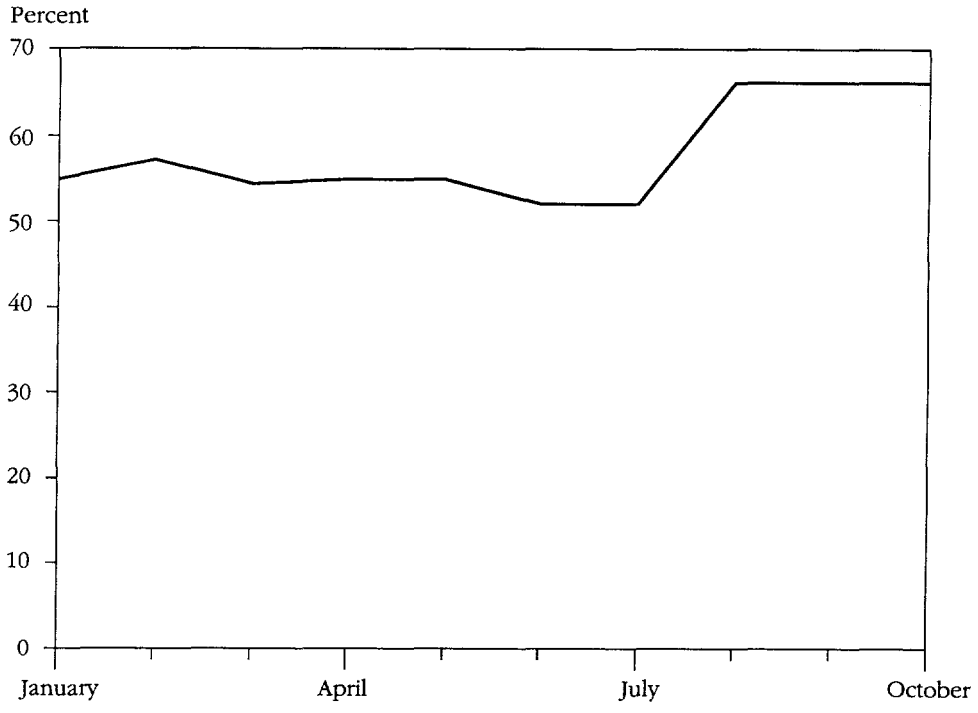
financial exchange rate prevents net capital flows through the capital account exchange market. Households cannot acquire foreign assets through the current account exchange market either, because proceeds from commercial transactions must be exchanged at the commercial exchange rate for domestic assets.

Assuming purchasing power parity and a fixed foreign price level normalized to one, the domestic price deflator is $P = S$. Then, using equations 3 and 4 and setting $\dot{F}/F = 0$ ($F$ is the initial stock of foreign assets at the onset of the dual exchange market) and $\dot{m} = [d(M/S)]/dt$, expected real asset accumulation is

$$d\,\frac{[(M + XF)/P]}{dt} = \dot{m} + (f - \delta)zF = \dot{m} + \left[h\left(\frac{m}{zF}\right) - \frac{i^*}{z}(1 + \delta) - \delta\right]zF.$$

The right-hand side of equation 5 specifies savings behavior. Planned real saving is assumed to depend positively on real income, $y$, and negatively on real wealth, $m + zF$. Using the national income identity and ignoring investment spending, saving represents the sum of the current account surplus and the

Figure 10. *Spread between the Official Financial and Commercial Exchange Rates in Jamaica, 1983*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

government budget deficit, all in real terms. Equation 5 describes the evolution of real wealth over time.

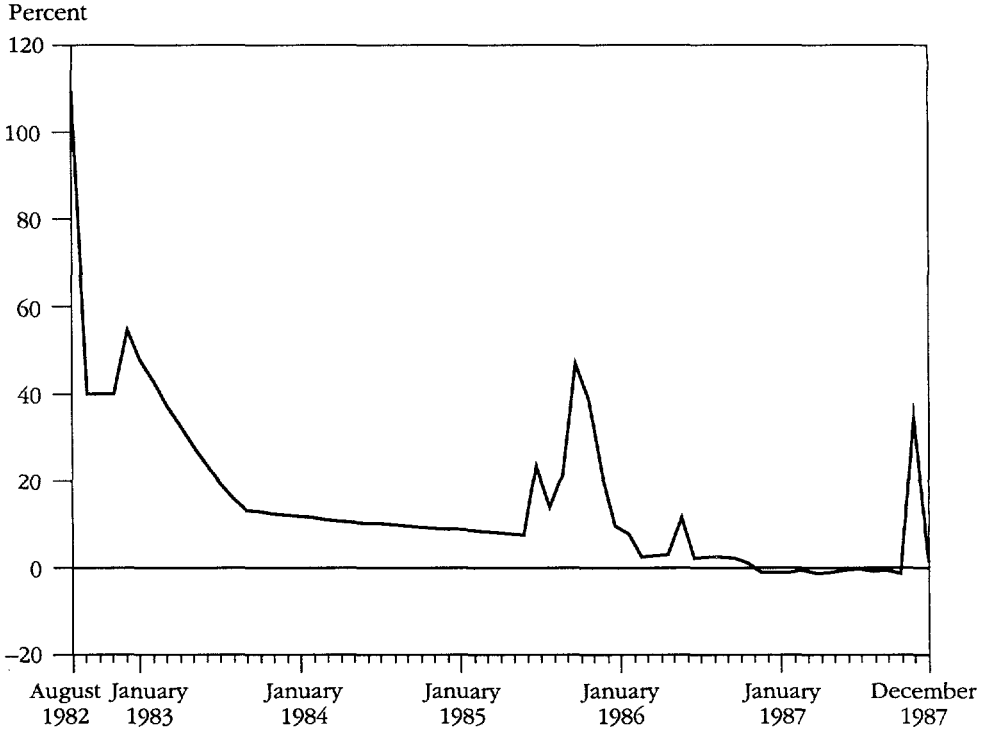
The steady state is attained when $\dot{z} = \dot{m} = 0$. In the steady state, the spread is constant. The financial and commercial exchange rates depreciate at the same rate. The real money stock is also constant. The growth of the nominal money stock is offset by the depreciation of the commercial exchange rate. The current account is balanced and the government deficit is financed by credit creation, with the depreciation of the commercial exchange rate taxing money balances.

Combining equations 4 and 5 and assuming convergence to the steady state, a reduced-form equation for the spread, $z$, depends on the foreign interest rate, the rate of depreciation of the commercial exchange rate, real income, and the initial stock of foreign assets at the onset of the dual exchange market:

$$z = z(i^*, \delta, y, F). \tag{6}$$

An increase in the foreign interest rate or the rate of depreciation of the commercial exchange rate shifts asset holders from money to foreign assets.

Figure 11. *Spread between the Official Financial and Commercial Exchange Rates in Mexico, 1982–87*



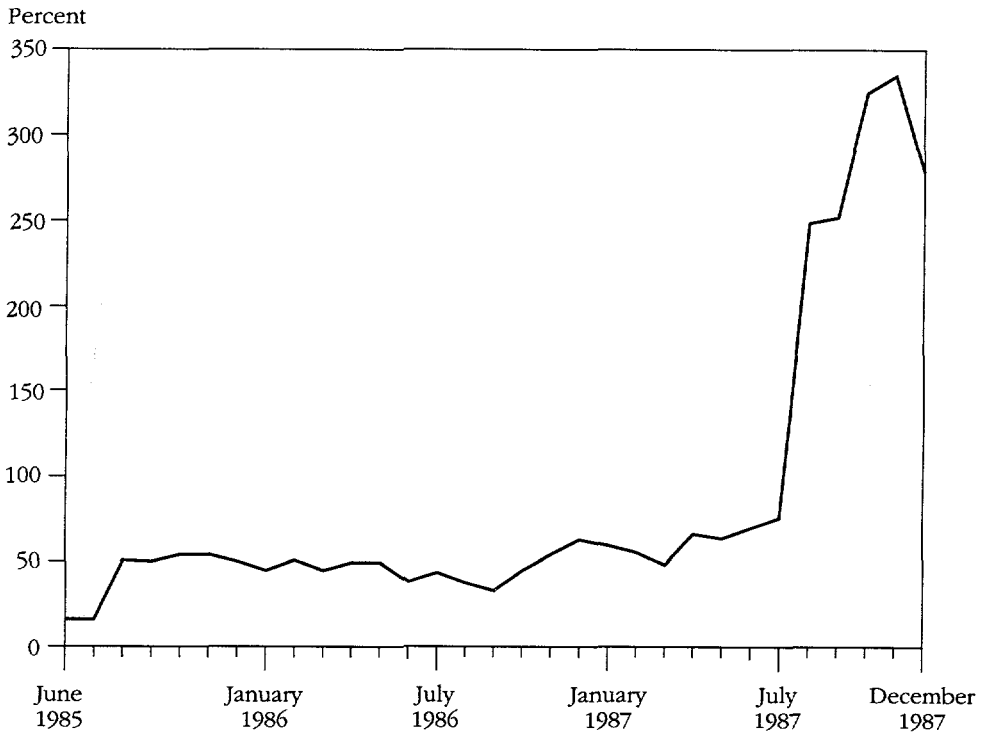

*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b) and unpublished country data.

Given the fixed supply of foreign assets, the spread must widen to restore portfolio balance. An increase in income will stimulate saving, generating current account surpluses that cause the real money stock to rise as the central bank intervenes in the commercial exchange market. The spread will rise as asset holders attempt to shift some of their new wealth into foreign assets.

An increase in the government budget deficit also widens the spread. With money demand inelastic with respect to inflation, an increase in the deficit financed by credit creation requires an increase in the rate of depreciation of the commercial rate to raise the needed inflation tax revenue. A faster depreciation of the commercial rate in turn widens the spread.

A shift in expectations also has an impact on the spread. For example, if agents believe that the government will increase the budget deficit and the rate of depreciation of the commercial exchange rate at some time in the future, they will now attempt a portfolio shift from money to foreign assets, which will lead to a jump in the spread. Thus, as described by Dornbusch (1986), changing expectations about the future course of fiscal policy can generate large fluctua-

Figure 12. *Spread between the Black Market and Official Commercial Exchange Rates in Peru, 1985–87*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b); Intenational Currency Analysis, Inc. (various issues).

tions in the spread. In addition, a shift in expectations about future income affects the spread because saving depends on some notion of expected permanent income.

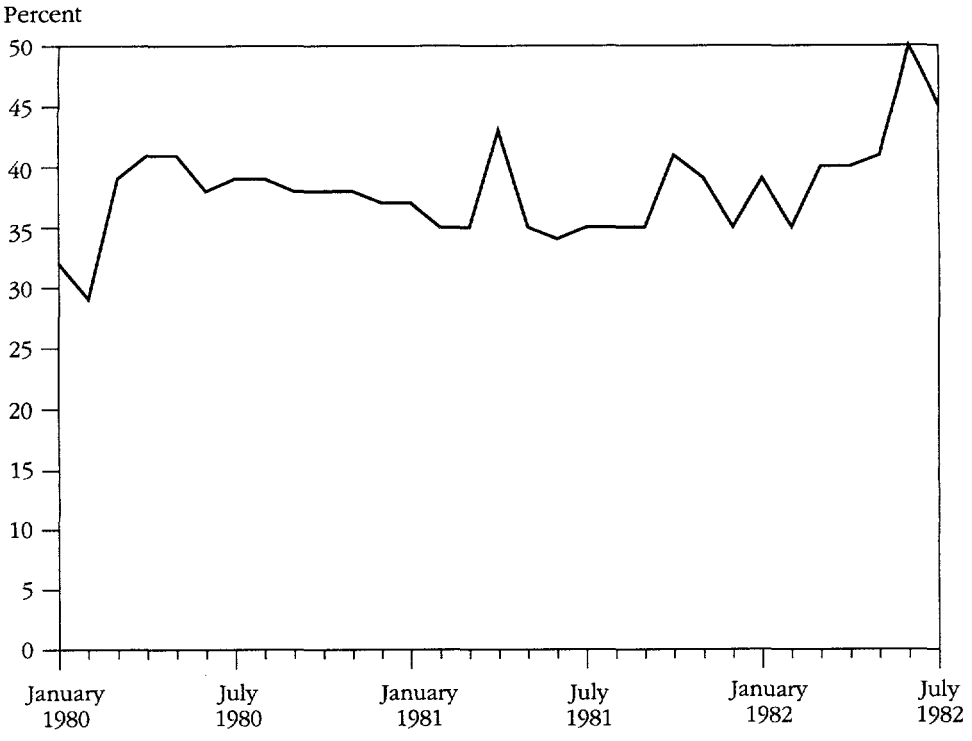
### *Empirical Estimation of the Spread*

Equation 6 guides the choice of economic variables in the regression analysis of the determinants of the spread in the European and Latin American dual exchange markets. The European spreads are estimated with quarterly data using the following initial specification:

$$\log z_t = \beta_0 + \beta_1 \log z_{t-1} + \beta_2(i_t^*) + \beta_3[L](\log y_t - \log y_{t-1}) + \beta_4[L](\log S_{t+1} - \log S_t) + \epsilon_t \tag{7}$$

where $\beta$[L] is a polynomial in the lead/lag operator L, $\log z_t$ is the log of the spread in time $t$, $i^*$ is the nominal foreign interest rate in percentage terms, ($\log y_t - \log y_{t-1}$) is the growth of real income, and ($\log S_{t+1} - \log S_t$) is the actual rate of depreciation of the commercial exchange rate, which is equal to the

Figure 13. *Spread between the Black Market and Official Commercial Exchange Rates in the Dominican Republic, 1980–82*




*Note:* The spread is the difference between the financial and commercial exchange rates in percentage terms. Monthly data are used.
*Source:* IMF (various issues b); International Currency Analysis, Inc. (various issues).

expected rate of depreciation plus a random component under rational expectations.

Although the spread is a real variable, its variance is dominated by the variance of the financial exchange rate because the commercial rate is generally fixed. Economists have had no success in explaining the variance of a flexible exchange rate solely in terms of fundamentals, so there is no reason to expect considerably more success in explaining the variance of the spread on the basis of fundamentals alone. Hence the lagged spread is included on the right-hand side as well. Dickey-Fuller tests confirm that the European spreads are stationary (as are the Latin American spreads), so the regression is run using the log-level spread rather than its rate of change. The variables on the right-hand side are also constructed to be stationary processes.[4]

4. Dickey-Fuller tests fail to reject the hypothesis of a unit root for the foreign interest rate, but on a priori grounds the interest rate should be bounded by reasonable limits.

Recognizing that the expected rate of depreciation of the commercial exchange rate may in fact deviate from its actual rate of depreciation, I also experiment with other specifications for expectations. I hypothesize that expectations may be influenced by factors such as movements in international reserves, the evolution of the real commercial exchange rate, and the size of the real budget deficit. A faster rate of reserve depletion brought about by increasing current account deficits may signal that the authorities are under pressure to increase the rate of depreciation of the commercial exchange rate. An increasing appreciation of the real commercial exchange rate that worsens the current account balance may also indicate the need to increase the rate of depreciation of the nominal commercial exchange rate. Finally, an increase in the real budget deficit may signal an eventual increase in the rate of depreciation of the commercial exchange rate to raise seigniorage revenue. Thus a modified version of equation 7 substitutes the terms $(\log R_t - \log R_{t-1})$, $(\log r_t - \log r_{t-1})$, and $(def_t)$ for the term $(\log S_{t+1} - \log S_t)$, where $R$ is the foreign currency value of international reserves, $r$ is the real commercial exchange rate, and $def$ is the real budget deficit. The deficit variable is not logged because some observations (surpluses) have negative values. I also experiment with various lag/lead structures for these variables. Note that the empirical work departs from the theoretical model by allowing for variation in the real commercial exchange rate.

In the regressions, the spread is calculated using end-of-quarter observations for the financial and commercial exchange rates. Foreign variables are proxied by U.S. data. The three-month U.S. Treasury bill rate is used for the foreign interest rate. Quarterly data on gross national product (GNP) are unavailable, so an index of industrial production is used for the income measure. Data on the net stock of foreign assets are also unavailable, but because the net stock does not change during the operation of dual exchange markets, $F$ can be subsumed in the constant term. The stock of nongold international reserves is chosen as the reserves variable. The real commercial exchange rate is measured by the consumer price–adjusted bilateral exchange rate with the United States. The real budget deficit is calculated by deflating the nominal budget deficit (in billions of national currency units) by the domestic consumer price index.

Ideally, the same set of explanatory variables would be used in the Latin American regressions. Unfortunately, data problems make this impossible. Income figures are often unavailable on a quarterly basis. For two countries, budget deficit figures are also lacking on a quarterly basis. Two compromises are needed to estimate similar regressions for the European and Latin American spreads. First, the European regressions are run with and without the income variables in order to have one version of the regression duplicate the Latin American regressions. Second, when required, figures for the annual budget deficit are used to compute quarterly estimates.

The first set of regressions based on equation 7 is reported in table 3. Because most dual exchange market episodes are too short for estimation procedures, the European data and the Latin American data are each pooled. Belgium, however,

Table 3. *Estimation Results for Models Using the Rate of Depreciation of the Commercial Exchange Rate*

| *Variable or statistic* | *Europe*[a] | *Belgium*[b] *With income terms* | *Belgium*[b] *Without income terms* | *Latin America*[c] |
|---|---|---|---|---|
| Constant | −0.008* | −0.008 | −0.004 | 0.075 |
| | (−1.65) | (−1.56) | (0.84) | (1.42) |
| Spread between exchange rates last period, log $z_{t-1}$ | 0.512** | 0.55** | 0.56** | 0.82** |
| | (6.26) | (6.88) | (6.85) | (18.57) |
| Nominal foreign interest rate, $i^*$ | 0.009** | 0.008** | 0.007** | −0.01 |
| | (3.20) | (2.83) | (2.36) | (−0.54) |
| Rate of depreciation of the commercial exchange rate | | | | |
| Next period log $S_{t+1}$ − log $S_t$ | 0.03 | 0.04 | 0.04 | 0.02 |
| | (0.96) | (1.12) | (1.34) | (0.36) |
| In two periods, log $S_{t+2}$ − log $S_{t+1}$ | 0.05 | 0.07** | 0.05 | −0.04 |
| | (1.51) | (2.05) | (1.56) | (−0.79) |
| Growth of real income | | | | |
| This period, log $y_t$ − log $y_{t-1}$ | 0.12 | 0.13* | | |
| | (1.64) | (1.78) | | |
| Next period, log $y_{t+1}$ − log $y_t$ | 0.14** | 0.16** | | |
| | (1.99) | (2.28) | | |
| $\bar{R}^2$ | 0.57 | 0.56 | 0.53 | 0.71 |
| Sample size | 103 | 94 | 94 | 201 |

* Significant at the 90 percent confidence level.

** Significant at the 95 percent confidence level.

*Note*: The dependent variable is the (log) spread between exchange rates in the financial and commercial foreign exchange markets. Quarterly data are used. *t*-statistics are in parentheses.

a. Regression uses pooled data for Belgium (second quarter 1963 to fourth quarter 1971 and first quarter 1973 to third quarter 1987), France (fourth quarter 1971 to second quarter 1973), and Italy (second quarter 1973 to third quarter 1973).

b. Regression uses data for Belgium (second quarter 1963 to fourth quarter 1971 and first quarter 1973 to third quarter 1987).

c. Regression uses pooled data for Argentina (fourth quarter 1971 to first quarter 1976), Chile (second quarter 1970 to second quarter 1975 and fourth quarter 1982 to second quarter 1988), Costa Rica (second quarter 1970 to second quarter 1972 and fourth quarter 1981 to first quarter 1983), the Dominican Republic (fourth quarter 1982 to second quarter 1984), El Salvador (fourth quarter 1982 to third quarter 1984), Guatemala (first quarter 1985 to fourth quarter 1988), Jamaica (third quarter 1977 to fourth quarter 1977), Mexico (fourth quarter 1982 to fourth quarter 1988), Paraguay (second quarter 1983 to fourth quarter 1986), Peru (second quarter 1971 to first quarter 1977 and second quarter 1987 to fourth quarter 1988), and Uruguay (second quarter 1972 to second quarter 1978).

*Source*: IMF (various issues b); unpublished country data.

had a long continuous experience with dual exchange markets, so separate regressions of the Belgium spread are also reported.

The first regression reported in table 3 uses pooled Belgian, French, and Italian data. The results show that all the explanatory variables affect the spread in ways consistent with the stock-flow model and that some are significant at the 95 percent confidence level. An increase in the foreign interest rate has a positive and highly significant effect on the spread, as does an expected increase in income growth. An expected increase in the rate of depreciation of the commercial exchange rate also widens the spread, but the effect is not significant. Together with the lagged spread, these variables explain 57 percent of the variation in the Eu-

ropean spreads. Since country and time dummies were all insignificant, they are not reported. Additional lags on the spread or leads on the rate of depreciation and on income growth were also insignificant.

The results are quite similar for the same regression run on Belgian data alone, although the expected rate of depreciation of the commercial exchange rate is highly significant as well. The strong correlation between the expected increase in the rate of depreciation and the spread helps explain the behavior of the Belgian spread in the first quarter of 1976 and from the third quarter of 1981 to the second quarter of 1982 (figure 2). The dramatic widening of the Belgian spread in the first quarter of 1976 corresponded to speculative pressures against the French franc that spilled over to the Belgian commercial franc when the French franc departed from the snake and the Belgian current account shifted into a deficit that quarter.[5] The wider spread during the period from the third quarter of 1981 to the second quarter of 1982 (7 to 11 percent range) reflected speculative activity surrounding three devaluations of the Belgian commercial franc against the deutsche mark that occurred within the European Monetary System in the fourth quarter of 1981 and the first and second quarters of 1982.

The Belgian regression is run a second time without the income terms in order to make it identical to the Latin American regression. The omission of income terms marginally reduces the explanatory power of the regression compared with the one for Belgium with income terms.

The regression for the Latin American spreads uses pooled data from eleven countries: Argentina, Chile, Costa Rica, the Dominican Republic, El Salvador, Guatemala, Jamaica, Mexico, Paraguay, Peru, and Uruguay. The panel contains the set of countries that operated official dual exchange markets in some part of the period between the first quarter of 1970 and the second quarter of 1989. The Argentine episodes of 1981 and 1982 and the Bolivian episode of 1982 were excluded from the panel because they contained only two quarterly observations. Country and time dummies are not reported.

In contrast to the European experience, in Latin America movements in the foreign interest rate have no effect on the spread. This finding suggests that while insulation of domestic interest rates from foreign interest rate disturbances may have been a motive for adopting the DRS in Europe, this was not the case in Latin America. The expression for the effective return in equation 1 indicates that the DRS may provide insulation from foreign interest rate shocks if such shocks generate offsetting movements in the spread. In the European and Belgian regressions, an increase in the foreign interest rate does in fact generate some offsetting increase in the spread. The same pattern of response is not evident in the Latin American data.

The regression results for Latin America also show that a change in the future rate of depreciation of the commercial exchange rate has no effect on the Latin

5. The snake refers to the close margins for the bilateral exchange rates of European currencies that were negotiated near the end of the Bretton Woods system.

American spreads. This result is not surprising because actual and expected rates of depreciation of the commercial exchange rate are likely to be quite different. Because the Latin American authorities always ended DRS episodes with a devaluation of the commercial exchange rate (see table 1), agents would naturally expect, with some positive probability, a devaluation in the next quarter or two even though the realized commercial rate stayed fixed for extended periods.

One final observation is that the coefficient attached to the lagged spread is a bigger fraction in the Latin American regression than in the European one. This result captures what was seen in the figures, namely, that the spreads are more per sistent in the Latin American cases. Indeed, the higher $R^2$ for the Latin American regression comes from the fact that the lagged spread is a very important determinant of the current spread.

Table 4 reports estimation results for variations on the regressions in table 3.

Table 4. *Estimation Results for Models Using Alternative Proxies for the Rate of Depreciation of the Commercial Exchange Rate*

| | *Belgium*[a] | | |
|---|---|---|---|
| *Variable or statistic* | *With income terms* | *Without income terms* | *Latin America*[b] |
| Constant | −0.003 | −0.001 | −0.03 |
| | (−0.72) | (−0.34) | (−0.63) |
| Spread between exchange rates last period, $z_{t-1}$ | 0.58** | 0.58** | 0.79** |
| | (8.02) | (8.0) | (17.92) |
| Nominal foreign interest rates, $i^*$ | 0.005* | 0.005* | 0.008 |
| | (1.82) | (1.67) | (0.31) |
| Real budget deficit, *def* | 0.36E-02* | 0.32E-02 | 0.32E-05** |
| | (1.71) | (1.54) | (2.41) |
| Change in foreign-currency value of international reserves, $\log R_t - \log R_{t-1}$ | −0.06** | −0.06** | −0.06** |
| | (−5.03) | (−5.52) | (−2.07) |
| Change in real commercial exchange rate, $\log r_t - \log r_{t-1}$ | −0.02 | −0.02 | −0.43** |
| | (−0.79) | (−0.67) | (−7.68) |
| Growth of real income | | | |
| This period, $\log y_t - \log y_{t-1}$ | 0.006 | | |
| | (0.09) | | |
| Next period, $\log y_{t+1} - \log y_t$ | 0.11* | | |
| | (1.67) | | |
| $\bar{R}^2$ | 0.63 | 0.63 | 0.81 |
| Sample size | 94 | 94 | 175 |

* Significant at the 90 percent confidence level.

** Significant at the 95 percent confidence level.

*Note*: The dependent variable is the (log) spread between exchange rates in the financial and commercial foreign exchange markets. Quarterly data are used. *t*-statistics are in parentheses.

a. Regression uses data for Belgium (second quarter 1963 to fourth quarter 1971 and first quarter 1973 to third quarter 1987).

b. Regression uses pooled data for Argentina (first quarter 1973 to first quarter 1976), Chile (first quarter 1974 to second quarter 1975 and fourth quarter 1982 to second quarter 1988), Costa Rica (first quarter 1971 to second quarter 1972 and fourth quarter 1981 to first quarter 1983), the Dominican Republic (fourth quarter 1982 to second quarter 1984), El Salvador (fourth quarter 1982 to third quarter 1984), Guatemala (first quarter 1985 to fourth quarter 1988), Jamaica (third quarter 1977 to fourth quarter 1977), Mexico (fourth quarter 1982 to fourth quarter 1988), Paraguay (second quarter 1983 to fourth quarter 1986), Peru (second quarter 1971 to first quarter 1977 and second quarter 1987 to fourth quarter 1988), and Uruguay (second quarter 1972 to second quarter 1974 and second quarter 1976 to second quarter 1978).

*Source*: IMF (various issues b); unpublished country data.

The regression results in table 4 use the real budget deficit and the rates of change in both international reserves and the real commercial exchange rate to capture expectations about future movements in the nominal commercial exchange rate. Because the European and Belgian regression results are nearly identical, only the Belgian regressions are reported.

The regression for Belgium with income terms shows that the variables proxying for expectations have sensible and important effects on the Belgian spread. An increase in the real budget deficit widens the spread, as does a more rapid rate of reserve depletion. The coefficient is significant at the 90 percent confidence level on the deficit variable and at the 95 percent level on the reserve variable. A faster real appreciation of the commercial exchange rate has no significant effect on the spread, however. Using a different real exchange rate specification, such as the bilateral rate with Germany, did not change the insignificance of the real commercial exchange rate variable. Expected future income growth and current increases in the foreign interest rate have positive and significant effects on the Belgian spread, as in the earlier regressions.

Although the coefficient on the real deficit variable is positive, it is not highly significant. One possible reason may be that, in practice, an increase in the budget deficit has opposing effects on the spread. As described by the theoretical model, a budget deficit financed by credit creation ultimately requires a more rapid depreciation of the commercial exchange rate, which raises the return on foreign assets, triggers an attempted portfolio shift, and widens the spread. However, an increase in the deficit may also be financed by bond sales. In that case, although the accumulation of domestic bonds tends to widen the spread as agents try to diversify their new wealth, higher domestic interest rates tend to narrow the spread as agents try to shift into the higher-yielding domestic assets. The positive coefficient on the deficit term suggests that the wealth effects may dominate the interest rate effect, but not by much. Additional regressions were run using both the foreign and domestic interest rates or the interest differential as right-hand-side variables. The results must be treated cautiously because the domestic interest rate is not really an exogenous variable. Nevertheless, it is interesting to note that, controlling for domestic interest rates, an increase in the real budget deficit has a highly significant positive effect on the spread.[6]

6. The regression using Belgian data is

$$\log z_t = \underset{(0.56)}{0.001} + \underset{(13.51)}{0.76 \log z_{t-1}} + \underset{(4.08)}{0.01(i_t^* - i_t)} - \underset{(-1.06)}{0.06(\log y_t - \log y_{t-1})} + \underset{(0.26)}{0.017(\log y_{t+1} - \log y_t)}$$

$$+ \underset{(2.55)}{0.44\text{E}{-05}\, def_t} - \underset{(-4.94)}{0.059(\log R_t - \log R_{t-1})} - \underset{(-1.95)}{0.064(\log r_t - \log r_{t-1})}$$

$$R^2 = 0.75, n = 94$$

($t$-statistics are in parentheses). A correction has been made for serial correlation. Regressions of the black market spread commonly include the interest differential on the right-hand side in the belief that activities in the black market do not affect domestic variables such as the interest rate. For the European DRS

The regression for Belgium without the income terms highlights determinants of the Belgian spread in a way that duplicates the regression for the Latin American spreads. The omission of the income terms has little effect on the estimation results.

In the regression for Latin America, expectational factors also seem to be important. Both increased budget deficits and a faster depletion of international reserves widen the spread and are highly significant explanatory variables. In addition, and in contrast to the Belgian case, a faster real appreciation of the commercial exchange rate also has a positive and highly significant effect on the spread. Indeed, the size of the coefficient indicates that real appreciations have economically significant effects on the Latin American spreads. Because real appreciations are more dramatic in Latin America, they apparently provide a stronger signal of a pending devaluation of the commercial exchange rate. As before, the lagged spread is an important determinant of the current spread, but the foreign interest rate has no effect on the spread.

To test whether the positive relation between the budget deficit and the Latin American spread is robust, the regression for Latin America was rerun over different samples. In some cases, the coefficient on the deficit variable is insignificant. Although there is convincing evidence that the budget deficit is a strong predictor of the black market premium, the spread between two official exchange rates appears to be less sensitive to this factor. What holds up across all samples and specifications is the importance of the lagged spread and real commercial exchange rate in determining the Latin American spread. Real appreciations of the commercial exchange rate appear to affect expectations strongly. Moreover, to the extent that these appreciations induce the authorities to shift some current account transactions that are in large deficit at the commercial exchange rate to the financial exchange market, they may help explain why the spreads are so much larger in Latin America than in Europe.

## III. Drawbacks in Empirical Estimation

Some important caveats about the empirical tests are in order. I have analyzed the spread on the assumption that there is a structural relationship between the spread and its determinants. However, because the dual exchange market is generally a temporary arrangement and because institutional practices vary during its operation, this assumption may not be appropriate.

Consider first the issue of temporariness. With the exception of Belgium, which used a dual exchange market for more than thirty years, countries adopt

---

episodes, however, the financial exchange rate and domestic interest rate are jointly determined endogenous variables.

When the first regression in table 4 was rerun using current and future values of the real budget deficit, the coefficient on the deficit one quarter ahead was negative and highly significant. One explanation consistent with this result is that an increase in the next period's budget deficit puts upward pressure on current domestic interest rates, causing a portfolio shift into domestic assets that narrows the current spread.

the DRS for a relatively short time. France used the arrangement for thirty-one months and Italy for fifteen months. Table 1 shows that the Latin American dual exchange markets were also of limited duration. Argentina used the official dual exchange market for six months or less in 1981 and again in 1982, Bolivia operated a DRS for seven months, Jamaica for ten months, and so on. When private agents believe the dual exchange market is temporary, the spread will be influenced by probability-weighted beliefs of the possible regimes to follow. These beliefs build into the data elements that are difficult to model in empirical work and contaminate the link between the spread and its determinants, particularly around the time of reunification.

Changes in institutional practices during the operation of a dual exchange market also affect the structural stability of the regression equation. Regressions for various sample periods of a single DRS episode show that determinants of the spread can differ in their importance across time as institutional practices change. There are numerous examples of these institutional changes. For instance, between the second quarter of 1971 and the first quarter of 1983, Belgium channeled both capital inflows and outflows through the financial market. Before and after that period, Belgium treated capital flows asymmetrically, with capital outflows assigned to the financial market and capital inflows free to go through either market. Italy introduced a 50 percent deposit requirement on capital exports six months after adopting its DRS. Costa Rica, El Salvador, and Jamaica heavily managed and even fixed the financial exchange rate for a time during the operation of their dual exchange markets.

The lack of data makes it impossible to measure the extent of European foreign exchange intervention in the financial market. In the Belgian case, it appears that neither systematic intervention over the long run nor large-scale, short-run intervention occurred in the financial exchange market (Bindert-Bogdanowicz 1979). In the French and Italian cases, the general view is that there was little or no management of the financial exchange rate, at least through direct foreign exchange intervention. More indirect ways of influencing the rate, such as encouraging public sector borrowing and lending through the financial market, were attempted.

All three European countries did have periods when both exchange rates floated. Between August and December 1971, the Belgian commercial franc floated against all currencies except the Dutch guilder. The Italian commercial lira floated between January and March in 1974. The Italian commercial lira began its float on February 13, 1973, just weeks after the Italian DRS was established. The Italian commercial lira floated in its separate exchange market until the DRS was abolished in March 1974.

The reduced-form equation for the spread and the estimation relied on the assumption that all current account transactions are channeled through the commercial exchange market and all capital account transactions through the financial market. In practice, officially sanctioned and fraudulent cross-market leakages occur. To the extent that these leakages depend on the size of the

spread, the reduced-form equation for the spread is unchanged. Nevertheless, these leakages have macroeconomic consequences and can moderate the response of the spread to various disturbances.

Consider the case where some current account transactions are officially channeled through the financial exchange market. This modified segmentation of the exchange market has several macroeconomic implications. The theoretical model is complex because there are three dynamic variables: the real money stock, the stock of foreign assets, and the spread. A flexible financial exchange rate no longer prevents the net accumulation of foreign assets. To keep balanced trade in the financial exchange market, net accumulation of foreign assets must accompany exports of goods and services channeled through that market. (Similarly a net decumulation accompanies imports.) If the authorities want to limit net accumulation of foreign assets even more, they must supplement the DRS with quantitative capital controls. Most Latin American countries, in fact, employed capital controls when operating dual exchange markets (see table 1), as did France and Italy to a lesser extent.

With some current account items traded in the financial exchange market, the aggregate price level becomes an expenditure-weighted function of both the financial and commercial exchange rates. In addition, relative prices are distorted, with the relative price of goods traded in the financial market measured by the spread. The DRS thus provides less insulation of domestic prices and reserves from financial shocks. Changes in expectations about the future commercial exchange rate, for example, generate portfolio shifts that alter the spread and affect aggregate and relative prices, spending, and the current account.

Even if the authorities try to achieve a complete separation of current account and capital account transactions, illegal leakages occur. When the spread widens, private agents have an incentive to buy foreign exchange at the commercial exchange rate and sell it at the financial exchange rate. Lanyi (1975) has described how these illegal transactions can take place. Exporters underinvoice their sales receipts and invest the unrecorded payments in foreign assets. The proceeds from the sales of these assets are then repatriated at the financial rate. This strategy succeeds if the authorities minimize administrative costs by requiring documentation only for the purchase and sale of foreign exchange in the commercial exchange market. Importers overinvoice and follow a similar strategy. Thus when the spread widens in response to a disturbance, leakages may quickly dampen the spread (Gros 1988). In addition, the authorities may encourage such leakages by officially reclassifying certain transactions. The correlation between economic factors and the spread may thus be difficult to detect except in high-frequency data.

It is important to look at how the European and Latin American countries segmented their exchange markets in practice. If current account and capital account transactions are broadly separated, the distortion created by the DRS is essentially an intertemporal one. If a number of current account transactions are

Table 5. *Division of the Foreign Exchange Market*

| *Country* | *Date* | *Principal market* | *Secondary market* |
|---|---|---|---|
| Argentina | 1981 | Imports, most exports, amortization payments on foreign loans (mixed rate for trade[a]) | All else |
| Argentina | 1982 | Imports, exports (mixed rate for trade[a]) | All else |
| Bolivia | 1982 | Wheat imports, public debt service payments | All else |
| Costa Rica | May 1983 | Payments of external debt service, imports of certain essential commodities, 40 percent of private sector debt service payments, 99 percent of export proceeds, 100 percent of official capital inflows | All else, including most imports |
| Dominican Republic (the) | December 1984 | Debt repayments, oil imports, specified imports | All else |
| El Salvador | December 1984 | 70 percent of export proceeds, 60 percent of import payments | All else |
| El Salvador | Late 1985 | Proceeds from coffee and sugar exports, inflows of foreign loans to public sector and banking system, some specified imports such as fuel and medicine, interest and amortization payments of the public sector and banking system | Imports of consumer goods, imports of most inputs and invisibles, some traditional exports, most proceeds from nontraditional exports of goods and services, private capital inflows, authorized capital outflows |

| | | | |
|---|---|---|---|
| Guatemala | 1984 | Most export receipts, proceeds from foreign borrowing, external public debt payments, certain private debt payments, all official transactions, essential imports (market handles about 55 percent of exports and 35 percent of imports); mixed export rate in effect | All else |
| Guatemala | June 1986 | Exports, imports, most capital transactions | Remittances, tourism, low-priority imports |
| Jamaica | 1983 | Essential imports, traditional exports, specified invisibles, exports to CARICOM (the non-Caribbean Common Market) countries | All else |
| Mexico | August 1982 | Priority imports, petroleum export proceeds, foreign debt payments | All else |
| Mexico | December 1982 | Merchandise export proceeds, foreign debt repayments including principal and interest, specified import payments | All else |

a. Mixed rate means that a set percentage of the transactions goes through the principal market and the remainder through the secondary market.
*Source*: IMF (various years a).

combined with capital account transactions in the financial exchange market, the distortion becomes as much a distortion of the goods market as one of the assets market.

In the 1970s all three European countries segmented the exchange market so as to bring about an almost complete separation of current account and capital account transactions. (Of course, private arbitrage activity made the separation less than perfect.) In May 1971, a period of heightened speculation against the dollar, Belgium made important changes in the classification of transactions—changes that resulted in an almost complete separation of the commercial and financial exchange markets. Nevertheless, a small number of current account transactions could be undertaken in either exchange market, and individual licenses could be granted to allow certain capital account transactions through the commercial market. In addition, domestic and foreign bank notes, representing private travel expenses and so forth, could be bought and sold on the financial market. In January 1974 outward payments of investment earnings could be channeled through either market.

Under the French scheme there was also a broad separation of commercial and financial transactions, but it was by no means complete. Some current account items, such as travel, tourism, investment income, workers' remittances, and bank note transactions, were channeled through the financial market. A few financial transactions, such as those related to commercial credits, were channeled through the commercial exchange market.

Under the Italian system, current account and capital account transactions were as nearly as possible separated into the corresponding exchange markets. Nevertheless, all purchases and sales of foreign bank notes, which accounted for a substantial portion of tourist expenditures and workers' remittances, were assigned to the financial exchange market. For details, see IMF (various years a).

In Latin America, the official segmentation was much less along commercial-financial lines. As seen in table 5, many current account items were assigned to the financial exchange market. In Argentina in 1981 and 1982, imports, most exports, and amortization payments on foreign loans were channeled through the commercial exchange market, and everything else went though the financial market. Yet in both episodes, a time-varying mixed rate, which specified that a percentage of export receipts and import payments had to be settled in the financial market, was quickly established. El Salvador and Guatemala also established a mixed rate. Costa Rica initially channeled specified imports, most exports, and certain debt repayments through the commercial exchange market, but eventually most imports were directed to the financial market. Mexico achieved the broadest separation between current account and capital account transactions, but it was by no means complete. The overall picture that emerges is that the Latin American dual exchange markets distorted relative prices of goods and services to a much greater extent than did the European dual exchange markets.

## IV. Conclusion

A number of theoretical models have been developed to analyze dual exchange markets, but no attempt has been made to develop some empirical regularities about the distortion created by this sort of exchange rate arrangement. This article tries to fill that gap by examining the data from three European countries and eleven Latin American countries that operated official dual exchange markets in the 1970s or 1980s.

Section I examined the behavior of the spread between the commercial and financial exchange rates. The spread and its evolution over time generate the intertemporal distortion that influences international capital flows. The spread also measures the distortion in relative prices if some current account transactions are directed to the financial exchange market. The spreads have been quite small in European dual exchange markets but quite large in Latin American ones.

Section II set out a standard stock-flow macroeconomic model to isolate some of the economic determinants of the spread between the two exchange rates. It then presented a reduced-form equation for the spread based on the model and fitted it to the data. The regressions of the spread showed that portfolio variables are important determinants of the spread in both Europe and Latin America. However, there are interesting differences in the relative importance of these variables. In Europe, the foreign interest rate and rate of change of international reserves, along with the lagged spread, explained almost two-thirds of the variation in the spread. An increase in the real government budget deficit also widened the spread, but the effect was not highly significant. The rate of change of the real commercial exchange rate was uncorrelated with the spread. In Latin America, the rate of change of the real commercial exchange rate and the rate of international reserve depletion were important determinants that, along with the lagged spread, explained around 80 percent of the variation of the Latin American spreads. Latin American spreads were positively related to budget deficits, but only in some specifications. Latin American spreads were not responsive to movements in foreign interest rates, however, suggesting that the DRS was not adopted to insulate domestic interest rates in Latin America. Instead, the dual exchange market was designed to delay an across-the-board devaluation of the currency.

One reason for the absence of empirical work on dual exchange markets is the transitional nature of most of these regimes. Section III acknowledged that the regression results presented in this article should be treated with caution because the inherent temporariness of dual exchange markets casts doubt as to the stability of the structural relationship between the spread and its economic determinants. Frequent rule changes during the operation of a dual exchange market further weaken the claim that a dual exchange market episode can be treated as a single event. In addition, officially sanctioned and fraudulent cross-market leakages that dampen the spread between the two exchange rates can

make it difficult to detect the links between the spread and its determinants in low-frequency data. Alternatively, reclassifying current account transactions that are in large deficit at the fixed commercial rate as "financial" transactions can widen the spread and also make it difficult to uncover the links between the spread and its other determinants. Finally, when official dual exchange markets are not broadly divided between commercial and financial transactions, the distortion created by dual exchange markets becomes as much a distortion of the goods market as of the assets market.

## References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Adams, Charles, and Jeremy Greenwood. 1985. "Dual Exchange Rate Systems and Capital Controls: An Investigation." *Journal of International Economics* 18(February):43–63.

Bhandari, Jagdeep, and Bernard Decaluwe. 1987. "A Stochastic Model of Incomplete Separation between Commercial and Financial Exchange Markets." *Journal of International Economics* 22(February):25–55.

Bindert-Bogdanowicz, Christine. 1979. "The Dual Exchange Rate System." *Revue de la Banque de Bruxelles*. June.

Central Bank of Bolivia. Various issues. *Buletin Estadistico*. Bogotá.

Dornbusch, Rudiger. 1986. "Special Exchange Rates for Capital Account Transactions." *The World Bank Economic Review* 1(1):3–33.

Dornbusch, Rudiger, Daniel Dantas, Clarice Pechman, Roberto de Rezende Rocha, and Demetrio Simoes. 1983. "The Black Market for Dollars in Brazil." *Quarterly Journal of Economics* 98(February):25–40.

Fleming, John Marcus. 1971. "Dual Exchange Rates for Current and Capital Transactions: A Theoretical Examination." In John Marcus Fleming, *Essays in International Economics*. Cambridge, Mass.: Harvard University Press.

Flood, Robert. 1978. "Exchange Rate Expectations in Dual Exchange Markets." *Journal of International Economics* 8(February):65–77.

Flood, Robert, and Nancy Marion. 1982. "Exchange-Rate Regimes in Transition: Italy 1974." *Journal of International Money and Finance* 2(December):279–94.

———. 1988. "Determinants of the Spread in a Two-Tier Foreign Exchange Market." *Economics Letters* 27(2):173–79.

Frenkel, Jacob, and Assaf Razin. 1986. "The Limited Viability of Dual Exchange-Rate Regimes." NBER Working Paper 1902. National Bureau of Economic Research, Washington, D.C.

Ghei, Nita, and Miguel Kiguel. 1992. "Dual and Multiple Exchange Rate Systems in Developing Countries: Some Empirical Evidence." Working Paper 881. World Bank, Country Economics Department, Washington, D.C. Processed.

Gros, Daniel. 1988. "Dual Exchange Rates in the Presence of Incomplete Market Separation." *International Monetary Fund Staff Papers* 35:437–60.

Guidotti, Pablo, and Carlos Vegh. 1988. "Macroeconomic Interdependence under Capital Controls: A Two-Country Model of Dual Exchange Rates." IMF Working Paper 88/74. International Monetary Fund, Washington, D.C.

IMF (International Monetary Fund). Various years a. *Annual Report on Exchange Arrangements and Exchange Restrictions*. Washington, D.C.

———. Various issues b. *International Financial Statistics*.

International Currency Analysis, Inc. Various issues. *World Currency Yearbook*. (formerly *Pick's Currency Yearbook*). Brooklyn, N.Y.

Lanyi, Anthony. 1975. "Separate Exchange Markets for Capital and Current Transactions." *International Monetary Fund Staff Papers* 22(November):714–49.

Lizondo, Jose Saul. 1987. "Unification of Dual Exchange Markets." *Journal of International Economics* 22(February):57–77.

Marion, Nancy. 1981. "Insulation Properties of a Two-Tier Exchange Market in a Portfolio Balance Model." *Economica* 48(February):61–70.

Pinto, Brian. 1989. "Black Market Premia, Exchange Rate Unification, and Inflation in Sub-Saharan Africa." *The World Bank Economic Review* 3(3):321–38.

# The Impact of Mexico's Retraining Program on Employment and Wages

Ana Revenga, Michelle Riboud, and Hong Tan

*This article analyzes the impact and effectiveness of the Mexican labor retraining program for unemployed and displaced workers—Programa de Becas de Capacitación para Trabajadores (PROBECAT). The strategy followed is to compare the post-training labor market experiences of trainees with those of a comparison group—a matched sample of unemployed individuals who were eligible for, but did not participate in, PROBECAT. The results of this exercise suggest that participation in PROBECAT reduced the mean duration of unemployment for both men and women trainees and increased the monthly earnings of men, but not of women. The results also indicate that the post-training earnings effect varied systematically by level of education attained, with the largest earnings increases (of about 28 to 37 percent) found for men with six to twelve years of education.*

In 1984, as a response to a growing economic crisis, the government of Mexico established a labor retraining program for unemployed and displaced workers—Programa de Becas de Capacitación para Trabajadores, or PROBECAT. Its objective was to dampen the social costs of major economic restructuring and rising unemployment. As adjustment efforts accelerated during the latter half of the decade, the need for policies targeting the unemployed and facilitating their reemployment became more pressing. As a result, in 1987 the retraining program was strengthened and its scope and coverage expanded. Since then, PROBECAT has provided short-term vocational training to more than 250,000 unemployed people.

The Mexican government is currently considering an extension of PROBECAT for several reasons. First, and most important, is concern about the impact of the North America Free Trade Agreement (NAFTA) on migration flows, especially from rural areas, and on unemployment. Second, although the adjustment

At the World Bank, Ana Revenga is in the Latin America and Caribbean Country Department II, Michelle Riboud is in the Europe and Central Asia Country Department IV, and Hong Tan is in the Private Sector Development Department. Ana Revenga also holds a joint appointment at the Banco de España. The authors thank the Dirección General de Empleo, Secretaría del Trabajo y Previsión Social, Mexico, for providing the data used in this article. They also acknowledge Feliciano Iglesias, Frank Lysy, Sweder van Wijnbergen, and the journal referees for their helpful comments, and Muriel Aza for providing competent research assistance.

process to date has taken place with relatively little effect on observed levels of unemployment, substantial labor reallocation between expanding and contracting sectors is likely to occur with further liberalization and privatization of the Mexican economy. Third, PROBECAT is the only unemployment program currently in place in Mexico.

In making decisions about the future of PROBECAT, policymakers in Mexico will need improved information about the labor market impacts of retraining on target populations as well as information about the cost-effectiveness of the program. Although evaluating the impact of such a program is an accepted practice in many industrial countries, it is less common among developing countries. A notable exception is the evaluation of Colombia's Servicio Nacional de Aprendizaje training program by Jimenez and Kugler (1987). In part, the lack of evaluation may be due to a paucity of relevant data and to lack of familiarity with program evaluation methodologies. For Mexico, the availability of longitudinal data on both a cohort of PROBECAT trainees and a comparison group of unemployed offers a unique opportunity to study the impact of retraining programs in a developing country.

The purpose of this article is to evaluate the impact of PROBECAT on the employment and incomes of trainees. We seek to address four key questions. First, what is the impact of training on the subsequent employment experiences of trainees? Second, does training increase the speed with which trainees move from unemployment to employment? Third, conditional upon finding employment, what effect does training have on the monthly earnings, work hours per week, and hourly wages of trainees? Fourth, do the monetary benefits from program participation outweigh the costs of providing retraining for the unemployed?

We address these issues by comparing the post-training labor market experiences of PROBECAT trainees with those of a comparison group—a sample of unemployed individuals who were eligible for, but did not participate in, PROBECAT. For the trainees, we use detailed data on the post-training experiences of the 1990 trainee cohort elicited in a retrospective survey conducted by the Secretaría del Trabajo y Previsión Social in 1992. For the comparison group, we use panel data on a random sample of unemployed individuals drawn from the 1990–91 quarterly urban labor force survey Encuesta Nacional de Empleo Urbano (ENEU).

This approach improves on previous evaluations of PROBECAT (see, for example, the reports by the Secretaría de Trabajo y Previsión Social 1988, 1989, and 1990 and Carlson 1991). Earlier studies were subject to several data and methodological limitations. One limitation was the crudeness of wage data. Earnings information was bracketed and reported only in reference to the minimum wage. Another limitation was that there were no comparison groups; the outcomes for training completers six months after completion of the training program were compared with the outcomes for training dropouts three months after dropping out. Both limitations have been overcome in our evaluation.

We use a statistical methodology to account for selection bias arising from the nonrandom selection of individuals into PROBECAT. An alternative, experimental evaluation methodology is to randomly assign individuals into two groups: participants and nonparticipants. Because random assignment avoids the issue of selection bias, the impact of the program can be evaluated by simply comparing outcomes for the two groups. The use of statistical methods for evaluating the programs has both detractors and supporters; several studies suggest that nonexperimental methods may be subject to misspecification error, whereas other studies question whether experimental evaluations are really necessary. LaLonde (1984) and Fraker and Maynard (1985) critique nonexperimental evaluation methodologies; Heckman and Hotz (1987, 1989) defend the statistical evaluation approach; and Levitan (1992) summarizes the advantages and disadvantages of both statistical and experimental approaches to program evaluation. We acknowledge that the statistical methods used here are not immune to criticism, and we therefore caution that the results be interpreted with care—more as initial estimates than as a definitive evaluation of PROBECAT. Nonetheless, we note that the experimental evaluation approach is both politically and practically difficult to implement and thus is not a viable option for many developing countries. In these countries, improving both the quality of data and methodologies used in program evaluation may result in greater payoffs.

Section I provides a broad overview of unemployment in Mexico and of PROBECAT. It also describes several surveys we have used in comparing trainees and unemployed individuals who did not participate in training. Section II discusses several methodological issues that arise in training program evaluations and describes our approach to resolving them. Sections III and IV report our estimates of the effects of training on the probabilities of employment, time to first job, monthly earnings, work hours per week, and hourly wages. Section V presents initial estimates of the cost-benefit ratios of PROBECAT training for men and women participants. We conclude by summarizing the most important findings and discussing their implications.

## I. Unemployment and PROBECAT

According to official statistics, the open unemployment rate in Mexico is relatively low. In 1992 it stood at 2.9 percent of the labor force, and even in the worst years of the adjustment crisis it did not rise beyond 6.1 percent. However, these figures have several shortcomings. First, they refer only to urban unemployment and thus exclude the sizable fraction of the Mexican population living in rural areas. Second, they are based on a loose definition of employment, in which an individual who works at least one hour a week is counted as employed. Third, they include only those individuals who are actively searching for a job. This last point is important because research suggests that the distinction between "unemployed" and "not in the labor force" based on intensity of search is usually very weak (see Clark and Summers 1979 and Summers 1986). This is

Table 1. *Distribution of the Unemployed by Age, Mexico, 1988*

| | Men | | Women |
|---|---|---|---|
| *Age* | *Standard definition* | *Expanded definition* | *Standard definition* |
| 12–15 | 4.2 | 10.7 | 2.5 |
| 16–20 | 33.3 | 33.5 | 36.6 |
| 21–25 | 24.7 | 19.3 | 38.1[a] |
| 26–30 | 13.3 | 10.5 | 38.1[a] |
| 31–40 | 8.8 | 9.0 | 11.7 |
| 41–50 | 8.0 | 9.5 | 6.9 |
| 51–60 | 5.1 | 5.9 | 2.5 |
| 61–70 | 2.6 | 1.6 | 1.7 |

*Note:* The data are from the Encuesta Nacional de Empleo (1988) survey, which covers all urban areas and a sample of the rural population and contains a sample of about 46,000 households. The standard definition of unemployment defines individuals as unemployed if they are actively looking for a job. In the expanded definition, the unemployed are defined as those under age 55 who are not working, not studying, and not retired but are able to work (not sick or disabled), regardless of whether they are actively searching for a job. Women who report being at home taking care of the house are not counted as unemployed.

a. Refers to the combined age categories 21–25 and 26–30

*Source*: Authors' calculations from the Encuesta Nacional de Empleo, 1988.

confirmed by our analysis of the employment data for Mexico (Revenga and Riboud 1993). We find a large fraction of idle men—men who are out of work but are able to work, who are not studying, and who are not taking care of a household. When the definition of unemployment is expanded to include these idle men, the aggregate unemployment rate in 1991 rises from 2.8 to 5.5 percent.[1]

Tables 1 and 2 report the distribution of the unemployed by age and by education, respectively. When the standard definition of unemployment is used, 75.5 percent of total unemployment for men is accounted for by individuals age 30 and below. The comparable figure for women is even higher: 77.2 percent. With regard to education, 53.5 percent of total male unemployment and 63.1 percent of female unemployment are accounted for by those with seven to twelve years of schooling. Individuals with completed secondary education (nine years of schooling) account for 20.4 percent of male unemployment and 18.9 percent of total female unemployment.[2] Those with a higher secondary education (ten to twelve years of schooling) account for an additional 20.2 percent of male unemployment and 35.6 percent of female unemployment.

When the expanded definition of unemployment is used, the overall unemployment rate rises as noted above, and the unemployment distribution of men by age and education changes. The proportion of unemployed men age 25 and

1. This figure is calculated for 1991 because that is the last year for which we have access to the detailed unemployment survey tapes. Note that we do not define a comparable group of idle female workers because family responsibilities tend to make their labor market behavior patterns much more complex, with frequent periods out of the labor force.

2. In Mexico there are six years of secondary education following primary school. The first three years are referred to as secondary education, the second three years as higher secondary education.

Table 2. *Distribution of the Unemployed by Education Level, Mexico, 1988*

| | *Men* | | *Women* |
|---|---|---|---|
| *Years of school* | *Standard definition* | *Expanded definition* | *Standard definition* |
| 0 | 2.4 | 4.4 | 1.6 |
| 1–5 | 9.7 | 12.5 | 8.6 |
| 6 | 16.0 | 20.0 | 14.6 |
| 7–8 | 12.9 | 12.8 | 8.6 |
| 9 | 20.4 | 18.8 | 18.9 |
| 10–12 | 20.2 | 18.1 | 35.6 |
| 13 or more | 18.4 | 13.4 | 12.1 |

*Note:* The data are from the Encuesta Nacional de Empleo (1988) survey, which covers all urban areas and a sample of the rural population and contains a sample of about 46,000 households. The standard definition of unemployment defines individuals as unemployed if they are actively looking for a job. In the expanded definition, the unemployed are defined as those under age 55 who are not working, not studying, and not retired but are able to work (not sick or disabled), regardless of whether they are actively searching for a job. Women who report being at home taking care of the house are not counted as unemployed.

*Source:* Authors' calculations from the Encuesta Nacional de Empleo, 1988.

below increases slightly, from 62.2 to 63.5 percent. Much more striking is the increase in the proportion of unemployed men with less than nine years of completed schooling: from 41 to 49.7 percent. This more economically meaningful, expanded definition of unemployment will be used throughout the analyses that follow.

### *Program Features of PROBECAT*

PROBECAT is administered through the network of state employment offices. Since 1987 it has trained 251,181 unemployed persons and provided 9,268 courses. During the training period, program participants receive a stipend equal to the minimum wage. Upon completion of the course, the local state employment office helps trainees find a job. (Most trainees surveyed, however, found jobs on their own and not through the state employment office.)

The majority of program participants enroll in classroom training, primarily in short-term vocational courses offered through contracts with local private and public institutions. Courses vary in duration from one to six months, the majority of courses (87 percent) lasting about three months. Training is provided in a variety of occupational areas: carpentry, construction, electricity, food preparation, graphic arts and design, handicrafts, machinery, mechanics, refrigeration, services and administration, shoe repair, textiles and apparel, and welding. In principle, courses are organized to respond to the needs of the local labor market and are designed to redress local shortages of workers with particular skills. These needs are determined through periodic studies of local labor market conditions.

Not everyone is eligible to participate in PROBECAT. The selection procedure gives variable weights to different criteria, including the number of economic dependents, attainment of certain levels of basic education, prior work experi-

ence, and unemployment of less than three months. The weighting scheme is quite complex and nonlinear, and only individuals with a total composite score exceeding a threshold level are eligible to join the program. In addition, participants must (in theory) be between the ages of 20 and 55 and be registered as job seekers at the local state employment office.[3] This nonrandom selection of individuals into PROBECAT poses potentially serious measurement problems for an evaluation of the training program.

### *Data Sources*

A number of surveys have been fielded to help monitor and evaluate PROBECAT. The first set of surveys, comprising follow-ups of trainees at three and six months after program completion, has been used in several reports by the Secretaría del Trabajo y Previsión Social (1988, 1989, and 1990). A second, more complete retrospective survey was administered to the 1990 cohort of trainees in early 1992.[4] It elicited a wealth of information on all jobs held between the completion of training and February 1992, including start and end dates for each job, monthly earnings, work hours per week, occupation, and industry. Our evaluation is based on this second PROBECAT survey.

As a comparison group for the trainees, we used a sample of unemployed individuals drawn from the 1990–91 ENEU. The ENEU, a household-based survey of the sixteen main urban areas in Mexico, elicited detailed information on employment status, jobs, monthly earnings, and work hours per week and was broadly comparable to the PROBECAT survey. The ENEU uses a quarterly rotation system so that each rotation group (of households) remains in the survey for five consecutive quarters and then leaves the sample. We obtained panel data for the rotation group that remained in the survey from the third quarter of 1990 to the third quarter of 1991—the period spanned by the trainee data—and drew our comparison group from this sample. This comparison group included all those who were unemployed in the third quarter of 1990 (whom we then tracked for a year). Note that the comparison group is based on the expanded definition of unemployment that includes all individuals who report being out of work the previous week, are able to work, and are not students or retirees, whether or not they are searching for a job. For certain analyses, we have augmented this comparison group with a second cohort of those who became unemployed in the fourth quarter of 1990 (and were not in the first cohort). For the latter cohort, only nine months of data are available.

3. The original age bracket was amended to allow a small number of participants between the ages of 16 and 20. Follow-up surveys also show the presence of a few participants above age 55.

4. This retrospective survey was based on a sample of 1,995 trainees who were administered a three-month follow-up in 1990. Of this original sample, 273 individuals could not be located for the 1992 retrospective survey. Consistency checks also revealed the presence of four individuals in the 1992 survey who were not part of the original sample.

Table 3 presents summary information on the demographic characteristics of the trainee and comparison group samples. In 1990 the average PROBECAT male trainee was 28 years old; the majority had completed primary schooling and some secondary education, and about 41 percent were married. The average female trainee was 29 years old; female trainees were less likely than men to have a higher education, and about 46 percent were married. Almost half the men (42.8 percent) identified themselves as being household heads. It is evident from table 3 that trainees differed from the general population of the unemployed. Compared with the sample drawn from the ENEU, trainees tended to be slightly older. They were more likely to be married, to be the household head, and to have completed secondary school. They also included a higher proportion of women (women were 49.0 percent of the trainee group but only 33.8 percent of the comparison group).

## II. The Evaluation Methodology

We estimated the impact of training on several outcome measures: the time taken to exit from unemployment (that is, time to first job); the probability of employment at three, six, and twelve months after the end of training; post-training monthly earnings; work hours per week; and hourly wages. Analysis of such a wide variety of outcome measures departs from the traditional focus of most training evaluations, which is on the impact of training on earnings, with relatively little attention paid to its impact on subsequent employment. An exception is Card and Sullivan (1988), which looks at the impact of training on post-training employment histories. Most studies focus on earnings outcomes, but typically without trying to disentangle the separate effects of training on employment, earnings, hours of work, and hourly wages.

We believe that our approach provides a more complete characterization of program effects than does the traditional approach. For example, earnings comparisons are contingent on having a job, and one impact of training may be to increase the likelihood of employment. Card and Sullivan (1988) provide evidence that a large part of the measured effect of training on earnings is a result of increases in the post-training employment of trainees. Similarly, monthly earnings are the product of hours of work and hourly wage rates, and training may have very different effects on each of these two outcomes.

The principal methodological issue that arises in evaluating the impact of PROBECAT is that of selectivity bias. As table 3 demonstrates, trainees are a nonrandom sample of the unemployed population. Failure to control for the differences in observed characteristics of trainees and the comparison group can lead to biases in estimated program impacts. These biases are potentially exacerbated by systematic differences across groups in unobserved (by the analyst) characteristics, such as motivation, ability, or tastes.

We addressed the selection bias problem using two approaches. The first approach was used in the analyses of the employment effects of training.

Table 3. *Demographic Characteristics of the Trainee and Comparison Group Samples*

| Characteristic | Men | | Women | |
|---|---|---|---|---|
| | Trainees | Comparison group | Trainees | Comparison group |
| Age (average years) | 27.9 | 24.6 | 29.0 | 23.6 |
| Married (percent) | 41.2 | 19.7 | 45.7 | 21.2 |
| Unmarried couple (percent) | 3.8 | 2.7 | 3.3 | 2.1 |
| Average years of school | 9.1 | 7.8 | 7.8 | 9.2 |
| Highest educational level reached (percent) | | | | |
| No formal education | 0.1 | 4.3 | 0.4 | 1.6 |
| Primary incomplete | 3.4 | 17.3 | 9.1 | 7.9 |
| Primary complete | 13.2 | 22.6 | 18.1 | 22.2 |
| Secondary incomplete[a] | 17.5 | 12.1 | 24.5 | 12.2 |
| Secondary complete | 30.5 | 18.9 | 29.7 | 15.9 |
| Higher secondary | 26.6 | 14.6 | 13.8 | 26.5 |
| University | 8.6 | 10.2 | 4.3 | 13.7 |
| Head of household (percent) | 42.8 | 23.7 | 11.6 | 6.3 |
| Sample size | 881 | 371 | 845 | 189 |

*Note*: The comparison group includes all unemployed individuals in the third quarter of the 1990 Encuesta Nacional de Empleo Urbano (ENEU).

a. In Mexico there are six years of secondary education following primary school. The first three years are referred to as secondary education and the second three years as higher secondary education.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

PROBECAT's own selection criteria were used to define a matched group of unemployed individuals with attributes similar to those of the trainees. This involves two steps. The first step is to estimate a probit model on the pooled trainee and unemployed samples relating the likelihood of program participation to the PROBECAT selection criteria for which we have data—marital status, number of children, number of dependents, years of education, and time in unemployment. The second step is to limit the comparison group to unemployed individuals with high predicted probability of program participation. A similar approach is followed in Westat (1981, 1984), Bassi (1983), and Geraci (1984).

The second approach, which was used in the analyses of monthly earnings, hours of work, and hourly wage outcomes, is based on the two-stage selectivity correction procedure developed by Heckman (1979). This involves, as before, estimating a model of selection into PROBECAT, calculating a variable to capture the individual's likelihood of program selection, and including this variable as a regressor in the outcome models to control for sample selectivity.

Our use of two different selectivity correction approaches for discrete and continuous outcomes is justified on econometric grounds. The Heckman approach—which we used for earnings and hours worked—relies critically on the nonlinearity of the first-stage probit for identification of the selection correction term in the second stage. Arguably, the same approach is not appropriate for analyzing employment outcomes because the second-stage outcome equations are themselves nonlinear and the observables in the first-stage participation probit are likely to be correlated with the unobservables in the second stage. Therefore, for this set of discrete outcomes, we adopted the matching procedure.

Neither approach to addressing the problem of selection bias is completely satisfactory. In both strategies, we were forced to address the selection issue through the use of cross-sectional control variables, such as level of education and demographic characteristics. This will yield correct estimates if selection into the program is determined solely by observable characteristics for which we are able to control. However, if selection occurs on the basis of unobservable variables, or if it is influenced by variables for which we cannot control, then our estimates could easily be biased. This no doubt detracts from the overall robustness of our estimates of program impacts.

## III. Effects of Training on Employment

We began the PROBECAT evaluation by assessing the impact of program participation on the likelihood of employment, both in the short term and over increasingly longer periods of time. First, we asked if participation in the training program had any effect on the time it takes trainees to move from unemployment into a first job. Next, we asked whether trainees systematically differed from the comparison group in their probability of employment after three, six, and twelve months of unemployment. For trainees, unemployment was mea-

sured after completion of training; for the comparison group, it was measured from the third quarter of 1990. Together, the two sets of analyses can be used to draw inferences about the proportion of time both groups spend in employment during the first twelve months.

### *Time to First Job*

What is the impact of PROBECAT on the time to first job? For trainees, it was straightforward to construct a continuous measure of time to first job (expressed in months) using information on the end dates for training and the start dates for the first job. For trainees who had not found a job within the sample period, the time-to-first-job variable was truncated (censored) at February 1992, and this censoring was taken into account in model estimation. Constructing a measure of the time to first job was more difficult for the comparison group. For this group, we had a continuous measure of time in unemployment up until the third quarter of 1990; subsequently, we observed the group's employment status only at discrete points in time (quarterly) over a one-year interval. The issue is that when an individual's employment status first changes from one quarter to the next, we must infer when, within a three-month period, the individual would find a job.

A number of assumptions may be used to estimate the commencement of employment. First, we can treat the unemployment duration reported by the comparison group in the third quarter of 1990 as being representative of the underlying distribution of incomplete unemployment spells. By appealing to steady-state assumptions, we can estimate the distribution of completed spells of unemployment by doubling the duration of incomplete spells reported (see Salant 1977). This assumption is fairly strong, but not absurd for the Mexican data. Revenga and Riboud (1993) show that, in fact, the distribution of completed unemployment spells in the ENEU is remarkably similar to that inferred from the distribution of incomplete spells.

A second approach is to exploit the panel nature of the ENEU data to identify the first quarter in which an individual's employment status changes (that is, when the individual finds a job) or, if the individual remains unemployed at the end of one year, to code the unemployment spell as censored. To compute the time to first job, we can assume that the job was found at the end, in the middle, or at the beginning of that interval. This corresponds to adding 3, 1.5, or 1 month(s) to incomplete unemployment spells first reported in the third quarter of 1990, plus the number of subsequent full quarters of unemployment.

Both approaches, and all three start-time assumptions, yielded similar results, namely, that the time to first job is always shorter for the trainees than for the comparison group. The assumption that the job is found at the beginning of the interval produces the lowest time to first job for the comparison group, as might be expected. The results reported below are based on the second approach,

which uses the most stringent start-time assumption (job found at the beginning of the interval).[5]

We corrected for selectivity bias by applying to the unemployed sample the same criteria used to select trainees into the program. We first estimated an equation for the probability of selection into PROBECAT, using the pooled trainee and unemployed samples. This probit model relates program participation to the criteria for which we have information: marital status, number of children and economic dependents, education, and time spent in unemployment at the selection point. We then limited the comparison group to "eligible" unemployed individuals with a high predicted probability of program participation. The cutoff point used to select these individuals was a predicted probability value of 0.6. All the employment results presented below are based on comparisons of the trainees with the "matched" (selectivity-corrected) sample of unemployed individuals who were eligible for, but did not participate in, PROBECAT.

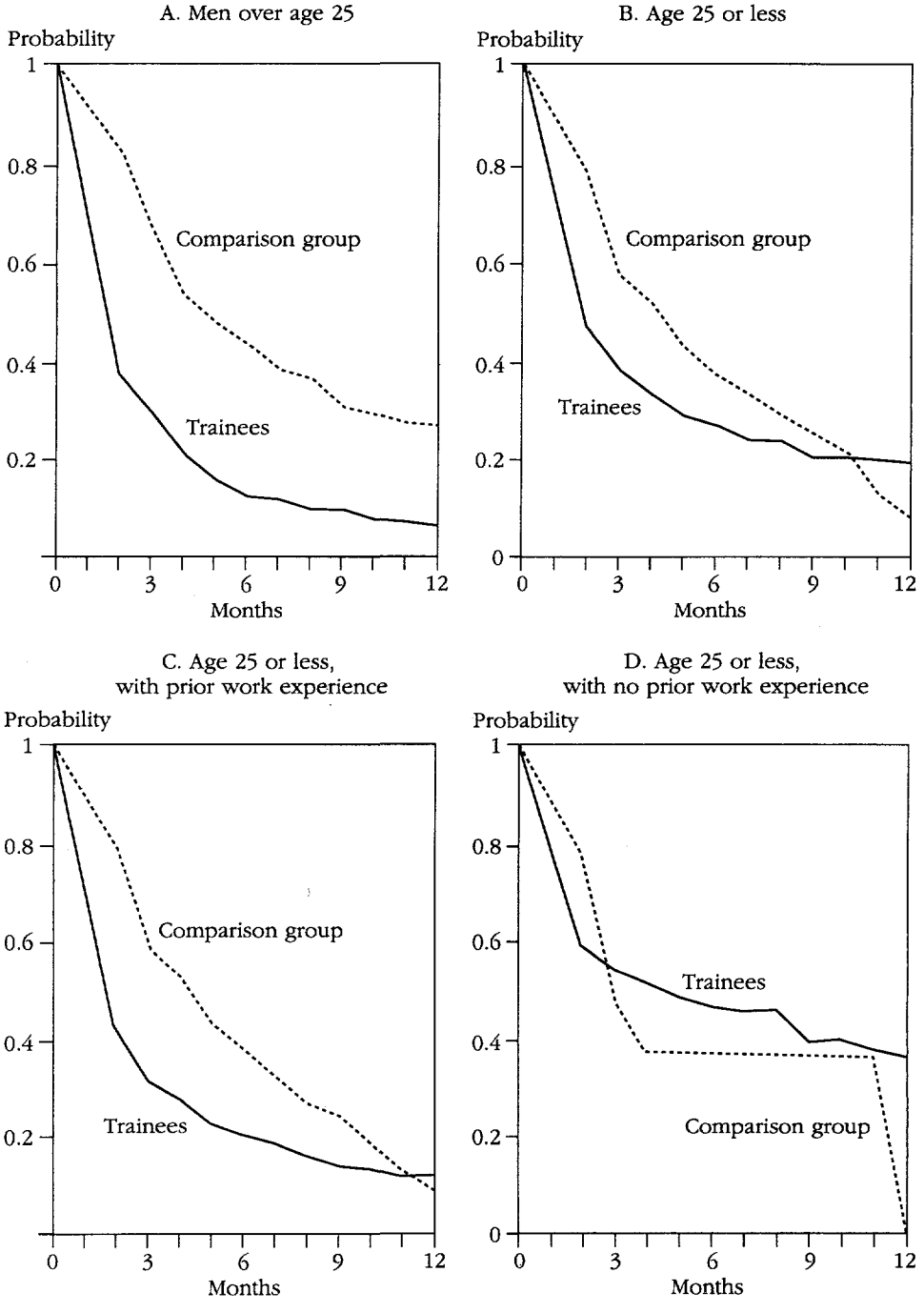
Figure 1 plots survival curves for male trainee and matched comparison groups using the raw duration data on the time to first job. These survival curves, defined as a function of time $t$ (in months), indicate the probability of remaining unemployed $t$ months after entering unemployment. We present separate survival curves for those over age 25 and for those age 25 and under. The survival curves clearly show that trainees exit unemployment more quickly than do individuals from the comparison group: at three months, 62 percent of young trainees have left unemployment compared with just 42 percent of the comparison group. The difference is more marked for the older trainees: 72 percent of them have left unemployment within three months, compared with 33 percent of the comparison group. We estimated that the average duration of unemployment for male trainees under 25 is 1.4 months shorter than that for the comparison group; for trainees over 25, the average duration of unemployment is 3.7 months shorter.

Figure 1 also shows survival curves for young men age 25 and under, both with and without previous work experience. Young trainees with work experience exit unemployment more quickly than comparable individuals in the comparison group. However, for new entrants into the labor force, these patterns are quite different, with some trainees exiting unemployment relatively quickly and others remaining unemployed for a long time. About 39 percent of young trainees without work experience remain unemployed twelve months after training completion. In contrast, all their young counterparts in the comparison group exit unemployment by twelve months. Not surprisingly, for men under age 25, we find that the average duration of unemployment for trainees exceeds that for the comparison group by 1.5 months.

With the exception of age, these graphical comparisons do not control for systematic differences in the demographic characteristics of the trainee and comparison groups. The unadjusted estimates may be misleading if unemployment

5. Results using the first approach and the more lax start-time assumptions are available on request.

Figure 1. *Survival Curves, Men*




*Note:* The sample sizes are as follows: for panel A, 119 in the comparison group and 437 trainees; for panel B, 252 in the comparison group and 444 trainees; for panel C, 121 in the comparison group and 330 trainees; and for panel D, 131 in the comparison group and 107 trainees.

*Source:* Authors' calculations.

duration is related to level of education attained or other individual and household characteristics.[6] To address this potential problem, we estimated a Cox proportional hazards model of unemployment duration on the pooled trainee and comparison group samples. This model decomposes the reemployment probabilities (the hazard rate) into a function of time (which is the same for all individuals) as well as other regressors. This regression approach allows us to investigate the impact of training on the time to first job, controlling for both individual and group differences in age, level of education, years of prior work experience, and household attributes.

Table 4 presents the Cox regression results for men. The estimated coefficients on the indicator variables for training are both positive and statistically significant in all cases, confirming the previous finding that trainees exit unemployment more quickly than their counterparts in the comparison group. The size of this estimated coefficient suggests that the average duration of unemployment for the comparison group is 30 percent longer than that for trainees.

Figure 2 plots survival curves for women in the trainee and comparison groups by age group. Like their male counterparts, female trainees appear to exit unemployment more quickly than women who did not undergo training. At three months, 50 percent of female trainees age 25 and under have found employment, compared with 32 percent of the comparison group. For the sample of young women, these differences disappear over the course of the first year; after nine months the survival curves for trainees and the comparison group are virtually identical. For the sample of older women, the difference between trainees and the comparison group increases over time so that 75 percent of trainees have left unemployment after twelve months, compared with just 47 percent of older women in the comparison group.

We also investigated the employment effects for women with different degrees of attachment to the labor force. If training is enhanced by initial skill or education endowments, we might expect training effectiveness to be diminished for women with low attachment to the labor force because of skill obsolescence (Mincer and Ofek 1982). To explore this hypothesis, we distinguished between women who had worked sometime in the six months prior to training and those who had been out of work for a longer period. In the selection process, preference was given to individuals who had been unemployed for less than three months. However, a number of participants surveyed (women in particular) reported being out of work for a longer period. Some women may have been drawn back into the labor force by the training program, and we attempted to flag them under the group that was unemployed for longer than three months. Figure 2 shows the survival curves for these two groups of women. In both cases, trainees fare better than those without training. However, consistent with the hypothesis of skill obsolescence, female trainees who have recently left em-

6. However, a parallel study shows that the only significant determinants of unemployment duration in Mexico are age and having economic dependents (Revenga and Riboud 1993).

Table 4. *Cox Regression Results for Men*

| *Independent variable* | *Regression* 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Age | −0.009 | −0.007 | −0.008 | −0.007 |
| | (−1.595) | (−1.297) | (−1.551) | (−1.261) |
| Years of schooling | −0.016 | | −0.016 | |
| | (−1.370) | | (−1.401) | |
| Number of children | −0.011 | −0.008 | −0.012 | −0.010 |
| | (−0.349) | (−0.269) | (−0.394) | (−0.307) |
| Dummy variables[a] | | | | |
| Education | | | | |
| No formal | | 0.110 | | 0.106 |
| | | (0.284) | | (0.272) |
| Primary incomplete | | −0.168 | | −0.169 |
| | | (−0.903) | | (−0.908) |
| Secondary incomplete[b] | | −0.015 | | −0.018 |
| | | (−0.116) | | (−0.142) |
| Secondary complete | | −0.139 | | −0.137 |
| | | (−1.200) | | (−1.184) |
| Higher secondary | | −0.031 | | −0.037 |
| | | (−0.262) | | (−0.309) |
| University | | −0.296 | | −0.299 |
| | | (−1.940) | | (−1.961) |
| Household head | 0.318 | 0.311 | 0.315 | 0.309 |
| | (3.233) | (3.154) | (3.203) | (3.135) |
| Prior work experience | 0.737 | 0.742 | 0.777 | 0.775 |
| | (5.593) | (5.612) | (5.565) | (5.538) |
| Training program | | | | |
| Participant | 0.355 | 0.340 | | |
| | (4.101) | (3.840) | | |
| Participant and unemployed six months or less | | | 0.395 | 0.373 |
| | | | (4.034) | (3.739) |
| Participant and unemployed more than six months | | | 0.321 | 0.312 |
| | | | (3.380) | (3.215) |

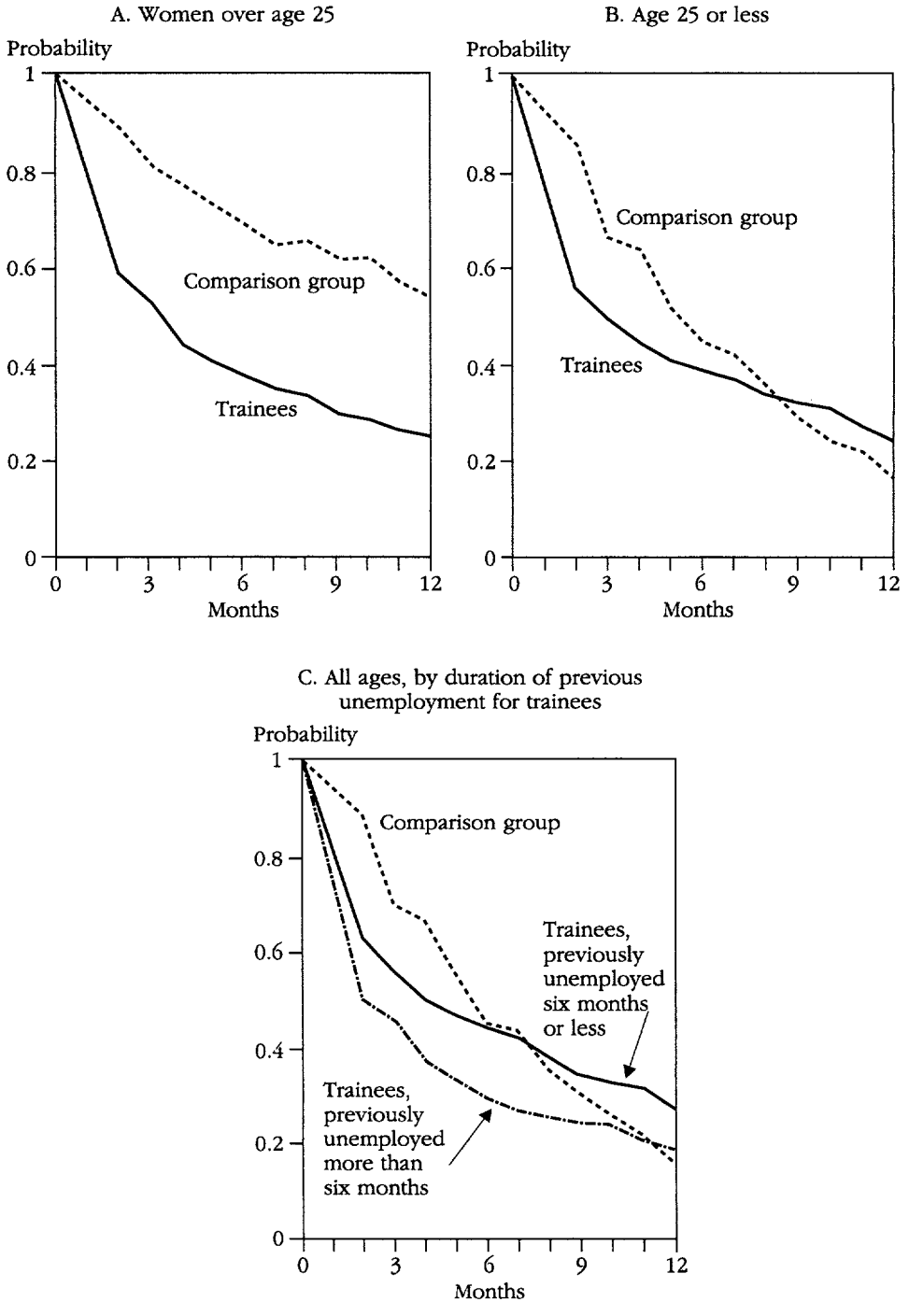
*Note*: The dependent variable is the log of duration of unemployment, in months. There were 814 observations. *t*-statistics are in parentheses.

a. = 1 if condition holds, = 0 otherwise.

b. In Mexico there are six years of secondary education after primary school. The first three years are referred to as secondary education and the second three years as higher secondary education.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

Figure 2. *Survival Curves, Women*




*Note:* The sample sizes are as follows: for panel A, 60 in the comparison group and 469 trainees; for panel B, 129 in the comparison group and 376 trainees; and for panel C, 189 in the comparison group, 493 trainees unemployeed six months or less, and 282 trainees unemployeed more than six months.

*Source:* Authors' calculations.

ployment exit unemployment after training more quickly than do trainees reentering the work force after a long inactive spell.

Table 5 presents the results for women of estimating a Cox proportional hazards model of unemployment duration for the pooled trainee–comparison group samples. The results suggest that differences in exit rates between trainees and the comparison group disappear once account is taken of several individual and household characteristics. The coefficient on the training program participant variable is close to zero. However, this result is due in large part to the differential effects of strong, compared with weak, labor force attachment, shown in figure 2. When an interaction term between training and duration of prior unemployment is included in the Cox model, the results suggest that women who enter training after a relatively short spell of unemployment exit more quickly than those who do not undergo training; those who enter training after a long spell out of the labor force exit more slowly.

### *Employment Probabilities over Time*

We also compared the employment probabilities of trainees and the comparison group over progressively longer intervals of time. The ENEU reports the labor market status of the unemployed sample at three, six, nine, and twelve months after the third quarter of 1990 (when we first observe them). For PROBECAT trainees, we used the start and end dates from their retrospective histories to define variables for labor market status for comparable intervals of time after the completion of training.

In tables 6 and 7 we begin with simple comparisons of the employment status of trainees and the unemployed sample *without* adjusting for program selection effects. These tables show the percentage of each group that reports being employed at three, six, nine, and twelve months, separately by sex and by prior work experience. Table 6 suggests that trainees, on average, are more likely than the comparison group to be employed during the year following training. For men, the difference is about 9 percentage points at three and six months and 5 percentage points at nine and twelve months. For women, the difference between trainees and the comparison group is somewhat smaller, averaging 4 to 5 percentage points over the year. In table 7 we differentiate between new labor force entrants and those with previous work experience. These figures suggest that training is much less effective for new entrants. Although trainees with work experience are usually more likely to be employed than the comparable individuals in the comparison group, trainees without work experience are slightly less likely to be employed at three months and much less likely at twelve months. A similar, but even more pronounced, pattern is found in the samples for women.

This example highlights the importance, in program evaluations, of controlling for group differences in demographic characteristics. As we noted earlier in the methodology section, simple comparisons can be very misleading if trainees differ systematically from the comparison group. In this case, the critical differ-

Table 5. *Cox Regression Results for Women*

| *Independent variable* | *Regression* 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Age | −0.014 | −0.014 | −0.011 | −0.011 |
| | (−1.866) | (−1.867) | (−1.467) | (−1.481) |
| Years of schooling | 0.005 | | 0.004 | |
| | (0.300) | | (0.240) | |
| Number of children | −0.002 | −0.007 | 0.006 | −0.0002 |
| | (−0.056) | (−0.203) | (0.160) | (−0.001) |
| Dummy variables[a] | | | | |
| Education | | | | |
| No formal | | 0.949 | | 0.876 |
| | | (1.305) | | (1.208) |
| Primary incomplete | | 0.042 | | 0.024 |
| | | (0.212) | | (0.119) |
| Secondary incomplete[b] | | −0.017 | | −0.041 |
| | | (−0.115) | | (−0.276) |
| Secondary complete | | −0.074 | | −0.109 |
| | | (−0.490) | | (−0.727) |
| Higher secondary | | 0.125 | | 0.123 |
| | | (0.787) | | (0.769) |
| University | | 0.043 | | −0.0002 |
| | | (0.187) | | (−0.001) |
| Household head | 0.399 | 0.390 | 0.457 | 0.449 |
| | (2.970) | (2.881) | (3.427) | (3.343) |
| Prior work experience | 0.507 | 0.506 | | |
| | (4.731) | (4.713) | | |
| Training program | | | | |
| Participant | 0.012 | 0.071 | | |
| | (0.098) | (0.543) | | |
| Participant and unemployed six months or less | | | 0.236 | 0.314 |
| | | | (1.631) | (2.078) |
| Participant and unemployed more than six months | | | −0.153 | −0.083 |
| | | | (−1.175) | (−0.608) |

*Note*: The dependent variable is the log of duration of unemployment, in months. There were 599 observations. *t*-statistics are in parentheses.

a: = 1 if condition holds, = 0 otherwise.

b. In Mexico there are six years of secondary education after primary school. The first three years are referred to as secondary education and the second three years as higher secondary education.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990-91 ENEU.

Table 6. *Employment Outcomes for Trainees and the Unadjusted Comparison Group*
(percentage employed)

| *Time interval (months)* | *Men* | | | *Women* | | |
|---|---|---|---|---|---|---|
| | *Trainees* | *Comparison group* | *Difference* | *Trainees* | *Comparison group* | *Difference* |
| 3 | 60 | 51 | 9 | 33 | 29 | 4 |
| 6 | 65 | 56 | 9 | 38 | 32 | 6 |
| 9 | 66 | 61 | 5 | 38 | 34 | 4 |
| 12 | 71 | 65 | 6 | 39 | 35 | 4 |

Note: The table uses the expanded definition of unemployment, in which the unemployed are defined as those under age 55 who are not working, not studying, and not retired but are able to work (not sick or disabled), regardless of whether they are actively searching for a job. Women who report being at home taking care of the house are not counted as unemployed. The time interval is the period from the first observation of the individual. These employment outcomes have not been adjusted for program selection effects. Sample sizes are 1,138 for men and 1,000 for women.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

ence between the two groups appears to be a much greater representation of new labor force entrants in the comparison group. This and other group differences, induced in part by program selection, are explicitly taken into account in the following analyses.

We estimate probit models in which the probability of employment—at three, six, and twelve months—is related to age, education, prior work experience, unemployment duration, a set of seasonal dummy variables, and an indicator variable for whether the individual participated in the PROBECAT program. Two different models are estimated. In one specification, no attempt is made to correct for selectivity bias, and the model is estimated on the pooled trainee and unadjusted comparison group samples. In the second model, the potential selectivity bias issue is addressed (as before) by pooling trainees with a "matched" comparison group—unemployed individuals with high predicted probability of program participation.

The effects of training on subsequent probabilities of employment are summarized in table 8 for men and women.[7] First, consider the results for men. The model without selectivity correction suggests that training produces a weak positive effect on the probability of employment at three months and a zero effect thereafter. Selectivity correction strengthens these results. The corrected estimates show a statistically significant effect of training on the probability of employment at three months, a smaller but still significant effect at six months, but no significant effect thereafter.

For women, the effects of training on employment have a slightly different pattern. When a continuous measure of education is used, the estimates without selectivity correction show that training has no statistically significant impact on the probability of employment. However, in specifications that include dummy

7. The probit regression results on which these estimates are taken are available from the authors.

Table 7. *Employment Outcomes and Work Experience for Trainees and the Unadjusted Comparison Group*
(percentage employed)

| *Time interval (months)* | *Trainees* | | | *Comparison group* | | |
|---|---|---|---|---|---|---|
| | *With work experience* | *No work experience* | *Difference* | *With work experience* | *No work experience* | *Difference* |
| | | | *Men* | | | |
| 3 | 65.0 | 32.2 | 5.0 | 60.0 | 34.6 | −2.4 |
| 6 | 70.6 | 35.7 | 0.8 | 69.8 | 33.1 | 2.6 |
| 9 | 71.2 | 37.4 | −2.0 | 73.2 | 40.4 | −3.0 |
| 12 | 76.4 | 40.9 | 1.9 | 74.5 | 49.3 | −8.4 |
| | | | *Women* | | | |
| 3 | 43.3 | 15.1 | 10.9 | 32.4 | 23.8 | −8.7 |
| 6 | 50.0 | 17.4 | 14.8 | 35.2 | 28.6 | −11.2 |
| 9 | 50.0 | 19.0 | 16.7 | 33.3 | 34.5 | −15.5 |
| 12 | 50.2 | 21.6 | 15.9 | 34.3 | 36.9 | −15.3 |

*Note*: The table uses the expanded definition of unemployment, in which the unemployed are defined as those under age 55 who are not working, not studying, and not retired but are able to work (not sick or disabled), regardless of whether they are actively searching for a job. Women who report being at home taking care of the house are not counted as unemployed. The time interval is the period from the first observation of the individual. These employment outcomes have not been adjusted for program selection effects. Sample sizes are 1,138 for men and 1,000 for women.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

variables for different levels of education, the selectivity-corrected estimates show a positive, statistically significant training effect. Prior experience also appears to be an important determinant of whether training is effective for women. In results not reported here, we find that training has a significantly positive effect on employment at three, six, and twelve months for women with prior work experience but has a negative and statistically significant training effect at three and twelve months for those without work experience.[8]

To summarize, participation in PROBECAT appears to affect subsequent employment probabilities of trainees but does so in quite different ways for men and women. For men, it increases their probability of being employed up to six months after the program but does not have an effect thereafter; this result, taken together with the previous finding that male trainees find jobs more quickly, suggests that they tend to be employed for a greater proportion of the post-training period than the men in the comparison group. For women, training appears to raise employment probabilities only for those with prior work experience, but, unlike with men, this positive training effect persists over the year. In contrast, women without any work experience benefit relatively little, if at all, from training.

8. A substantial proportion of the female trainee sample had no prior work experience. To investigate the potential importance of this variable, we modified the specification of the employment equation to include an interaction term between training and a dummy variable for prior work experience. The results of this model specification are available from the authors.

Table 8. *The Estimated Effects of Training on Employment*
(difference in predicted employment probabilities)

| Time interval (months) | Model with education dummy variables: Unadjusted comparison group | Model with education dummy variables: Adjusted comparison group | Model with continuous school variable: Unadjusted comparison group | Model with continuous school variable: Adjusted comparison group |
|---|---|---|---|---|
| | *Men* | | | |
| 3 | 0.055*** | 0.084** | 0.072** | 0.098** |
| 6 | 0.011 | 0.055*** | 0.033 | 0.077** |
| 12 | 0.008 | 0.042 | 0.015 | 0.050 |
| Sample size | 1,138 | 943 | 1,138 | 943 |
| | *Women* | | | |
| 3 | 0.048 | 0.089** | 0.052 | 0.054 |
| 6 | 0.069** | 0.130** | 0.055 | 0.066*** |
| 12 | 0.061*** | 0.109** | 0.053 | 0.056 |
| Sample size | 1,000 | 916 | 1,000 | 916 |

** Denotes statistical significance at the 0.05 level.
*** Denotes statistical significance at the 0.10 level.
*Note*: Values are the estimated differences in predicted employment probabilities due to participation in training. The "unadjusted comparison group" columns are based on a probit model estimated on the pooled trainee and unadjusted comparison groups. The "adjusted comparison group" columns are based on a probit model with selectivity correction in which the matched comparison group is selected from the comparison group sample according to selection criteria used for trainees.
*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990-91 ENEU.

## IV. Monthly Earnings, Hours of Work, and Hourly Wages

The above analysis suggests that PROBECAT training has a positive but moderate impact on the post-training employment rates of participants. The next step is to investigate whether PROBECAT training also translates into an increase in the post-training earnings of participants.

### *Data and Summary Statistics*

The data set used was constructed from the retrospective PROBECAT survey and a comparison group drawn from two ENEU unemployed cohorts. The first cohort included individuals who were unemployed in the third quarter of 1990 and were tracked for twelve months. The second cohort included individuals who became unemployed in the fourth quarter of 1990 and were not in the first cohort; for this latter cohort, only nine months of information were available. We pooled all observations reporting positive (and usable) earnings at any time during the period of the PROBECAT survey and during the twelve- or nine-month interval in the case of the comparison group.[9] The data set thus contained multiple observations on each individual—observations for every job spell expe-

9. We define usable data as less than 5 million pesos of positive monthly earnings and less than 85 hours of work a week.

Table 9. *Monthly Salary, Hours Worked, and Hourly Wage for Trainees and the Comparison Group*
(average)

| | *Men* | | *Women* | |
|---|---|---|---|---|
| *Outcome variable* | *Trainees* | *Comparison group* | *Trainees* | *Comparison group* |
| Monthly earnings (1,000 pesos) | 681.59 | 637.67 | 531.85 | 571.52 |
| Work hours per week | 45.81 | 43.59 | 42.77 | 39.51 |
| Hourly wage (1,000 pesos) | 3.984 | 4.016 | 3.476 | 4.198 |
| Sample size | 1,212 | 1,051 | 681 | 300 |

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

rienced by trainees and for every quarter in which individuals in the comparison group were observed to be employed.

The final data set contained 1,212 trainee observations and 1,051 comparison observations for men, and 681 trainee observations and 300 comparison observations for women. To accommodate the specific structure of this data set, we created (and included) two kinds of variables. The first is a variable for the number of months between the date salaries are reported and time $t_0$, which is either the completion of training or the initial date of unemployment for the comparison group. The second is a set of quarterly dummy variables to account for inflation in salaries over the base period.

Means of the three outcome variables for men and women are reported in table 9, separately for the trainee and comparison samples. On average, male trainees report monthly earnings of approximately 682,000 pesos, compared with 638,000 pesos for the comparison group. In other words, earnings of trainees are about 7 percent higher than those of the comparison group. However, higher earnings may partly reflect inflation because trainee salaries include those reported in the first quarter of 1992, whereas comparison group salaries end in the third quarter of 1991. Trainees also report slightly higher hours worked—45.8 hours a week compared with 43.6 hours for the comparison group. Finally, the hourly wage of 3,984 pesos for trainees is slightly lower than the 4,016 pesos estimated for the comparison group.

In the women's sample, trainees report monthly earnings that are about 7 percent lower than those of the comparison group: approximately 532,000 pesos versus 572,000 pesos for trainees and the comparison group, respectively. Like their male trainee counterparts, women trainees who worked did so for three hours longer a week than women in the comparison group. However, their hourly wage was about 700 pesos less than the wage received by the women in the comparison group.

### *Overall Program Effects*

We analyzed the separate effects of training on the logarithm of monthly earnings, hours worked per week, and the logarithm of hourly wages. Each of these outcome measures was regressed on a vector of explanatory variables,

including a quadratic measure of potential work experience, level of education, prior work experience, unemployment duration at time 0, quarterly dummy variables, and an indicator variable for whether the individual was a participant in PROBECAT training. We also experimented with interaction terms between training and levels of education to see if training effects would vary across different educational levels (the results are reported below).

In this set of analyses, we follow the statistical adjustment suggested by Heckman (1979) to correct for selectivity bias from nonrandom selection into the training program. As before, we first estimated a probit model that related program participation to the selection criteria for which we had data—marital status, number of children, education, and time unemployed prior to training. We then used the probit estimates to compute an inverse Mills ratio for all individuals—both trainees and the unemployed comparison group—and included this variable as a regressor in the outcome equations to correct for selectivity bias. This adjustment, however, does not address a potentially important second source of bias that arises because earnings outcomes are observed only if the person has a job. We note, but defer to future research, the difficult task of jointly modeling the two sources of selectivity bias.

Table 10 summarizes the overall impacts of program participation for men and women.[10] The first model specification in table 10 is a simple ordinary least squares (OLS) regression. The second is a model that corrects for nonrandom selection into the PROBECAT program. The reported coefficients for monthly salary and hourly wage may be interpreted as the average percentage change in the variables attributable to the individual's participation in PROBECAT. The coefficients for hours of work per week are mean changes attributable to participation in training.

In the raw data, male trainees reported monthly earnings that, on average, were 7 percent higher than those of the comparison group. Results of the simple OLS model specification in table 10 indicate that, without selectivity correction for program participation, the monthly earnings of male trainees are 10.8 percent lower than those of the comparison group. This result suggests that the two groups have very different attributes. In the model that corrects for selectivity bias (and these group differences), this earnings differential is now reversed. The selectivity-corrected estimates show that the monthly earnings of trainees are 17.7 percent higher than the earnings of the comparison group, a difference that is statistically significant.

The results also suggest that training is associated with increased labor supply for the men's sample. In the raw data, trainees reported working about two hours more a week than individuals in the comparison group. The estimated coefficient for hours of work in the regression without selectivity correction (simple OLS) reveals no significant differences between the two groups in the

10. The full set of results is available on request.

Table 10. *Summary of the Effects of Participation in the Training Program*

| | *Men* | | *Women* | |
|---|---|---|---|---|
| *Variable* | *Simple OLS model* | *OLS model corrected for selectivity bias* | *Simple OLS model* | *OLS model corrected for selectivity bias* |
| Log of monthly earnings | −0.108 | 0.177 | −0.122 | 0.033 |
| | (−3.26) | (2.19) | (−2.42) | (0.25) |
| Work hours per week | 0.978 | 7.796 | 4.484 | 6.234 |
| | (1.32) | (4.32) | (4.19) | (2.06) |
| Log of hourly wage | −0.095 | −0.007 | −0.261 | −0.105 |
| | (−2.61) | −0.07 | (−4.48) | (−0.68) |

*Note*: *t*-statistics are in parentheses.
*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

number of hours worked. However, the model corrected for selectivity bias shows that trainees supply, on average, 7.8 more hours per week than the comparison group once a correction is included for selectivity. Similarly, in the hourly wage results, selectivity correction reduces the negative effect of PROBECAT on hourly wage as compared with the simple OLS model. In fact, the final outcome of the selectivity correction is that there is no significant difference in hourly wages between the two groups. Together, these results suggest that training, on average, raises monthly earnings of male trainees through a greater supply of hours worked per week, not through higher hourly wages.

A similar pattern of training effects is found for women. In the aggregate data, women trainees received lower monthly earnings and hourly wages, but worked more hours per week, than the comparison group. In the OLS model, these program effects on earnings, hourly wage, and hours worked are generally statistically significant. In the model that corrects for selectivity bias, however, many of these differences disappear. The only statistically significant effect of PROBECAT is in the results for hours of work, which suggest that women trainees work approximately six hours more a week than women in the comparison group.

### *Training Effects by Level of Education*

Thus far, we have assumed implicitly that program effects are invariant across different groups of trainees. This may not be a good assumption if the effectiveness of training is shaped by the initial skill endowments that trainees bring to the program. If education helps trainees get more out of training, we would expect training effectiveness to increase (at least over some range) with level of education. We addressed this possibility by including interaction terms between training and indicator variables for each level of education. As before, a separate set of dummy variables for each level (except one) was included to control for education effects common to both trainee and comparison groups.

Table 11 presents the results of estimating these expanded model specifications for men. To conserve space, we report results only for the models estimated

Table 11. *Results of Estimating Expanded Model Specifications for Men*

| | *Dependent variable* | | |
|---|---|---|---|
| *Explanatory variable* | *Log monthly salary* | *Work hours per week* | *Log hourly wage* |
| Constant | 5.726* | 42.587* | 0.574* |
| General experience | 0.023* | 0.134 | 0.019* |
| Experience-squared | −0.000* | −0.004* | −0.000* |
| *Education* | | | |
| No formal | −0.203* | −1.024 | −0.194 |
| Primary incomplete | −0.005 | −0.797 | 0.003 |
| Secondary incomplete | 0.062 | 0.430 | 0.026 |
| Secondary complete | 0.165* | −1.790 | 0.186* |
| Higher secondary | 0.279* | −0.704 | 0.295* |
| University | 0.571* | −0.455 | 0.555* |
| *Education-training interaction* | | | |
| No formal | −0.058 | 14.820 | −0.485 |
| Primary incomplete | −0.063 | 9.101* | −0.270* |
| Primary complete | 0.212* | 8.409* | −0.002 |
| Secondary incomplete[a] | 0.267* | 5.424* | 0.161 |
| Secondary complete | 0.199* | 10.132* | −0.039 |
| Higher secondary | 0.171** | 7.848* | −0.029 |
| University | 0.045 | 2.769 | 0.032 |
| Inverse Mills ratio | −0.179* | −4.650* | −0.049 |
| $\bar{R}^2$ | 0.157 | 0.054 | 0.145 |
| Sample size | 2,330 | 2,271 | 2,271 |

* Statistically significant at the 0.01 level.

** Statistically significant at the 0.05 level.

*Note*: These results are for the OLS model estimated with the selectivity correction. The regressions include duration of previous unemployment, self-employment status, and time dummy variables.

a. In Mexico there are six years of secondary education following primary school. The first three years are referred to as secondary education and the second three years as higher secondary education.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

with the selectivity correction. Table 11 suggests that training has positive and statistically significant effects on monthly earnings and hours of work by level of education. In general, these earnings-and-hours effects exhibit an inverted-U pattern, being lowest for the least-educated men, rising with years of education to a peak at the secondary school level (seven to nine years of schooling), and then declining for the most-educated individuals (those with postsecondary education). For women (not reported here), the education-training interaction terms are generally insignificant, suggesting that training effects on these outcomes are broadly similar across educational levels.

The results for men—positive impacts on monthly earnings and hours worked but no systematic effect on hourly wages—raise questions about whether training actually increases productivity, in which case one might expect higher hourly wages, or whether it raises earnings by inducing greater work effort among trainees.[11] We believe the answer lies in the kinds of jobs that trainees find upon completing training. The raw data suggest that, in relation to the comparison

11. These results do not appear to be specific to Mexico. An anonymous referee points out that similar findings have been reported in evaluation studies of U.S. training programs.

group, a higher proportion of trainees eventually find jobs in large enterprises. For the sample as a whole, employment in large firms is associated with longer hours of work per week and higher monthly salaries, which may partially explain the results that we find. To explore this hypothesis more rigorously, we estimated an ordered logit model for the probability of employment in ten (increasingly larger) firm-size categories. As regressors, we included measures of experience, duration of previous unemployment, time dummy variables, and an indicator variable for participation in PROBECAT.

Table 12 reports the results for three different model specifications: training by itself, training interacted with a quadratic measure of years of education, and training interacted with indicator variables for each level of education. All three specifications suggest that trainees are more likely to find jobs in larger firms than are individuals in the comparison group. The fully interacted model specification reveals an inverted-U pattern of effects by level of education, similar to the previous findings for earnings and hours of work. In short, PROBECAT appears to raise trainees' monthly earnings and hours of work by facilitating their entry into larger firms offering higher pay and more stable, full-time employment. PROBECAT may achieve this result either by retraining the unemployed in skills for which there is demand—that is, through a matching effect—or by making them more trainable—that is, by providing them with learning skills. PROBECAT may also indirectly affect future earnings potential by placing trainees in larger firms that tend to provide more on-the-job training.[12] The trainees will have to be followed over a longer time period for us to verify this hypothesis.

To summarize, the results suggest that participation in PROBECAT increases monthly earnings of male trainees and that this occurs primarily because of their increased hours of work. The disaggregated analysis by educational level reveals that this effect varies with the level of education attained. The effects of training on monthly earnings are largest for those with secondary education (seven to nine years of schooling). The effects of hours of work are large and positive for most groups, except possibly for those with the lowest and highest levels of education. For women, there is some evidence that work hours are increased by training, but these do not translate into higher monthly earnings. Unlike their male counterparts, the earnings-and-hours effects of training do not vary by level of education. The results also suggest that, for men, program participation increases the probability of finding employment in a large firm. Because large firms tend to pay higher wages, provide more training opportunities, and thus have steeper earnings profiles, finding a job in a large firm is likely to imply increased earnings opportunities over time for trainees in relation to the comparison group.

12. Estimates based on the 1988 National Employment Survey show that the proportion of workers receiving training in the workplace varies from 5 percent in microenterprises to 23 percent in large firms employing more than 250 workers.

Table 12. *The Probability for Men of Post-Training Employment in Larger Firms*

| Explanatory variable | Maximum likelihood ordered logit model 1 | 2 | 3 |
|---|---|---|---|
| *Education* | | | |
| No formal | −0.262 | −0.163 | −0.229 |
| Primary incomplete | −0.014 | 0.102 | −0.051 |
| Secondary incomplete[a] | 0.289* | 0.236** | −0.092 |
| Secondary complete | 0.202** | 0.189 | 0.095 |
| Higher secondary | 0.517* | 0.633* | 0.541* |
| University | 0.482* | 0.986* | 0.845* |
| *Trainee* | | | |
| Dummy (1,0) | 0.786* | | |
| Interaction with education | | 0.246* | |
| Education-squared | | −0.017* | |
| No formal education | | | −0.236 |
| Primary incomplete | | | 0.827* |
| Primary complete | | | 0.707* |
| Secondary incomplete | | | 1.314* |
| Secondary complete | | | 0.893* |
| Higher secondary | | | 0.688* |
| University | | | −0.057 |
| Inverse Mills ratio | −0.354* | −0.282* | −0.338* |

* Statistically significant at the 0.01 level.

** Statistically significant at the 0.05 level.

*Note*: The dependent variable is firm-size category. The model estimates the probability of employment in ten (increasingly larger) firm-size categories. The regressions include quadratic specification of general experience, the duration of previous unemployment, and time dummy variables. The sample size is 2,330 observations.

a. In Mexico there are six years of secondary education following primary school. The first three years are referred to as secondary education and the second three years as higher secondary education.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

## V. The Cost-Effectiveness of PROBECAT

What do these findings imply about the cost-effectiveness of PROBECAT? To answer this question, we focus on two of the more significant labor market outcomes identified in the previous analyses: first, the impact of program participation on the speed with which trainees find jobs, and, second, the impact of program participation on monthly earnings. These two impacts, and their implications for the benefit streams associated with program participation, are combined with cost estimates to arrive at some back-of-the-envelope calculations on the cost-effectiveness of PROBECAT.

We include the direct training costs as well as the indirect costs associated with participation in the program. Direct costs are costs for instructors, training materials, and program administration. From data provided to us by the Secretaría del Trabajo y Previsión Social, the average operating cost per course completer in 1991 was about 350,400 pesos. Indirect costs are measured in terms of search time forgone by joining the training program. We assume that at time $t_0$ each unemployed worker faces two possible strategies: immediately initiate a job

search (the strategy for the comparison group), or enter a training course and thus delay a job search by the length of the course (the strategy for the trainees). From our previous analysis, we know that a job search after training is shorter on average (by about 2.5 months) than a job search without training (the case for the comparison group). However, we must also take into account the costs of deferring a search when trainees participate in PROBECAT. Thus, we calculate indirect costs for trainees by adding to search time the time spent in training (an average of 2.9 months). The benefit measures are calculated from the previous estimates of the effects of training on monthly earnings. Monthly earnings are predicted for trainees and the comparison group using sample means of all regressors.

Table 13 summarizes the calculation of these cost and benefit measures. The first and third columns show the mean duration of search for the trainee and comparison groups, respectively. The difference between the two columns measures the decrease in search time (the employment effect) attributable to the program. The second column adds the average duration of training (2.9 months) to the search time for trainees. The fourth column reports the difference in total time out of work for the two groups. For male trainees, participation in PROBECAT increases the total time to first job (search plus training time) by 0.4 months in relation to the comparison group, whereas for female trainees, training increases the total time to first job by about 1 month. Note, however, that the figure for women hides very substantial differences by demographic group. For women with prior work experience, training reduces the time to first job by 1.1 months (even when we include job search delay because of training).

On the benefit side, the fifth column of table 13 reports the estimated wage effect of training. The sixth column reports the corresponding predicted monthly wage evaluated at sample means. The positive wage impact attributable to training is quite large for men, averaging about 152,000 pesos. Although predicted wages for female trainees are slightly higher than for the comparison group, the difference is not statistically significant.

The calculation of the net benefits of the training program is shown in table 14. The first column shows the direct, average cost of providing training—350,400 pesos per trainee. (The training stipend is not taken into account as it is simply an income transfer and not an economic cost.) The second column is the indirect cost of training, which is the monetized value of incremental job search costs (forgone earnings) associated with attending training, valued at the average wage of the comparison group (that is, the wage trainees would have received had they not participated in PROBECAT). On average, these indirect costs are about 196,000 pesos for men and 435,000 pesos for women. The fourth and fifth columns summarize the benefits of training associated with increased wages over three months and twelve months, respectively. The final two columns show the net benefit (benefits minus costs) associated with participation in the training program.

Table 13. *Summary of the Effects of the Training Program*

| Trainee type | Average duration of search for employment (months): Trainees: Excluding training period | Average duration of search for employment (months): Trainees: Including training period[a] | Average duration of search for employment (months): Comparison group | Difference in time to first job for trainees and comparison group (months) | Wage effect[b] | Predicted monthly wage[c] (thousands of pesos): Trainee | Predicted monthly wage[c] (thousands of pesos): Comparison group |
|---|---|---|---|---|---|---|---|
| Men | 4.0 | 6.9 | 6.5 | 0.4 | 0.27 | 642 | 490 |
| Women | 5.9 | 8.8 | 7.8 | 1.0 | 0.02 | 444 | 435 |

*Note*: The samples included 881 male trainees and 845 female trainees and 371 men and 189 women in the comparison group.

a. The average duration of the training period is 2.9 months.

b. The wage effect, $\ln W_t - \ln W_c$, is the coefficient on the training variable from a regression of log monthly wages on experience, experience squared, education, quarterly dummies, self-employment status, duration of unemployment prior to training or first observation, prior experience, and interactions for training and education status and for training and age.

c. Predicted monthly wages at sample means from the same wage regression as the wage effect.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

Table 14. *Costs and Benefits of the Training Program*
(thousands of pesos per trainee)

| | *Costs* | | | *Benefits (increase in monthly wage compared with that of comparison group)* | | *Net benefits* | |
|---|---|---|---|---|---|---|---|
| *Trainee type* | *Direct costs* | *Search costs*[a] | *Total costs* | *Over 3 months* | *Over 12 months* | *Over 3 months* | *Over 12 months* |
| Men | 350.4 | 196.0 | 546.4 | 456.0 | 1,824.0 | −90.4 | 1,277.6 |
| Women | 350.4 | 435.0 | 785.4 | 27.0 | 108.0 | −758.4 | −677.4 |

*Note*: The samples included 881 men and 845 women.

a. The additional time trainees take to find a job because of training times the opportunity cost of that time, which equals the comparison group's wage, that is, the wage trainees would have received without training.

*Source*: Authors' calculations from the 1992 PROBECAT survey and the 1990–91 ENEU.

These estimates, although very crude, nonetheless suggest the following findings: for men, the benefits of program participation outweigh the costs within a year of finishing training; for women as a whole, the costs exceed the benefits of training. As the previous analyses showed, however, there are substantial differences in outcomes depending upon whether women enter training with or without prior work experience. For women with work experience, benefits from earlier employment clearly offset the costs of participation in the program.

## VI. Conclusions

This evaluation of Mexico's PROBECAT sought to measure the impact of training on the employment and earnings of participants. Training outcomes were estimated by comparing PROBECAT trainees with a comparison group of unemployed individuals.

On the whole, the results suggest that PROBECAT was fairly effective in shortening the duration of unemployment for certain target groups, namely the trainees with prior work experience (both men and women). It also appeared to have improved the likelihood of employment for participants over a longer period of time. Compared with those who did not participate in the program, male trainees were more likely to be employed three and six months after training; female trainees with prior work experience also benefited, but unlike the case with male trainees, these positive employment effects appeared to have persisted over a full year. As for earnings, the evaluation suggests that program participation raised the post-training earnings of men but not of women. For male trainees, these earnings effects varied systematically by level of education, being greatest for those with seven to nine years of schooling. Finally, for both men and women, training induced an increase in the number of hours worked per week.

The disparity of training outcomes across different demographic groups indicated that the unemployed constitute a very heterogenous group and, conse-

quently, that eligibility criteria used for program participation can have important implications for the program's cost-effectiveness. In the specific case of Mexico, the analyses suggested that PROBECAT's selection criteria should be modified to target those demographic groups most likely to benefit from the program—the unemployed with prior work experience, slightly older workers (over 25 years old), and those with six to twelve years of schooling. For certain other groups—for example, the young, new entrants into the labor force, and those with low levels of education—it may be more appropriate for the government to provide adult basic education, facilitate return to school for the young, or introduce firm-based apprenticeship programs to give work experience to new entrants in the labor market.

More broadly, our study confirms that program evaluation results can be very sensitive to the way in which training effects are measured. One key source of bias is that arising from nonrandom selection of participants into the training program. In our evaluation of Mexico's PROBECAT program, we sought to correct for this one source of selectivity bias by using a variety of statistical methodologies. Several statistical issues remain, and future evaluations should endeavor to address them both through collection of better comparison group data and through more rigorous econometric modeling. These evaluations should also focus on other dimensions of PROBECAT not investigated here—training duration, type of training, the mix of theory and practice, and the relative effectiveness of different training providers.

## REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Bassi, Laurie. 1983. "The Effect of CETA on the Postprogram Earnings of Participants." *Journal of Human Resources* 18(4):539–56.

Clark, Kim B., and Lawrence H. Summers. 1979. "Labor Market Dynamics and Unemployment: A Reconsideration." *Brookings Papers on Economic Activity* 1:13–60.

Card, David, and Daniel Sullivan. 1988. "Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment." *Econometrica* 56(3, May):497–530.

Carlson, Samuel. 1991. *Mexico Labor Retraining Program: Poverty Alleviation and Contribution to Growth.* LATHR Report 6. World Bank, Latin America and the Caribbean Technical Department, Regional Studies Program, Washington, D.C.

Fraker, Thomas, and Rebecca Maynard. 1985. "The Use of Comparison Group Designs in Evaluations of Employment-related Programs." N.J.: Mathematica Policy Research, Princeton. Processed.

Geraci, Vincent. 1984. "Short-Term Indicators of Job Training Program Effects on Long-Term Participant Earnings." Report prepared for the U.S. Department of Labor. Washington, D.C. Processed.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):154–61.

Heckman, James, and V. Joseph Hotz. 1987. "Are Classical Experiments Necessary for Evaluating the Impact of Manpower Training Programs? A Critical Assessment." In *Proceedings of the Industrial Relations Research Association.*

———. 1989. "Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* (408):862–74.

Jimenez, Emmanuel, and Bernardo Kugler. 1987. "The Earnings Impact of Training Duration in a Developing Country." *The Journal of Human Resources* 22(5, Spring):228–47.

LaLonde, Robert. 1984. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." Working Paper 183. Princeton University, Industrial Relations Section, Princeton, N.J. Processed.

Levitan, Sar A. 1992. "Evaluation of Federal Social Programs: An Uncertain Impact." Occasional Paper. George Washington University, Center for Social Policy Studies, Washington, D.C. Processed.

Mincer, Jacob, and Haim Ofek. 1982. "Interrupted Work Careers, Depreciation, and Restoration of Human Capital." *Journal of Human Resources* 17(1):3–24.

Revenga, Ana, and Michelle Riboud. 1993. "Unemployment in Mexico: An Analysis of Its Characteristics and Determinants." World Bank Policy Research Working Paper 1230. Washington, D.C. Processed.

Salant, Stephen. 1977. "Search Theory and Duration Data: A Theory of Sorts." *Quarterly Journal of Economics* 91(1):39–57.

Secretaría de Trabajo y Previsión Social, Mexico. 1988, 1989, 1990. *Reporte de Seguimiento del Programa de Becas de Capacitación para Trabajadores.* Mexico City.

Summers, Lawrence H. 1986. "Why Is the Unemployment Rate So Very High near Full Employment?" *Brookings Papers on Economic Activity* 0(2):339–83.

Westat, Inc. 1981. "Continuous Longitudinal Manpower Survey Net Impact Report No. 1." Report prepared for the U.S. Department of Labor. Washington, D.C. Processed.

———. 1984. "Summary of Net Impact Results." Report prepared for the U.S. Department of Labor. Washington, D.C. Processed.

# The Distribution of Subsidies through Public Health Services in Indonesia, 1978–87

Dominique van de Walle

*Indonesia has made great progress during the past fifteen years in enhancing the command of the poor over privately provided goods, such as food, clothing, and housing. Has similar progress been made in improving their access to publicly provided social services? The article looks at how the use of health services and the incidence of subsidies in the health sector varied across socioeconomic groups in Indonesia in 1987. It also examines how the distributions of utilization and subsidies altered between 1978 and 1987. The findings indicate that changes in utilization patterns and in the incidence of subsidies have been pro-poor. Disparities in access and utilization have diminished. However, public spending on health care is not yet well targeted.*

There is an emerging consensus that poverty reduction and human development call for the expansion of access to certain publicly provided social services, such as basic education and health care.[1] Both privately and publicly provided goods and services matter for well-being. But increasing concern about this role of the public sector raises the issue of how the benefits of public spending are distributed.

Many public services are publicly provided private goods, and for them, utilization is the key determinant of benefits derived. Measuring utilization clearly requires going beyond aggregate social indicators such as infant mortality rates. Household-level data sets reveal how the utilization of, for example, public health services varies with other aspects of living standards, such as consumption of private goods, and with other relevant household characteristics, such as urban or rural location and region of residence.

Patterns of utilization incidence are the outcome of the interlinkage of public sector inputs and household behavioral responses. At the household level, deci-

1. This view has been recently articulated in Drèze and Sen (1989), UNDP (1990), World Bank (1990). On the relative importance of income poverty reduction and public services in promoting human development, see Anand and Ravallion (1993).

Dominique van de Walle is with the Policy Research Department at the World Bank. The author gratefully acknowledges support from RPO 67642. The author thanks Anupa Bhaumik for her research assistance; Anil Deolalikar and Benu Bedani for providing tabulated data; and Shankar Acharya, Emmanuel Jimenez, Paul Gertler, Samuel Lieberman, Kim Nead, Nicholas Prescott, Martin Ravallion, and the referees for helpful comments.

sions are influenced by household endowments, the prices faced, and various exogenous characteristics of households and individuals. The key determinants in the choice of a health facility are the full cost of using the facility (comprising the price charged, transport costs, and forgone income), the quality of the medical care, and any disutility incurred. Utilization and distributional outcomes are thus influenced by public sector decisions affecting these variables, including aggregate resources spent on the sector, allocation within the sector, the degree of private financing, pricing policies, and the organization of sector inputs.

This article looks at three specific empirical questions related to the public health care system in Indonesia. First, how does the utilization of different services vary by household living standards? A key point here is whether there is evidence of self-selection: do the nonpoor opt out of the public health care system? Second, what is the combined effect of utilization and pricing policy in determining the incidence of public health care subsidies? Third, how have the answers to the first and second questions changed over time?

These questions are particularly pertinent for Indonesia. The country made great progress in alleviating income poverty during the 1980s; the incidence of poverty is estimated to have declined from 40 percent in 1980 to 22 percent in 1987 (Ravallion and Huppi 1991; World Bank 1991b). But there has been some concern that improvements in certain social indicators (such as infant mortality rates and life expectancy) have not kept pace with those in the poverty measures. In particular, have increases in the incomes of the poor been commensurately matched by greater access to and utilization of health and education services (World Bank 1991a)?

The period from the late 1970s through the 1980s in Indonesia is particularly interesting because it coincides, first, with substantial declines in income poverty and, second, with considerable emphasis by the government on primary health care (World Bank 1991a, 1991b; Yahya and Roesin 1990). Large investments were made and substantial initiatives were begun in the primary health care system, including the integrated family planning and health post (*posyandu*) system. Progress in these areas may have come under threat in the mid-1980s, when Indonesia sustained various external shocks. The shocks led to substantial deterioration in the external terms of trade and a subsequent macroeconomic adjustment program involving, among other things, cuts in public expenditures, although (in relation to other sectors) spending on social services appears to have been somewhat protected (see World Bank 1993).

A limited set of policy instruments is typically available for alleviating poverty in developing countries. A long-standing concern in Indonesia has been the effectiveness of health sector expenditures in reaching the poor. Of special interest is identification of the intrasector services and facilities that can best be used to target in-kind transfers to the poor.

Section I discusses Indonesia's health care system and summarizes key policy issues of relevance for the study. The main methodological issues are addressed

in section II. Section III looks at the utilization of health care facilities across various groups in 1987 and how it has changed since 1978. Section IV examines the incidence of public expenditures in the health sector, again starting with an analysis of the situation in 1987 and then turning to how it has changed since 1978. Finally, section V offers conclusions.

## I. The Setting

Indonesia's health care problems are dominated by communicable diseases (such as malaria and tuberculosis), respiratory and diarrheal diseases, and injuries. Maternal and perinatal morbidity is high, as are anemia and vitamin and micronutrient deficiencies. Indonesia's health profile is widely seen as requiring a significant continuing role for public intervention.

Overall public spending in Indonesia is very low in the health sector (Griffin 1992; World Bank 1991a). Although it has risen considerably from its 1975 level of about 0.2 percent of gross national product (GNP), in 1985 it still remained under 1 percent of GNP—below the Asian mean of 1.3 percent of GNP (Griffin 1992). In addition, government per capita spending on health declined by about 25 percent between 1982/83 and 1987/88, although it has increased again since then (Ministry of Health 1991b). At least half of total public outlays are disbursed to hospitals. Private spending accounts for more than half of total expenditures on health.

District, provincial, and central governments all contribute to the funding of Indonesia's public health care system. However, the center is by far the main source of funding through its routine and development budgets and various grants and subsidies. The political economy of health sector budget allocations to the regions is unclear. Most of the allocations seem to be dictated by an underlying view of egalitarianism with respect to inputs. But the result is not necessarily an equitable interregional distribution of outcomes. For example, INPRES grants—central transfers to the regions intended for primary health care—are based on one of three criteria: population size, number of villages, or equal absolute amounts to each region (World Bank 1993). The distribution of funding can also be partly explained by past investment decisions. Large, specialized hospitals, built at a time when development dogma championed such investments, engender important commitments to future recurrent expenditures.

Whatever the political economy dynamics, outcomes in the way resources are allocated across geographic areas are generally inequitable. On a per capita basis, allocations from the center exhibit high variance across provinces. And the individual components of the funding (excluding spending on salaries) tend to be positively correlated with provincial per capita incomes and the capacity to raise local revenue (World Bank 1991a). Griffin (1992) finds the distribution of central government health resources to the provinces to be inversely correlated with need as revealed by infant mortality rates.

Public treatment options include hospitals and the various facilities that make up the primary health care system: health centers (*puskesmas*), subcenters

(*puskesmas pembantu*), and integrated health and family planning posts (*posyandus*). The health centers are intended to provide preventive and curative care and to have a doctor in charge, with paramedics, nurses, and a midwife in attendance. Health centers in Java have an average of 15 staff members, but in the Outer Islands staffs are a fraction of that number (World Bank 1991a). A handful of health centers provide inpatient services. Subcenters are headed by a nurse or midwife, have small staffs, and offer curative and maternal and child care. The integrated health and family planning posts are makeshift clinics set up on a monthly basis in villages with support from the health centers and village health volunteers. The goal of the health and family planning clinics is to provide preventive services to children under five and to pregnant and nursing women. Although the health system is designed to function on the basis of referrals from the lowest service up, the sick tend to go directly to the highest accessible tier, where better quality of both staff and equipment can be expected.

Medical care is also available from the private sector, consisting of traditional medicine and modern health care services delivered by private doctors and paramedics (both are often public sector employees who set up private practice after hours), polyclinics (private outpatient clinics), and private hospitals.[2] In 1985, 255 private hospitals provided 34 percent of all inpatient days. Many private hospitals are maternity or industrial hospitals that cater only to company employees and their families.

Access to facilities is likely to be an important determinant of the cost of usage. In terms of health infrastructure, Indonesia had six hospital beds per 10,000 population in 1987. There were 5,639 public health centers, 17,382 subcenters, and 23,084 doctors for a population estimated at 172 million in 1987 (BPS 1990).

The availability of health facilities has improved considerably since the late 1970s. Although the number of hospitals (private and public) increased from 602 in 1978 to 703 in 1987, the number of beds just kept pace with population growth. A better measure of progress is at the community and primary health care level, where development began in earnest in 1978. By 1987 the number of health centers had increased by 30 percent (from 4,353 in 1978/79), of subcenters by 162 percent, and of physicians by 121 percent. The implementation of the health and family planning post system was just getting under way in 1978. In 1987/88 there were 13,754 government-trained paramedical graduates; in 1979/80 there were 2,789. During the same period the number of new government health posts increased from 5,651 to 11,907, while Indonesia's population grew by 17 percent (BPS 1983).

Despite these improvements, the availability of health services continues to vary significantly across geographic regions and to be inequitable. In general, all medical facilities are more readily accessible in urban areas. Although a majority of villages now have access to integrated health and family planning posts, other

2. In some cases the outpatient units of public hospitals are also termed polyclinics.

facilities remain sparse in rural Indonesia. The national hospital bed figure quoted above masks large regional disparities. To illustrate, Jakarta, the wealthiest region, has 1.24 beds per 1,000 population, whereas Lampung, one of the poorest provinces, has only 0.18. The average distance to a health center varies widely across provinces, from 0.8 kilometers in Jakarta to 33 kilometers in Irian Jaya. Health centers in those regions averaged 1.8 and 0.35 doctors per center, respectively. Subcenters and health and family planning posts tend to be more accessible, but average distances depend on location: in Jakarta the average patient must travel 0.2 kilometers to reach a health and family planning post; in Irian Jaya, 14 kilometers (World Bank 1991a).

The evidence suggests that travel time and costs in seeking medical care may still be prohibitive for households in many of Indonesia's rural areas. One reason is that population densities are uneven and extremely low in some parts of the country. Such areas also tend to have low per capita incomes. Of course, these circumstances are unlikely to attract private providers of health care.

The quality of medical facilities also varies markedly. Quality is often lower in poorer, more isolated regions, taking the form of inadequate and unreliable drug supplies, a narrower range of services, and fewer and less-qualified staff. For example, public health centers in poor, remote areas often find it difficult to attract and keep a trained doctor and so function without one.

Many observers have remarked on the low utilization rates of hospitals and health centers (see World Bank 1991a). Distances appear to be an important part of the explanation. Low utilization rates are also found to be correlated with low regional per capita expenditure. And there is evidence linking low quality with low utilization, as well. For example, the presence of surgical services and specialist doctors in a facility has been shown to have a high positive effect on utilization rates (World Bank 1991a).

In addition to the private time and transport costs entailed in traveling to public facilities, Indonesians must pay for health care. A set fee of 300 rupiah per health center visit covers consultation and three days' worth of drugs. Visits to health and family planning posts are free. Hospitals charge a series of fees differentiated according to whether inpatient or outpatient care is provided and, for inpatient care, by the class of room. Additional charges are imposed for a range of special services. In general, private facilities charge much higher fees. Many public sector doctors and paramedics successfully charge significantly higher rates in after-hours private practices. Their availability at more convenient times appears to partly explain the demand for their services.

Quality and travel times vary across expenditure groups. To what extent do prices vary across expenditure groups? The government has made some attempts to protect the poor from unaffordable health care fees. The fixed fee for health centers is considered within the reach of most households. Provision of low-cost beds is required of all public and private hospitals. The government has also promoted the *surat kataranaan lurah,* an affidavit of indigence that poor individuals who are sick can obtain from the village head. The *surat* exempts the

Table 1. *Monthly Per Capita Expenditure on Health Care by Component, Area, and Household Consumption Expenditure Decile, Indonesia, 1987*
(rupiahs per month)

| *Expenditure category and geographic area* | *Decile*[a] *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* | *Expenditure elasticity*[b] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Doctors* | | | | | | | | | | | |
| All-Indonesia | 1.90 | 5.06 | 6.13 | 8.03 | 8.23 | 10.73 | 19.44 | 25.49 | 39.64 | 86.04 | 1.85 (17.4) |
| Urban | 7.27 | 10.13 | 15.47 | 26.72 | 33.65 | 25.95 | 37.34 | 47.11 | 73.27 | 116.62 | 1.40 (12.6) |
| Rural | 1.53 | 2.94 | 5.29 | 7.10 | 8.22 | 8.06 | 8.80 | 16.08 | 20.81 | 63.29 | 2.05 (17.1) |
| *Inpatient care* | | | | | | | | | | | |
| All-Indonesia | 0.75 | 1.53 | 7.39 | 1.79 | 2.93 | 3.76 | 4.23 | 11.03 | 14.35 | 63.41 | 2.01 (6.9) |
| Urban | 4.31 | 1.27 | 5.74 | 4.21 | 7.12 | 13.33 | 7.85 | 21.60 | 89.29 | 97.48 | 2.05 (5.6) |
| Rural | 0.95 | 0.34 | 2.58 | 8.23 | 1.88 | 3.67 | 3.26 | 3.10 | 10.72 | 23.54 | 1.99 (4.0) |
| *Nurses and midwives* | | | | | | | | | | | |
| All-Indonesia | 0.25 | 0.77 | 0.75 | 1.53 | 2.21 | 1.90 | 3.35 | 7.47 | 9.45 | 10.03 | 1.93 (7.7) |
| Urban | 1.95 | 2.02 | 1.96 | 6.72 | 10.47 | 7.29 | 11.04 | 7.92 | 8.84 | 6.81 | 0.82 (2.6) |
| Rural | 0.15 | 0.97 | 0.38 | 0.29 | 1.62 | 2.60 | 2.14 | 1.99 | 5.96 | 11.38 | 2.47 (6.1) |
| *Paramedics* | | | | | | | | | | | |
| All-Indonesia | 4.68 | 7.29 | 7.07 | 9.21 | 8.65 | 9.57 | 8.52 | 9.27 | 8.74 | 6.27 | 0.10 (0.8) |
| Urban | 5.55 | 4.93 | 3.77 | 4.03 | 3.56 | 2.93 | 5.08 | 2.47 | 1.58 | 2.12 | −0.56 (3.7) |
| Rural | 4.16 | 7.51 | 6.79 | 8.13 | 9.38 | 10.11 | 10.72 | 10.45 | 13.05 | 15.18 | 0.69 (7.4) |
| *Birth control* | | | | | | | | | | | |
| All-Indonesia | 0.06 | 0.15 | 0.29 | 0.44 | 0.46 | 0.25 | 1.17 | 1.20 | 1.02 | 2.09 | 1.67 (6.0) |
| Urban | 0.02 | 0.36 | 0.43 | 1.60 | 2.07 | 1.81 | 1.29 | 0.90 | 2.37 | 3.70 | 1.98 (3.5) |
| Rural | 0.04 | 0.21 | 0.06 | 0.43 | 0.54 | 0.47 | 0.28 | 0.52 | 1.08 | 0.71 | 1.65 (3.4) |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Traditional healers* | | | | | | | | | | | |
| All-Indonesia | 2.79 | 2.58 | 3.05 | 4.89 | 4.68 | 3.99 | 4.58 | 3.88 | 3.38 | 5.50 | 0.28 (2.4) |
| Urban | 2.25 | 7.98 | 3.68 | 4.36 | 1.96 | 1.51 | 3.32 | 3.06 | 5.84 | 2.16 | −0.14 (0.5) |
| Rural | 3.08 | 2.75 | 2.79 | 2.88 | 5.11 | 3.66 | 4.35 | 4.80 | 4.75 | 6.36 | 0.51 (4.6) |
| *Drugs prescribed by a doctor* | | | | | | | | | | | |
| All-Indonesia | 1.09 | 1.76 | 3.35 | 3.23 | 5.74 | 7.41 | 16.91 | 24.34 | 40.27 | 98.42 | 2.40 (16.3) |
| Urban | 5.95 | 8.17 | 14.76 | 38.80 | 36.43 | 35.19 | 42.59 | 68.48 | 82.50 | 161.97 | 1.66 (9.6) |
| Rural | 1.13 | 0.97 | 2.66 | 2.50 | 2.08 | 5.34 | 4.72 | 9.40 | 14.26 | 49.85 | 2.37 (11.4) |
| *Drugs not prescribed by a doctor* | | | | | | | | | | | |
| All-Indonesia | 3.98 | 5.53 | 6.98 | 8.50 | 8.52 | 12.21 | 12.96 | 13.77 | 16.56 | 24.95 | 0.90 (12.8) |
| Urban | 9.66 | 8.94 | 15.73 | 12.48 | 14.46 | 15.32 | 16.96 | 18.11 | 26.02 | 30.68 | 0.61 (8.1) |
| Rural | 3.38 | 5.10 | 5.95 | 7.41 | 7.72 | 8.59 | 10.87 | 12.68 | 12.95 | 18.90 | 1.00 (12.9) |
| *Other health goods and services* | | | | | | | | | | | |
| All-Indonesia | 1.47 | 2.04 | 2.31 | 3.22 | 3.55 | 3.79 | 5.16 | 6.66 | 5.90 | 13.02 | 1.07 (15.2) |
| Urban | 3.61 | 4.16 | 4.97 | 8.11 | 8.41 | 5.63 | 7.09 | 5.00 | 9.77 | 26.36 | 0.81 (4.5) |
| Rural | 1.61 | 1.38 | 2.14 | 2.74 | 2.95 | 2.69 | 4.27 | 3.70 | 5.45 | 6.70 | 0.96 (8.3) |
| *Total health care expenditures* | | | | | | | | | | | |
| All-Indonesia | 16.95 | 26.71 | 37.32 | 40.85 | 44.96 | 53.61 | 76.33 | 103.11 | 139.31 | 309.74 | 1.43 (34.9) |
| Urban | 40.57 | 47.95 | 66.50 | 107.03 | 118.14 | 108.95 | 132.57 | 174.66 | 299.48 | 447.90 | 1.25 (14.7) |
| Rural | 16.03 | 22.17 | 28.64 | 39.72 | 39.50 | 45.18 | 49.41 | 62.71 | 89.03 | 195.90 | 1.43 (27.3) |

a. Deciles group persons by total household consumption expenditure per capita.

b. Expenditure elasticities were derived by an ordinary-least-squares regression of each health expenditure component on the decile mean per capita expenditure level for each region, both variables in logs. Numbers in parentheses are *t*-ratios.

*Source:* 1987 SUSENAS data tapes.

recipient from paying the fees associated with one medical treatment. However, anecdotal evidence suggests that the *surat* is very little used.

There is also limited health insurance coverage under PHB (Perum Umum Husada Bhakti; formerly, ASKES). This government-run insurance scheme covers active and retired public servants and their dependents for treatment in public facilities. Estimates of the number of people covered vary from around 10.5 million to 14 million for 1986 (World Bank 1991a). The scheme is financed through a 2 percent levy on the base salaries of all government workers and the pension payments of retired government workers. Those covered are rarely poor (World Bank 1991a).

How important are private expenditures on health care? Table 1 presents monthly per capita expenditures on health according to deciles of total household consumption per capita as recorded in the 1987 National Socio-Economic Survey (SUSENAS). Absolute magnitudes are low. The bottom decile spent most of its health care expenditure on non-doctor-prescribed drugs, paramedics, and, in urban areas only, private doctors. In contrast, the highest expenditure components for the top decile were doctor-prescribed medicines and private doctors. Generally, the amounts increase with overall living standards, and they also increase more than proportionately with consumption. (The elasticity of expenditures in each category with respect to total consumption expenditure tends to exceed one.) As a simple summary measure, the least-squares elasticities are recorded in the last column.

The above observations suggest that in Indonesia potential patients choose between subsidized public health care, where choice of quantity and quality is limited, and unsubsidized private care, where choice is greater but at a higher price (assumed here to include the full cost). Under certain conditions there will then be an equilibrium in which the rich (whose perceived net benefits from attending public facilities are assumed to be lower than for others) opt out of the public health sector.[3] However, access to private care may cost so much that the demand for private health care becomes zero. For example, in many rural areas, individuals, whether rich or poor, would have to travel very long distances—with a high opportunity cost of time—to reach private facilities. In these circumstances, everyone opts out of the private sector.

A similar situation arises in choosing between different public services. The highest levels—hospitals—dispense higher-quality care but are also more expensive, both because of higher user fees and because they are scarcer and hence entail greater traveling costs for most patients. Some self-selection across expenditure groups can be expected. Finally, the low quality of services and the significant distances in some regions are likely to deter even the poor and result in continued reliance on traditional medicine or self-treatment, neglect of pre-

3. For example, with additively separable utility between health care and other goods, declining marginal utility of consumption, common preferences, and a competitive private sector, there will exist a unique expenditure switch point below which people use only public facilities and above which they opt for private care.

ventive care, and delays in seeking treatment except in emergencies. Increased availability of facilities, as well as higher living standards, can be expected to relax some of these constraints and improve access to health care.

## II. The Methodology

A household's standard of living depends on its command over both private goods and the benefits derived from publicly provided goods such as education and health care. Thus an assessment of the interhousehold distribution of the benefits of public expenditures should compare the distribution of living standards without government spending with the distribution attained with publicly provided services.

Commonly used indicators of living standards, such as household per capita expenditures, exclude the monetary value derived from publicly provided goods. However, for several reasons, these indicators provide only a rough approximation of the distribution that would be obtained prior to government intervention. Household expenditures on private goods are influenced by what governments spend on public services. Public services may displace private spending: for example, when outpatient care in a public hospital is provided at a subsidized rate, people will spend less on private doctors. Public services may also augment private spending: for example, subsidized schooling may encourage households that might not otherwise send their children to a private school to spend income on children's clothing. Furthermore, the distribution of living standards is influenced by the outcomes (such as good health) of past public spending. These effects are difficult to quantify. Here I follow common practice in assuming that total household expenditure on privately supplied goods and services ("consumption" for short) is an adequate proxy for living standards in the absence of publicly provided goods. Thus the distribution of the benefits of publicly provided goods and services across households ranked by consumption is the basis for assessing the impact of public provisioning on living standards.

It should be pointed out that the study makes a risk-neutral valuation of the subsidies. That is, it examines the transfer benefits, not the risk benefits, of public provision of health care. Risk benefits depend on the availability of risk markets, including insurance, and on how well these markets perform. It could be argued that risk markets work less well for the poor, so that there are potentially large risk benefits from public health care subsidies for the poor, and for many of the nonpoor in rural areas. But there is no way of measuring this.

### *Benefit Incidence Analysis*

Utilization is measured as the proportion of an eligible subgroup that makes use of a health facility. The estimated unit costs to the government are then attributed across households according to utilization patterns. This approach, usually referred to as "benefit incidence" analysis, became popular in the late 1970s, spurred in part by the increased availability of and improvements in

household-level surveys.[4] The best-known applications for developing countries are the studies of Malaysia by Meerman (1979) and of Colombia by Selowsky (1979).

It is notoriously difficult to measure the benefits from publicly provided goods and services (see Cornes 1992). The problems associated with the benefit incidence approach are well documented (see, for example, Selden and Wasylenko 1992). Here, it may be useful to point out what are likely to be some of the more important concerns in the present context. A key question has always been how well the standard methodology approximates the distribution of the value of the benefits. Utilization need not fully reflect the actual benefits derived from a health facility. "Need," as measured by reported illness, is often juxtaposed with treatment received to judge equity of access and value of benefits. Yet when medical need is based on whether the household reports a member as having been sick in the prior week, there is no information about the severity of the illness. It may be reasonable to assume that poor households tend to ignore illnesses (out of necessity) more than rich households do. Chernikovsky and Meesook (1986) also speculate that access to health services influences the reporting of illness because the likelihood of being treated encourages recognition of poor health. In either case the assessment of the degree of need is biased, and this impairs the ability to assess the distribution of health spending. The probable direction of the bias in recall will be to underestimate the needs of the poor.

Another weakness of the methodology is that all facilities dispensing a certain type of service are treated identically. Yet differential quality of service is an important characteristic of health care in Indonesia. This fact is relevant in allocating subsidies in that the per-unit cost of a low-quality service will generally not equal that of a high-quality one, and the methodology will tend to underestimate the disparities in how benefits are distributed.[5] Policy implications may also be affected. Finally, the approach does not allow for the private costs of participation. These are likely to be correlated with living standards and so could be important in assessing results and the implications of incidence estimates.

In estimating unit costs, this study (like most such studies) concentrates on variable and semivariable or "recurrent" costs. It does not account for capital costs. This omission may lead to biases in the qualitative results. Meerman (1979) found that failure to account for public capital in Malaysia leads to serious underestimation of the total community resources used to provide medi-

4. In contrast, "expenditure incidence" studies examine the question of who receives government expenditures through, for example, being employed by the public sector (doctors, nurses, teachers, and so on). For a detailed review of the past and present state of benefit incidence analysis, see Selden and Wasylenko (1992).

5. The per-unit cost of a low-quality service may in fact be higher than that of a high-quality service if, for example, low quality is the result of the costs of reaching the area in which the service is located. But the costs of reaching the area may be the result of low public spending in the past. What time horizon to use and how to treat the incidence of capital expenditures are problematic. In any case, the benefits of health care subsidies will not be equal across regions.

cal care. For example, allowing for imputed capital service costs in the Malaysia data increased total costs per inpatient-day by 78 percent. Capital costs may also matter to policy decisions. Higher-level services necessitate more costly capital inputs and are likely to be used more by wealthier groups. Ignoring capital will then result in an underestimation of the inequality in the distribution of public expenditures. In addition, in allocating spending between sectors in the most cost-effective way, total public costs will often be more relevant than recurrent expenditures on their own. It is important to keep these points in mind when drawing conclusions from the incidence estimates.

### *Implementation for Indonesia*

The data used here are from Indonesia's 1978 and 1987 National Socio-Economic Surveys (SUSENAS), which are detailed consumption surveys based on large, nationally representative samples. These surveys provide the best source of household-level data for Indonesia.[6] It is therefore natural to use consumption expenditures as the welfare indicator.

The analysis is carried out along two separate dimensions. At one level, an attempt is made to provide a broad profile of utilization and of the incidence of subsidies for 1987. Incidence analysis at one point in time attempts to estimate how average benefits are distributed. It says nothing about whether increments to public expenditures are well distributed or pro-poor. For example, the rich may receive a large share of the inframarginal subsidies and the poor benefit most from the marginal spending. One way to get at this is by comparing incidence at two points in time. At a second level, the article attempts to characterize the changes in incidence since the late 1970s. A study by Meesook and Chernikovsky used the 1978 SUSENAS and the benefit incidence methodology to examine the incidence of public health expenditures (Meesook 1984; Chernikovsky and Meesook 1986). That study provides a benchmark for comparing utilization incidence across income groups between the two dates.

The work by Meesook and Chernikovsky is based on the May subround of the 1978 SUSENAS, covering 6,000 households. In 1987 the entire survey, comprising 55,000 households, was held during January. Overall, survey methodologies and questionnaires are generally comparable across the two surveys. Any dissimilarities are discussed in the text when they arise.

There is a worry that the timing of the surveys—May versus January—might affect the results of the comparison over time. For instance, the incidence of illnesses may vary from season to season, as may the opportunity cost of seeking treatment, and this variability may differ across expenditure groups. Little is known about seasonality in sickness, in health facility use, or in employment in Indonesia. Nor is much known about the link between seasonality in employment and in agriculture. Indonesia is in an equatorial region with no real dry

6. For a description of total household expenditures in the SUSENAS surveys and of the data generally, see van de Walle (1988).

season (Walsh 1981). There appears to be some seasonality in agriculture, but it varies significantly between and within islands. In addition, because Indonesia has a well-diversified rural economy, seasonality in work is less likely to be closely linked with crop seasonality than in many rural settings. Statistics on average maximum and minimum temperatures, wind velocity, and relative humidity across provinces and months indicate negligible differences between January and May (BPS 1990). Average rainfall tends to be lower in May, but this is not systematic across regions. It seems unlikely that there is any significant seasonal difference between May 1978 and January 1987, at least in the aggregate.

The analysis of health costs and budgets in Indonesia is not straightforward, for several reasons. There are numerous budgetary sources for the sector, and numerous ministries, as well as foreign funds, contribute to overall expenditure. But there is no central accounting system that keeps track of total spending, and the composition of expenditures is not clear from accounting classifications of outlays. Calculating total recurrent spending, let alone per-visit subsidies, is therefore a complex task (see World Bank 1991a). With these difficulties in mind, the article tries to follow the methodology detailed in Meesook (1984) as closely as the available data permit.

Implementation of the approach requires that costs per visit be calculated for the various public health facilities. The government spends on health care through hospitals and the primary health care system, as well as on training and communicable disease control. The study focuses on the apportionment of the benefits of expenditures on hospitals and public health centers (including subcenters) for which utilization is identifiable from the household-level data.

The requisite data on health financing are taken from a careful compilation of recurrent expenditures on public hospitals and health centers (World Bank 1991a). The same source estimated cost recovery to have been 3 percent of total recurrent expenditures on the health center system and 19.9 percent of recurrent expenditures on hospitals. These amounts are subtracted from recurrent expenditures to get the net government subsidy. Finally, the numbers of yearly visits to hospitals and to primary health centers are derived directly from the SUSENAS and, together with the recurrent expenditure levels, are used to calculate per-visit subsidies.

The earlier study (Meesook 1984) added up the 1980/81 routine budgets from all government levels, assumed that about two-thirds went to health care as an estimate of recurrent expenditures, and apportioned that amount between hospitals and public health centers. Fees collected from users and ASKES insurance were then subtracted to get the total yearly subsidy. Total annual visits to different health facilities were assessed from the 1980 census.

One difficulty arises from the fact that hospital care is also provided by private facilities in Indonesia. Public hospitals accounted for 66 percent of total inpatient days and 72 percent of all outpatient visits in 1985 (72 percent of total

hospital visits).[7] The subsidy for each hospital visit is calculated to be Rp5,200 when no distinction is made between public and private hospitals. This appears to be what was done by Meesook (1984) and thus provides the only basis for comparison with her results.

A different approach is adopted in presenting subsidy incidence estimates for 1987. Although it is not possible to identify visits to public, as opposed to private, facilities from the 1987 SUSENAS, this information is available from the special health module included in the 1990 SUSENAS. Visits to public hospitals as a proportion of total hospital visits derived from the 1990 data are given in table 2 and underlie the 1987 distribution of subsidies across consumption expenditure deciles presented later. This is the first time that such information has been available at the household level for Indonesia. Although the rich are widely believed to self-select away from public facilities, table 2 provides little evidence for this in rural areas. Indeed, the absence of any pattern across consumption deciles in the rural distribution of total outpatient or inpatient visits to public hospitals is striking. In urban Indonesia there is some evidence of a negative correlation between the share of visits to public hospitals and consumption levels. The rural numbers no doubt reflect lower rural densities and the consequent lack of a feasible choice between public and private hospitals.

Hospital visits include both inpatient and outpatient care, and different subsidy magnitudes are associated with each. In addition, as a proportion of total hospital visits, inpatient visits tend to increase with consumption. The level of each subsidy can be obtained by solving for $x_I$ in the identity $H = x_O N_O + x_I N_I$, where $H$ denotes the hospital budget net of user fees, $x_O$ is the average subsidy to a hospital outpatient visit, $x_I$ is the average subsidy to one inpatient day, $N_O$ stands for the number of public outpatient visits, and $N_I$ for the number of public inpatient visits. $N_I$ and $N_O$ are known from the SUSENAS, and $H$ is also known, as discussed earlier. An estimate of the outpatient-to-inpatient ratio of the rate of subsidy $(x_O/x_I)$ is needed.

Several studies have evaluated unit costs for individual health facilities in Indonesia.[8] Unit costs are consistently found to vary widely across facilities. This study is unable to take this variation into consideration and must therefore average over various estimates. One study of a sample of 40 hospitals calculated average unit costs in 1986/87 to be Rp3,593 for an outpatient visit and Rp12,803 for one inpatient day (Wirakartakusumah and others 1988).[9] Data

7. Visits to public hospitals as a percentage of total hospital visits are calculated from Ministry of Health data reported in table 2.4 in World Bank (1991a). The calculation is based on the total number of discharges from hospitals and the total number of outpatient visits, and the proportion of those that were private.

8. In these studies, unit costs are derived by adding up the individual cost components for a specific service output. This method represents a very different approach from the one pursued here, but it shares some of the same difficulties, including those encountered in collecting the data.

9. These are averages for class D and class C hospitals. Note also that the SUSENAS do not contain details on the length of hospitalization episodes. Here I assume that each inpatient visit is worth one subsidy amount.

Table 2. *Public Share of Hospital Visits by Region, Type of Visit, and Household Consumption Expenditure Decile, Indonesia, 1990*
(percentage of total hospital visits)

| *Visits to public hospitals by geographic area and type of visit* | *Decile*[a] | | | | | | | | | | *All visits to public hospitals* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* | |
| *All-Indonesia* | | | | | | | | | | | |
| All visits | 68.72 | 76.60 | 61.76 | 69.70 | 64.41 | 55.17 | 62.20 | 64.57 | 63.99 | 53.65 | 61.16 |
| Inpatient visits | 57.87 | 78.80 | 67.22 | 82.92 | 77.74 | 73.94 | 66.57 | 71.91 | 67.00 | 52.82 | 65.83 |
| *Urban areas* | | | | | | | | | | | |
| All visits | 62.45 | 49.20 | 56.10 | 50.54 | 59.98 | 62.58 | 49.79 | 64.16 | 56.74 | 39.47 | 54.38 |
| Inpatient visits | 75.48 | 83.40 | 79.93 | 63.74 | 64.78 | 69.25 | 55.59 | 60.80 | 52.31 | 45.61 | 59.36 |
| *Rural areas* | | | | | | | | | | | |
| All visits | 56.99 | 80.69 | 68.94 | 72.88 | 69.35 | 75.59 | 67.39 | 57.16 | 77.22 | 68.06 | 69.58 |
| Inpatient visits | 68.47 | 53.48 | 75.04 | 65.32 | 82.51 | 82.37 | 71.90 | 67.38 | 74.80 | 68.88 | 71.67 |

*Note:* The numbers give the percentages of total hospital and inpatient visits to public hospitals.
a. Deciles group persons by total household consumption expenditure per capita.
*Source:* 1990 SUSENAS data tapes.

on tariffs charged by a number of facilities for specific hospital treatments indicate an average fee of Rp300 for outpatient and Rp2,089 for inpatient care (Ministry of Health 1991a). This establishes a subsidy ratio of 0.307. On the basis of this information, the hospital outpatient subsidy is estimated to be Rp4,500 and the inpatient subsidy Rp14,600.

Finally, another factor tending to influence the distribution of health subsidies is government health insurance coverage under PHB. Coverage is thought to boost utilization of both primary health care centers and government hospitals, where free care is accorded to PHB cardholders. It can be presumed that PHB subscribers use these facilities relatively more than other groups, other things being equal. However, there is some controversy about what this implies for the incidence of subsidies. From existing evidence, it is probable that PHB contributions do not cover costs; what is less clear is whether those who are covered are more or less subsidized than others. It has been claimed that PHB reimburses health facilities at the official tariff rates (and perhaps at even lower rates; World Bank 1991a). This would imply higher subsidy rates to PHB patients (because tariffs are lower than average prices) and an underestimation of the regressivity of the distribution of health care subsidies. However, others claim that PHB reimbursements are actually higher than what other non-PHB patients pay in user fees, making the subsidies to civil servants lower than to others. This would in turn tend to imply a more progressive distribution of overall subsidies. It is unfortunately not clear how to take account of this without data identifying PHB recipients. For lack of any better evidence, I shall assume that the rate of subsidy for those covered by PHB is the same as for those not covered.

The average subsidy from recurrent expenditures for a visit to a health center is calculated to be Rp500. A study based on a survey of 42 rural health centers in five provinces in 1986/87 found average unit costs for curative care to average around Rp900 a visit, varying from a low of Rp526 for mother and child health care to a high of Rp1,337 for family planning consultations (Gani, Najib, and Wangsaraharja 1988). Although the fee recommended by the Ministry of Health in 1986/87 for a visit to a health center was Rp150, many local governments charged fees between Rp300 and Rp1,000 (World Bank 1991a). On the basis of a fee of Rp300, the estimate of the average unit cost points to a subsidy of Rp600 per visit, not too far off from the estimate here for primary health centers. Some patients pay less, and some are treated gratis if they have a letter of indigence from the village headman. Total visits from the SUSENAS data set include consultations at health and family planning posts, which have low unit costs and which were not considered in the Gani, Najib, and Wangsaraharja (1988) study. Although estimates of the per-unit subsidies must be viewed as rough, they do permit an idea of the relative orders of magnitude.

## III. The Utilization of Health Services

This section examines the utilization of health care services across consumption expenditure groups in urban and rural areas in 1987. This utilization profile is then compared with one drawn up for 1978.

### *The Picture in 1987*

According to the 1987 SUSENAS, 65 percent of all those reporting ill during the preceding week also reported seeking treatment outside the family. Of all outside treatments in 1987, the greatest numbers consulted a primary health center (43 percent), followed by paramedics (22 percent), private doctors (17 percent), hospitals (8 percent), traditional healers (6 percent), and polyclinics (4.5 percent). (Note that the surveys only record one treatment option per reported illness.) The 1978 SUSENAS implied that 23 percent of visits were to public health centers, 19 percent to hospitals, and 14 percent to private doctors (Meesook 1984). Thus there has been a sizable increase in the relative importance of public health centers.

For a variety of reasons, including both the availability of and the demand for services, the use of health facilities often differs between expenditure groups. Table 3 provides evidence from the 1987 SUSENAS of how individuals, ranked into deciles of per capita household expenditures (with decile 1 being the poorest and 10 the wealthiest), responded to illness. This information is presented for the all-Indonesia, urban, and rural distributions. It is clear that area of residence also bears on the use of facilities.

The percentage of reported illnesses treated by private doctors and hospitals is an increasing function of per capita expenditures, ranging from just under 2 percent for the poorest 20 percent in rural areas to 47 percent for the richest 10 percent in urban areas. Visits to private doctors exceeded those to hospitals for all groups and also increased much more steeply across household expenditure deciles. Both options were more common in urban than in rural areas. Conversely, the share of individuals who were not treated or were treated exclusively by themselves or their families falls across deciles, from 46 percent of total reported illnesses for the poorest decile to 23 percent for the wealthiest. The disparity is even more pronounced for those who did not seek any treatment: 12 percent did not do so among the poorest decile in rural areas, but the share was 0.3 percent for the richest urban decile.

Recourse to primary health centers drops systematically for the 6th through 10th deciles in urban areas. Use does not appear to have been linked to consumption levels in rural areas, where the proportion of total illnesses treated in primary health centers ranged from around 27 to 32 percent. The use of polyclinics was consistently low and appears to have been unrelated to household living standards. Use of paramedics declined with expenditures in urban areas, as well as at the all-Indonesia level. In rural areas, in contrast, the use of paramedics was maintained at around 15 to 17 percent across the deciles. The percentage of illnesses attended to by traditional healers was generally low,

being lowest for urban individuals. Use of traditional healers does not seem to have been significantly influenced by household expenditure levels, although this is less true for urban areas.

Table 3 also shows what proportion of those reporting an illness received inpatient care and where. Again, the evidence suggests that the incidence of inpatient care is correlated with living standards. The proportion of the sick who went on to be treated as inpatients was larger in urban than in rural areas. Across deciles, a majority of these inpatients were admitted to hospitals. In rural areas, primary health centers and hospitals shared the burden of inpatient care. The homes of paramedics were a popular option for the bottom deciles but less so for the middle ones. Finally, traditional healers also played a role in rural areas.

Table 4 presents additional detail on annual absolute utilization rates per person for modern health providers. The table reveals how the number of total and provider-specific visits differs across consumption deciles and sectors. It is clear from table 4 that the rate at which morbidity is treated varies across deciles, rising with consumption. The latter effect is more pronounced in rural areas, where individuals in the poorest decile visited one of the modern facilities an average of 1.4 times a year, whereas individuals in the wealthiest decile did so 3 times a year. In urban Indonesia the variability was lower, ranging from 1.8 to 2.2 visits per person per year.

### *Changes in the Incidence of the Utilization of Health Services between 1978 and 1987*

Table 5 presents statistics on individuals reporting illnesses and where they were treated, as recorded in both the 1978 and 1987 surveys. The table provides insights into how the type of treatment sought by different subgroups altered between 1978 and 1987. For example, of all those reporting ill in urban Java in the lower 40 percent of the per capita expenditure distribution, 31.8 percent did not seek treatment outside the home in 1987. The 1978 results, taken from Chernikovsky and Meesook (1986), are also given in the table. In 1978, 58 percent of urban Javanese in the poorest expenditure group who claimed to have been ill sought treatment.

The following observations can be made about each treatment option.

*Self, family, or no treatment.* At both dates, the lowest income groups were least likely to seek treatment outside the home. Indeed, self, family, or no treatment was consistently their most common course of action. But the use of facilities outside the home by the poorest 40 percent clearly increased after 1978. For example, in urban Java, treatment by self or family or no treatment declined from 58 to 32 percent. Urban residents were also more likely than rural residents to pursue outside treatment at any given consumption level.

*Primary health centers.* The 1978 results indicate that for rural areas, primary health centers were predominantly used by middle expenditure house-

Table 3. *Treatment of the Ill by Region, Type of Visit, and Household Consumption Expenditure Decile, Indonesia, 1987*
(percentage of people reporting an illness in the preceding week)

| | *Decile*[a] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Geographic area and type of treatment* | *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* |
| *All-Indonesia* | | | | | | | | | | |
| Percentage treated by | | | | | | | | | | |
| Private doctor | 2.15 | 2.54 | 3.43 | 5.28 | 6.82 | 8.62 | 12.21 | 14.18 | 20.43 | 31.65 |
| Hospital | 1.99 | 2.25 | 2.42 | 4.01 | 4.49 | 4.33 | 6.59 | 6.43 | 7.32 | 11.42 |
| Primary health center | 26.75 | 29.35 | 28.05 | 29.10 | 27.21 | 29.47 | 29.49 | 32.15 | 27.72 | 19.48 |
| Polyclinic | 3.44 | 2.73 | 2.02 | 3.82 | 3.56 | 2.16 | 2.26 | 3.14 | 3.42 | 2.85 |
| Paramedic | 14.93 | 16.20 | 16.92 | 14.64 | 15.01 | 16.51 | 15.57 | 12.32 | 11.60 | 8.73 |
| Traditional healer | 4.39 | 4.45 | 4.34 | 4.64 | 4.83 | 3.80 | 3.55 | 3.24 | 3.88 | 2.39 |
| Self or family | 35.72 | 34.60 | 33.26 | 31.12 | 32.67 | 30.33 | 25.79 | 23.97 | 21.60 | 20.74 |
| No medication (not treated) | 10.63 | 7.87 | 9.57 | 7.41 | 5.42 | 4.78 | 4.55 | 4.59 | 4.04 | 2.75 |
| Percentage receiving inpatient care | 1.92 | 1.95 | 1.99 | 1.77 | 2.12 | 2.20 | 2.75 | 2.97 | 3.03 | 6.41 |
| At a primary health center | 36.84 | 44.95 | 53.83 | 29.31 | 21.30 | 34.31 | 14.70 | 36.95 | 23.92 | 20.28 |
| At a hospital | 26.17 | 34.51 | 31.35 | 40.98 | 53.29 | 54.08 | 75.00 | 57.28 | 66.04 | 74.43 |
| By a paramedic | 32.59 | 7.80 | 6.09 | 14.87 | 9.82 | 6.82 | 6.98 | 2.18 | 3.66 | 4.67 |
| By a traditional healer | 4.39 | 12.74 | 8.72 | 14.83 | 15.59 | 4.79 | 3.31 | 3.60 | 6.39 | 0.63 |
| *Urban areas* | | | | | | | | | | |
| Percentage treated by | | | | | | | | | | |
| Private doctor | 7.59 | 10.61 | 19.03 | 21.53 | 18.76 | 22.15 | 32.69 | 31.54 | 36.60 | 46.77 |
| Hospital | 7.14 | 5.65 | 8.67 | 12.28 | 13.72 | 9.28 | 8.83 | 13.10 | 15.77 | 15.72 |
| Primary health center | 26.98 | 27.64 | 29.31 | 28.76 | 29.90 | 33.04 | 24.75 | 21.22 | 16.46 | 11.00 |
| Polyclinic | 1.06 | 5.41 | 1.56 | 0.51 | 3.03 | 3.66 | 5.50 | 0.65 | 1.09 | 3.48 |
| Paramedic | 13.38 | 14.11 | 10.28 | 11.02 | 4.86 | 3.17 | 3.84 | 6.91 | 3.19 | 3.32 |
| Traditional healer | 2.94 | 2.28 | 0.89 | 2.08 | 0.67 | 1.05 | 1.34 | 2.58 | 1.44 | 1.44 |
| Self or family | 33.40 | 30.08 | 25.91 | 20.73 | 26.40 | 22.55 | 19.45 | 20.42 | 23.17 | 18.01 |
| No medication (not treated) | 7.51 | 4.22 | 4.35 | 3.10 | 2.64 | 5.10 | 3.61 | 3.59 | 2.28 | 0.28 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Percentage receiving inpatient care | 2.36 | 0.44 | 2.96 | 3.25 | 2.54 | 2.81 | 3.07 | 4.47 | 5.04 | 9.44 |
| At primary health centers | 39.25 | 0.00 | 3.10 | 5.20 | 21.77 | 6.46 | 15.51 | 19.39 | 12.47 | 1.64 |
| At hospitals | 45.37 | 100.00 | 96.50 | 94.80 | 78.23 | 89.72 | 64.33 | 78.07 | 84.84 | 92.46 |
| By paramedics | 0.00 | 0.00 | 0.39 | 0.00 | 0.00 | 3.39 | 3.16 | 2.54 | 0.55 | 5.90 |
| By traditional healers | 15.38 | 0.00 | 0.00 | 0.00 | 0.00 | 0.43 | 17.00 | 0.00 | 2.15 | 0.00 |
| *Rural* | | | | | | | | | | |
| Percentage treated by | | | | | | | | | | |
| Private doctor | 1.81 | 1.73 | 2.28 | 3.42 | 4.51 | 6.09 | 5.84 | 8.01 | 11.52 | 15.94 |
| Hospital | 1.56 | 2.21 | 1.66 | 2.66 | 3.54 | 3.83 | 3.06 | 4.42 | 3.87 | 5.60 |
| Primary health center | 26.80 | 28.67 | 29.21 | 29.00 | 28.61 | 27.85 | 29.76 | 29.29 | 32.13 | 27.05 |
| Polyclinic | 3.64 | 2.65 | 3.14 | 2.13 | 4.08 | 2.94 | 1.82 | 2.81 | 3.43 | 3.57 |
| Paramedic | 15.49 | 15.23 | 17.29 | 17.17 | 14.59 | 15.69 | 16.73 | 17.91 | 15.93 | 17.13 |
| Traditional healer | 4.85 | 5.23 | 3.32 | 5.18 | 4.31 | 5.15 | 5.23 | 4.15 | 4.52 | 4.72 |
| Self or family | 34.21 | 36.35 | 33.04 | 33.22 | 31.36 | 32.66 | 32.89 | 28.21 | 23.81 | 21.64 |
| No medication (not treated) | 11.64 | 7.92 | 10.06 | 7.24 | 9.01 | 5.81 | 4.68 | 5.20 | 4.78 | 4.36 |
| Percentage receiving inpatient care | 1.81 | 1.50 | 2.08 | 2.18 | 2.11 | 2.78 | 1.85 | 2.46 | 2.86 | 4.25 |
| At primary health centers | 26.81 | 55.49 | 43.97 | 54.32 | 28.58 | 23.06 | 36.74 | 29.51 | 40.95 | 37.95 |
| At hospitals | 34.81 | 7.43 | 43.03 | 25.57 | 44.92 | 51.60 | 42.21 | 60.02 | 48.89 | 54.05 |
| By paramedics | 33.07 | 21.44 | 6.40 | 11.21 | 13.80 | 10.63 | 10.62 | 6.18 | 5.15 | 5.54 |
| By traditional healers | 5.32 | 15.64 | 6.60 | 8.90 | 12.70 | 14.72 | 10.43 | 4.29 | 5.01 | 2.46 |

*Note:* Treatment refers to treatment received in the preceding week.

a. Deciles group persons by total household consumption expenditure per capita.

*Source:* 1987 SUSENAS data tapes.

Table 4. *Utilization of Modern Health Providers by Area, Type of Provider, and Household Consumption Expenditure Decile, Indonesia, 1987*
(annual visits per capita)

| *Geographic area and health care provider* | *Decile*[a] *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* | *All expenditure groups* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *All-Indonesia* | | | | | | | | | | | |
| All visits | 1.44 | 1.71 | 1.66 | 1.88 | 1.91 | 2.19 | 2.45 | 2.51 | 2.61 | 2.30 | 2.07 |
| Private doctor | 0.06 | 0.08 | 0.11 | 0.17 | 0.23 | 0.31 | 0.45 | 0.52 | 0.76 | 0.98 | 0.37 |
| Hospital | 0.06 | 0.07 | 0.08 | 0.13 | 0.15 | 0.16 | 0.24 | 0.24 | 0.27 | 0.35 | 0.18 |
| Primary health center | 0.78 | 0.94 | 0.88 | 0.96 | 0.91 | 1.06 | 1.09 | 1.18 | 1.03 | 0.60 | 0.94 |
| Polyclinic | 0.10 | 0.09 | 0.06 | 0.13 | 0.12 | 0.08 | 0.08 | 0.12 | 0.13 | 0.09 | 0.10 |
| Paramedic | 0.44 | 0.52 | 0.53 | 0.48 | 0.50 | 0.59 | 0.58 | 0.45 | 0.43 | 0.27 | 0.48 |
| *Urban areas* | | | | | | | | | | | |
| All visits | 1.76 | 2.50 | 2.16 | 2.39 | 2.06 | 2.16 | 2.25 | 2.10 | 1.95 | 2.17 | 2.15 |
| Private doctor | 0.24 | 0.42 | 0.60 | 0.69 | 0.55 | 0.67 | 0.97 | 0.90 | 0.98 | 1.26 | 0.73 |
| Hospital | 0.22 | 0.22 | 0.27 | 0.40 | 0.40 | 0.28 | 0.26 | 0.37 | 0.42 | 0.42 | 0.33 |
| Primary health center | 0.85 | 1.09 | 0.92 | 0.93 | 0.88 | 1.00 | 0.74 | 0.61 | 0.44 | 0.30 | 0.77 |
| Polyclinic | 0.03 | 0.21 | 0.05 | 0.02 | 0.09 | 0.11 | 0.16 | 0.02 | 0.03 | 0.09 | 0.08 |
| Paramedic | 0.42 | 0.56 | 0.32 | 0.35 | 0.14 | 0.10 | 0.11 | 0.20 | 0.08 | 0.09 | 0.24 |
| *Rural areas* | | | | | | | | | | | |
| All visits | 1.41 | 1.70 | 1.56 | 1.75 | 1.78 | 1.83 | 2.12 | 2.38 | 2.82 | 3.00 | 2.03 |
| Private doctor | 0.05 | 0.06 | 0.07 | 0.11 | 0.14 | 0.20 | 0.22 | 0.30 | 0.49 | 0.69 | 0.23 |
| Hospital | 0.04 | 0.07 | 0.05 | 0.09 | 0.11 | 0.12 | 0.11 | 0.17 | 0.16 | 0.24 | 0.12 |
| Primary health center | 0.77 | 0.96 | 0.85 | 0.93 | 0.92 | 0.90 | 1.10 | 1.11 | 1.35 | 1.17 | 1.01 |
| Polyclinic | 0.10 | 0.09 | 0.09 | 0.07 | 0.13 | 0.10 | 0.07 | 0.11 | 0.14 | 0.15 | 0.11 |
| Paramedic | 0.44 | 0.51 | 0.50 | 0.55 | 0.47 | 0.51 | 0.62 | 0.68 | 0.67 | 0.74 | 0.57 |

a. Deciles group persons by total household consumption expenditure per capita.
*Source:* 1987 SUSENAS data tapes.

Table 5. *Treatment of the Ill by Provider, Area, and Household Consumption Expenditure Group, Indonesia, 1978 and 1987*
(percentage of people reporting an illness)

| | *Java* | | | | | | *Outer Islands* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Urban areas (household consumption expenditure group)*[a] | | | *Rural areas (household consumption expenditure group)*[a] | | | *Urban areas (household consumption expenditure group)*[a] | | | *Rural areas (household consumption expenditure group)*[a] | | |
| *Health care provider and year* | *Lower 40 percent* | *Middle 40 percent* | *Upper 20 percent* | *Lower 40 percent* | *Middle 40 percent* | *Upper 20 percent* | *Lower 40 percent* | *Middle 40 percent* | *Upper 20 percent* | *Lower 40 percent* | *Middle 40 percent* | *Upper 20 percent* |
| *Self, family, or no treatment* | | | | | | | | | | | | |
| 1987 | 31.8 | 26.4 | 19.6 | 45.7 | 37.6 | 27.6 | 34.7 | 23.7 | 27.9 | 41.2 | 35.6 | 28.0 |
| 1978 | 58 | 27 | 12 | 53 | 41 | 40 | 33 | 52 | 26 | 43 | 39 | 33 |
| *Primary health center* | | | | | | | | | | | | |
| 1987 | 26.9 | 26.3 | 14.4 | 30.5 | 31.5 | 31.0 | 31.2 | 30.1 | 14.5 | 25.5 | 25.8 | 28.0 |
| 1978 | 19 | 22 | 15 | 17 | 37 | 21 | 27 | 10 | 22 | 11 | 35 | 23 |
| *Private doctor* | | | | | | | | | | | | |
| 1987[b] | 29.3 | 33.0 | 46.2 | 19.0 | 23.0 | 32.7 | 19.1 | 26.5 | 39.4 | 17.9 | 22.3 | 26.8 |
| 1978[b] | 13 | 34 | 58 | 22 | 12 | 29 | 17 | 27 | 38 | 15 | 9 | 25 |
| *Hospital* | | | | | | | | | | | | |
| 1987 | 8.21 | 9.63 | 16.5 | 1.11 | 3.31 | 3.69 | 9.01 | 14.9 | 14.1 | 3.07 | 4.56 | 5.73 |
| 1978 | 0 | 14 | 5 | 1 | 1 | 7 | 7 | 5 | 11 | 1 | 2 | 6 |
| *Private clinic* | | | | | | | | | | | | |
| 1987[c] | 2.41 | 3.43 | 2.58 | 1.55 | 2.31 | 1.64 | 2.06 | 2.61 | 1.89 | 4.51 | 3.97 | 5.15 |
| 1978[d] | 0 | 0 | 9 | 1 | 2 | 0 | 0 | 3 | 2 | 8 | 5 | 3 |
| *Traditional healer* | | | | | | | | | | | | |
| 1987 | 1.46 | 1.19 | 0.88 | 2.12 | 2.31 | 3.35 | 3.92 | 2.19 | 2.23 | 7.78 | 7.68 | 6.37 |
| 1978 | 10 | 3 | 1 | 6 | 7 | 3 | 15 | 4 | 1 | 22 | 10 | 10 |

*Note:* Each number indicates the proportion of all those reporting ill in each expenditure group by geographic area and type of health care provider.
a. Grouped by household consumption expenditure per capita.
b. Includes paramedics.
c. Includes polyclinics.
d. Includes maternity hospitals and clinics.
*Source:* 1987 SUSENAS data tapes; Chernikovsky and Meesook (1986).

holds. The poor used these facilities relatively little (many going without treatment), and the rich tended to use other facilities, such as private doctors, more intensively. In urban areas the pattern differed between Java, where use was also highest for the middle expenditure group (22 percent), and the Outer Islands, where the poorest were the most common users (27 percent) and the middle the least common (10 percent). By 1987 the use of primary health centers had increased for most subgroups. There were two exceptions: use dropped for the wealthiest groups in the urban areas of both Java and the Outer Islands and for the middle expenditure group in rural areas.

In the urban areas of both Java and the Outer Islands the use of health centers declined with consumption expenditures, although only mildly between the bottom and middle groups. In marked contrast, use of rural primary health centers appears relatively equal across expenditure groups. The upper 20 percent were just as likely (if not slightly more likely) to use them as the lower 40 percent. One could not conclude from these data for 1987 and from tables 2 and 3 that subsidizing primary health care in rural areas is inherently pro-poor; the benefits are quite uniformly distributed. However, the benefits of subsidized primary health care will tend to be more pro-poor in urban areas.

*Private doctors.* Unlike visits to public health centers, visits to private doctors increased markedly with expenditure levels in both rural and urban areas. Use was also higher in urban areas at any consumption level. Between 1978 and 1987, use of private doctors increased for five out of six subgroups in the Outer Islands but for only half of all subgroups in Java. This difference may reflect the relatively lower availability of cheaper yet acceptable alternatives in the Outer Islands.

*Hospitals.* Hospital treatment also increased with urban residence and household per capita expenditure. The rich used hospitals less in rural areas and more in urban areas in 1987 than in 1978. Other groups mostly increased their use of hospitals.

*Private clinics.* The categories listed in the 1978 and 1987 surveys do not correspond exactly for private clinics. The 1978 survey asked about maternity hospitals and clinics, whereas the 1987 survey listed polyclinics. Both categories cover private facilities, which often offer better quality than the available public facilities and charge more for it. Therefore the two are compared here. The proportion of individuals using these facilities in 1987 tended to be highest for the middle 40 percent expenditure group. The exception was the rural Outer Islands, where both the lower and upper expenditure groups exhibited higher usage. By contrast, in 1978 usage of private clinics was very uneven, making a pattern difficult to discern. After 1978 usage generally decreased in the Outer Islands and increased in Java.

*Traditional healers.* The importance of traditional healers declined almost consistently over the decade, although this form of treatment retained many

followers in the rural areas of the Outer Islands. But there was on the whole much less differentiation in use across expenditure classes in 1987 than in 1978. It seems that where health centers exist, the local poor use them in preference to traditional medical practitioners.[10]

## IV. Public Expenditures in the Health Sector

This section estimates the distribution of public subsidies in the health sector in 1987. Changes in the distribution of subsidies since 1978 are also examined.

### *The Incidence of Government Health Subsidies in 1987*

Tables 6 and 7 present data on health subsidies in 1987 for the urban, rural, and total population distributions. Table 6 characterizes the decile-specific distribution of public subsidies to hospitals and primary health centers. Hospital subsidies are calculated by differentiating between inpatient and outpatient visits, as described in section II, and under the assumption that the decile distribution of public hospital visits in 1987 is as indicated by the 1990 SUSENAS health module (see table 2).

The overall subsidy is found to have been mildly progressive in that the subsidy as a percentage of household consumption tended to be higher for the poor. Absolute subsidy levels tended to increase with per capita expenditure levels but decline as a proportion of them. Hence, the subsidies were inequality reducing. However, the programs were not particularly well targeted; indeed, uniform (untargeted) provision of lump-sum transfers to the whole population would have been much more progressive. The magnitude of hospital subsidies tended to increase much more as per capita expenditures rose than was the case with the health center subsidies. The latter were generally flatter across deciles, although they tended to increase for the top three rural deciles and to decrease for the top three urban ones. This result is in line with the findings in section III that in rural areas all groups used public health centers but that in urban areas the poor used them relatively more than others did.

Table 7 summarizes Indonesia's "household health account" for 1987. Outlays on health care by both the government and the household (see table 1) are juxtaposed across consumption expenditure deciles. Total per capita health spending is found to have generally increased (with some ups and downs) the higher the decile. Both public and private expenditures followed a similar upward trend, although public outlays exceeded private outlays for most deciles. The exceptions occurred for the top decile in all groups and for the 7th through

10. The term traditional healers encompasses various types of practitioners of "traditional" as opposed to "modern" medicine. The SUSENAS include bonesetters but probably exclude traditional midwives. There appears to be declining demand for some types of traditional healers. For example, from casual observation, *dukuns* (a broad category of traditional healer), who are trained by the government, seem to be in steady decline, whereas bonesetters maintain high popularity in many areas of Indonesia.

Table 6. *Distribution of Public Subsidies to Hospitals and Primary Health Centers by Household Consumption Expenditure Decile, Indonesia, 1987*
(rupiahs per capita per month, except as specified)

| | Decile[a] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Geographic area and health care provider* | *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* | *Average* |
| *All-Indonesia* | | | | | | | | | | | |
| Hospital subsidy | 21.79 | 34.67 | 28.30 | 50.57 | 60.20 | 57.98 | 98.58 | 93.86 | 105.48 | 135.33 | 68.68 |
| Public health center subsidy | 32.06 | 38.77 | 36.20 | 39.46 | 37.38 | 43.48 | 44.96 | 48.61 | 42.26 | 24.86 | 38.80 |
| Total subsidy | 53.85 | 73.44 | 64.50 | 90.03 | 97.58 | 101.46 | 143.54 | 142.47 | 147.73 | 160.19 | 107.48 |
| As share of per capita expenditure (percent) | 0.67 | 0.69 | 0.52 | 0.63 | 0.60 | 0.55 | 0.67 | 0.55 | 0.45 | 0.26 | 0.49 |
| *Urban areas* | | | | | | | | | | | |
| Hospital subsidy | 72.89 | 52.31 | 116.02 | 126.49 | 120.61 | 109.02 | 75.58 | 139.36 | 137.86 | 151.27 | 110.14 |
| Public health center subsidy | 34.84 | 44.70 | 37.82 | 38.07 | 36.02 | 41.14 | 30.31 | 24.96 | 18.04 | 12.21 | 31.81 |
| Total subsidy | 107.73 | 97.02 | 153.84 | 164.56 | 156.63 | 150.16 | 105.89 | 164.33 | 155.90 | 163.48 | 141.95 |
| As share of per capita expenditure (percent) | 0.95 | 0.63 | 0.82 | 0.75 | 0.62 | 0.52 | 0.32 | 0.42 | 0.32 | 0.19 | 0.43 |
| *Rural areas* | | | | | | | | | | | |
| Hospital subsidy | 19.61 | 23.84 | 28.71 | 32.73 | 50.01 | 66.53 | 45.52 | 67.00 | 83.16 | 117.99 | 53.51 |
| Public health center subsidy | 31.46 | 39.61 | 35.05 | 38.29 | 37.77 | 37.11 | 45.36 | 45.82 | 55.64 | 48.16 | 41.43 |
| Total subsidy | 51.08 | 63.45 | 63.75 | 71.01 | 87.78 | 103.64 | 90.88 | 112.82 | 138.80 | 166.15 | 94.94 |
| As share of per capita expenditure (percent) | 0.67 | 0.64 | 0.56 | 0.55 | 0.61 | 0.65 | 0.50 | 0.54 | 0.55 | 0.39 | 0.53 |

a. Deciles group persons by total household consumption expenditure per capita.
*Source:* Author's calculations from 1987 and 1990 SUSENAS data tapes.

Table 7. *Household Health Account by Area and Household Consumption Expenditure Decile, Indonesia, 1987*
(rupiahs per capita per month)

| | Decile[a] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Geographic area and expenditure item* | *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* | *Average* |
| *All-Indonesia* | | | | | | | | | | | |
| Total expenditure on health care | 70.81 | 100.15 | 101.82 | 130.88 | 142.54 | 155.07 | 219.87 | 245.58 | 287.04 | 469.93 | 192.37 |
| Expenditure by household | 16.96 | 26.71 | 37.32 | 40.85 | 44.96 | 53.61 | 76.33 | 103.11 | 139.31 | 309.74 | 84.89 |
| Subsidy from government | 53.85 | 73.44 | 64.50 | 90.03 | 97.58 | 101.46 | 143.54 | 142.47 | 147.73 | 160.19 | 107.48 |
| Mean total consumption | 8,007 | 10,621 | 12,421 | 14,212 | 16,160 | 18,501 | 21,460 | 25,764 | 32,997 | 60,757 | 22,090 |
| *Urban areas* | | | | | | | | | | | |
| Total expenditure on health care | 148.30 | 144.97 | 220.34 | 271.59 | 274.77 | 259.11 | 238.46 | 338.99 | 455.38 | 611.39 | 296.33 |
| Expenditure by household | 40.57 | 47.95 | 66.50 | 107.03 | 118.14 | 108.95 | 132.57 | 174.66 | 299.48 | 447.90 | 154.38 |
| Subsidy from government | 107.73 | 97.02 | 153.84 | 164.56 | 156.63 | 150.16 | 105.89 | 164.33 | 155.90 | 163.48 | 141.95 |
| Mean total consumption | 11,372 | 15,503 | 18,785 | 21,903 | 25,194 | 28,803 | 33,383 | 39,522 | 49,378 | 88,144 | 33,199 |
| *Rural areas* | | | | | | | | | | | |
| Total expenditure on health care | 67.12 | 85.62 | 92.39 | 110.73 | 127.28 | 148.82 | 140.29 | 175.53 | 227.83 | 362.05 | 153.77 |
| Expenditure by household | 16.04 | 22.17 | 28.64 | 39.72 | 39.50 | 45.18 | 49.41 | 62.71 | 89.03 | 195.90 | 58.83 |
| Subsidy from government | 51.08 | 63.45 | 63.75 | 71.01 | 87.78 | 103.64 | 90.88 | 112.82 | 138.80 | 166.15 | 94.94 |
| Mean total consumption | 7,595 | 9,909 | 11,432 | 12,860 | 14,373 | 16,065 | 18,123 | 20,841 | 25,429 | 42,614 | 17,924 |

a. Deciles group persons by total household consumption expenditure per capita.
*Source:* Author's calculations from 1987 and 1990 SUSENAS data tapes.

10th deciles in urban Indonesia. Again, for the poor, public provisioning was relatively more important than private provisioning.

Variations in household expenditures per capita across consumption expenditure groups result from various factors. Specifically, spending per individual is the product of the number of illnesses reported per person, the share of reported illnesses treated, and the level of expenditure per treatment. Table 8 presents the results of this decomposition. Private per capita outlays followed an upward trend because expenditure per treatment rose and because the share of reported illnesses that were treated increased as total household per capita consumption increased. By contrast, reported illnesses appear not to have varied much with total expenditures, although there was a tendency for reported illnesses to diminish in urban Indonesia and to increase (more markedly) in rural areas.

*Changes in the Distribution of Health Subsidies between 1978 and 1987*

Table 9 compares data for 1978 and 1987 on the distribution of health subsidies to hospitals and public health centers across consumption groups and geographic areas.[11] It should be noted that the underlying population distribution is likely to have altered between the two dates, particularly because of urbanization. The population shares given in the last two rows of table 9 can help in judging the equity of subsidy shares in rural, compared with urban, areas and in Java, compared with the Outer Islands. Data are not available to enable a comparison across consumption groups in specific regions. The last column, which gives the shares for the total Indonesian population, clearly shows that at the national level the distribution of health subsidies became more equitable. The lower 40 percent expenditure group gained substantially. This result appears to have been driven by gains to the urban poor.

The distribution does not suggest that public health care expenditures are well targeted. Geographically, urban areas appropriated much more per capita than did rural areas; (see also table 7). If anything, this disparity appears to have become more pronounced after 1978, particularly so in the Outer Islands. Conversely, the overall share going to rural areas dwindled, most dramatically in rural Java.

The distribution of health sector subsidies became decidedly more equitable after 1978. To make this point more forcefully, it is useful to contrast the results here, based on 1987 patterns of use, with the results of an exercise aimed at making a rough estimate of the distribution of health spending in 1985/86 using the 1978 pattern of use (Griffin 1992). Griffin combined the 1978 utilization incidence with 1985/86 public health expenditures on hospitals and health centers, using essentially the same budget data as have been used here.[12] His approximation produced an extremely skewed distribution in which the poorest 40 percent of the population captured about 17 percent, the middle 30 percent

11. Recall that this comparison is made under the same assumptions as Meesook (1984).

12. The budget data used by Griffin have not been updated to 1986/87 as done here, and it is not clear whether cost recovery has been withheld from the budget totals to get net subsidy amounts.

Table 8. *Monthly Household Per Capita Expenditure on Health Care by Household Consumption Expenditure Decile, Indonesia, 1987*

| | Decile[a] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Geographic area and indicator* | *1 (lowest)* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10 (highest)* |
| *All-Indonesia* | | | | | | | | | | |
| Expenditure per treatment (rupiahs) | 130.05 | 173.29 | 249.41 | 241.61 | 260.65 | 276.09 | 354.35 | 470.73 | 605.99 | 1,564.38 |
| Share of illnesses treated | 0.54 | 0.58 | 0.57 | 0.62 | 0.62 | 0.65 | 0.70 | 0.71 | 0.74 | 0.77 |
| Number of illnesses per person | 0.24 | 0.27 | 0.26 | 0.28 | 0.28 | 0.30 | 0.31 | 0.31 | 0.31 | 0.26 |
| Expenditure per person (rupiahs) | 16.96 | 26.71 | 37.32 | 40.85 | 44.96 | 53.61 | 76.33 | 103.11 | 139.31 | 309.74 |
| *Urban areas* | | | | | | | | | | |
| Expenditure per treatment (rupiahs) | 262.11 | 222.59 | 364.37 | 523.51 | 681.62 | 596.51 | 693.90 | 963.45 | 1,807.02 | 2,434.39 |
| Share of illnesses treated | 0.59 | 0.66 | 0.70 | 0.76 | 0.71 | 0.72 | 0.77 | 0.76 | 0.75 | 0.82 |
| Number of illnesses per person | 0.26 | 0.33 | 0.26 | 0.27 | 0.24 | 0.25 | 0.25 | 0.24 | 0.22 | 0.23 |
| Expenditure per person (rupiahs) | 40.57 | 47.95 | 66.50 | 107.03 | 118.14 | 108.95 | 132.57 | 174.66 | 299.48 | 447.90 |
| *Rural areas* | | | | | | | | | | |
| Expenditure per treatment (rupiahs) | 124.37 | 141.98 | 206.85 | 249.12 | 247.35 | 271.73 | 256.06 | 296.87 | 355.09 | 733.28 |
| Share of illnesses treated | 0.54 | 0.56 | 0.57 | 0.60 | 0.60 | 0.62 | 0.62 | 0.67 | 0.71 | 0.74 |
| Number of illnesses per person | 0.24 | 0.28 | 0.24 | 0.27 | 0.27 | 0.27 | 0.31 | 0.32 | 0.35 | 0.36 |
| Expenditure per person (rupiahs) | 16.04 | 22.17 | 28.64 | 39.72 | 39.50 | 45.18 | 49.41 | 62.71 | 89.03 | 195.90 |

*Note:* Expenditure per person is the product of the number of illnesses reported per person multiplied by the share of illnesses treated multiplied by expenditure per treatment.

a. Deciles group persons by total household consumption expenditure per capita.

*Source:* 1987 SUSENAS data tapes.

Table 9. *Distribution of Government Health Subsidies by Household Consumption Expenditure Group and Area, Indonesia, 1978 and 1987*
(percentage of government health subsidies to hospitals and public health centers)

| *Household economic group and year*[a] | *Java* | | | *Outer Islands* | | | *Indonesia* | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Urban* | *Rural* | *Total* | *Urban* | *Rural* | *Total* | *Urban* | *Rural* | *Total* |
| *Lowest 40 percent* | | | | | | | | | |
| 1987 | 11 | 7 | 18 | 4 | 9 | 13 | 15 | 16 | 31 |
| 1978 | 1 | 14 | 15 | 0 | 4 | 4 | 1 | 18 | 19 |
| *Middle 30 percent* | | | | | | | | | |
| 1987 | 8 | 9 | 17 | 4 | 8 | 12 | 12 | 17 | 30 |
| 1978 | 3 | 21 | 25 | 2 | 9 | 11 | 5 | 31 | 36 |
| *Upper 30 percent* | | | | | | | | | |
| 1987 | 9 | 14 | 23 | 4 | 12 | 16 | 14 | 25 | 39 |
| 1978 | 12 | 15 | 27 | 4 | 14 | 18 | 16 | 29 | 45 |
| *Total for area* | | | | | | | | | |
| 1987 | 29 | 30 | 59 | 13 | 29 | 41 | 41 | 59 | 100 |
| 1978 | 16 | 50 | 67 | 6 | 27 | 33 | 23 | 77 | 100 |
| *Percentage of population in area* | | | | | | | | | |
| 1987 | 20 | 42 | 62 | 8 | 30 | 38 | 27 | 73 | 100 |
| 1978 | 12 | 52 | 64 | 7 | 29 | 36 | 19 | 81 | 100 |

a. Households are grouped by household consumption expenditure per capita.
*Source:* 1987 SUSENAS data tapes; Meesook (1984).

captured 31 percent, and the wealthiest 30 percent captured about 52 percent of total health care outlays. By contrast, the shares shown in table 9 are 31, 30, and 39 percent, respectively. The bias in Griffin's results arises from his assumption that the pattern of utilization was static.

## V. Conclusions

The past fifteen years have witnessed a concerted government effort to increase the aggregate provision of basic health care services in Indonesia. Little is known, however, about differences in access to and utilization of these publicly provided services and, hence, about how the benefits of health expenditures are distributed across socioeconomic groups. This article has characterized the profile of the utilization of health facilities and the incidence of health sector subsidies using household-level data for 1987. It has also examined how the utilization and subsidy incidence profiles have changed between the late 1970s and 1987—a period that has seen a steady fall in absolute income poverty in Indonesia.

The health sector has undergone significant changes. Public policy efforts at achieving widespread provision of primary health care in rural Indonesia are reflected in the utilization data. The article finds that there has been increased recourse to medical service by all those who report being ill—whether poor or otherwise—together with a drop in the use of traditional medicine. The changes since 1978 are most striking for the poorest groups. Nonetheless, in 1987 it remained true that whether an illness resulted in outpatient or inpatient care was highly correlated with living standards; the likelihood of visiting a private doctor or a hospital was still lower for the poor.

The use of primary health centers in rural areas spread and equalized over the consumption groups. The poorer groups used these services much more in 1987 than they did in 1978. In rural areas in 1987, rich and poor appeared to be equally likely to seek treatment in these facilities. This result suggests that public subsidies to primary health care centers in rural Indonesia are not as pro-poor as seems to be widely believed. It also suggests that a more pro-poor distribution of benefits would require price discrimination in the absence of increased private sector provision, although it is unclear how feasible that would be in rural areas. In 1987 the usage of health centers in urban areas contrasted with that in rural areas in that it declined much more with rising living standards, suggesting self-selection. Subsidized primary health care was therefore more pro-poor in the urban sector.

The overall health subsidy is found to be progressive (in that it tends to represent a larger relative share of consumption expenditure by the poor), but only mildly so. Subsidies are not particularly well targeted in that the absolute subsidy received tends to be higher for the nonpoor. As expected, the incidence of subsidies to hospitals increased with consumption, but for primary health centers the incidence was generally constant across deciles. For all but the high-

est consumption groups, public spending on health care exceeded private spending on health care.

All in all, usage patterns altered enough to make the distribution of public subsidies to the health sector more equitable in 1987 than in 1978. The lowest 40 percent of the consumption distribution experienced considerable gains, primarily among the urban poor. Although the aggregate distribution of the subsidies improved after 1978, benefits in 1987 were still far from being focused on the poor. Urban areas continued to be relatively favored and rural ones to be shortchanged, and this tendency appears to have increased.

From the point of view of efforts to alleviate poverty in Indonesia, the article's findings indicate that—within the health sector—subsidies to basic primary health care provide the best instrument for reaching the poor. But on the basis of the recent usage patterns reviewed in this article, even this instrument is far from ideal.

## References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Anand, Sudhir, and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services." *Journal of Economic Perspectives* 7(1):133–50.

BPS (Biro Pusat Statistik). 1983. *Statistik Indonesia: 1982*. Jakarta.

———. 1990. *Statistik Indonesia: 1989*. Jakarta.

Chernikovsky, Dov, and Oey A. Meesook. 1986. "Utilization of Health Services in Indonesia." *Social Science and Medicine* 23(6):611–20.

Cornes, Richard. 1992. "Measuring the Distributional Impact of Public Goods." In Dominique van de Walle and Kimberly Nead, eds., "Public Spending and the Poor: Theory and Evidence." World Bank, Policy Research Department, Washington, D.C. Processed.

Drèze, Jean, and Amartya K. Sen. 1989. *Hunger and Public Action*. Oxford: Oxford University Press.

Gani, Ascobat, Mardiati Najib, and Mary Wangsaraharja. 1988. "Indonesia Rural Health Services Cost Study. Final Report." Development Studies Project, University of Indonesia Faculty of Public Health, Jakarta, and the Johns Hopkins University School of Public Health, Baltimore, Md. Processed.

Griffin, Charles. 1992. *Health Care in Asia: A Comparative Study of Cost and Financing*. World Bank, Washington, D.C.

Meerman, Jacob. 1979. *Public Expenditure in Malaysia: Who Benefits and Why?* New York: Oxford University Press.

Meesook, Oey A. 1984. *Financing and Equity in the Social Sectors in Indonesia: Some Policy Questions*. World Bank Staff Working Paper 703. Washington, D.C.

Ministry of Health. 1991a. "Study on Operational and Maintenance Cost of Government Hospitals, Fiscal Year 1988/1989." Health Sector Financing Project Monograph 8. Bureau of Planning.

———. 1991b. "Analysis of Health Financing in Indonesia, 1982/83–1986/89 (Data Updating)." Health Sector Financing Monograph 10. Bureau of Planning.

Ravallion, Martin, and Monika Huppi. 1991. "Measuring Changes in Poverty: A Methodological Case Study of Indonesia during an Adjustment Period." *The World Bank Economic Review* 5(1):57–82.

Selden, Thomas, and Michael Wasylenko. 1992, "Benefit Incidence Analysis in Developing Countries." WPS 1015. World Bank, Country Economics Department, Washington, D.C. Processed.

Selowsky, Marcelo. 1979. *Who Benefits from Government Expenditures? A Case Study of Colombia.* New York: Oxford University Press.

UNDP (United Nations Development Programme). 1990. *Human Development Report.* Oxford: Oxford University Press.

van de Walle, Dominique. 1988. "On the Use of the SUSENAS for Modelling Consumer Behavior." *Bulletin of Indonesian Economic Studies* 24:107–22.

Walsh, R. P. D. 1981. "The Nature of Climactic Seasonality." In Robert Chambers, Richard Longhurst, and Arnold Pacey, eds., *Seasonal Dimensions to Rural Poverty.* London: Frances Pinter.

Wirakartakusumah, Djuhari M., Prijono Tjiptoherijanto, Azwini Kartoyo, Ricardi W. Alibasjah, Darwis Hartono, Eko Ganiarto, and Erinos Muslim Tanjung. 1988. "Phase II Evaluation and Analysis of Hospital Costs." Bureau of Planning, Department of Health, and University of Indonesia, School of Economics, Demographic Institute, Jakarta. Processed.

World Bank. 1990. *World Development Report 1990.* New York: Oxford University Press.

———. 1991a. *Indonesia: Health Planning and Budgeting.* A World Bank Country Study. Washington, D.C.

———. 1991b. *Indonesia: A Strategy for a Sustained Reduction in Poverty.* A World Bank Country Study. Washington, D.C.

———. 1993. "Indonesia: Public Expenditures, Prices, and the Poor." Country Department III East Asia and Pacific Region, Washington, D.C. Processed.

Yahya, Suyono, and Runizar Roesin. 1990. "Indonesia: Implementation of the Health for All Strategy." In E. Tarimo and Andrew Creese, eds., *Achieving Health for All by the Year 2000: Midway Reports of Country Experiences.* Geneva: International Labour Office.

# *World Bank Publications Order Coupon*

**CUSTOMERS IN THE UNITED STATES:**
Complete this coupon and return to

**The World Bank**
**Box 7247-8619**
**Philadelphia, PA 19170-8619**
**U.S.A.**

Charge by credit card by calling (202) 473-1155 or fax this completed order coupon to (202) 676-0581.

**CUSTOMERS OUTSIDE THE UNITED STATES:**
Contact your local World Bank Publications distributor for information on prices in local currency and payment terms. (See opposite page for a complete list of distributors.) If no distributor is listed for your country, use this order form and return it to the U.S. address. **Orders that are sent to the U.S. address from countries with distributors will be returned to the customer.**

| Quantity | Title | Stock Number | Price | Total |
|---|---|---|---|---|
| | Adjustment in Africa (full report), English | 60994 | $19.95 | |
| | Adjustment in Africa (full report), French | 12530 | $19.95 | |
| | Adjustment in Africa (summary), English | 12795 | $6.95 | |
| | Adjustment in Africa (summary), French | 12796 | $6.95 | |
| | Adjustment in Africa: Lessons from Country Case Studies | 12787 | $26.95 | |

Subtotal US$ _____

Postage and handling* US$ _____

Total US$ _____

* If a purchase order is used, actual postage will be charged. If payment is by check or credit card, postage and handling charges are US$5.00 per order. For air mail delivery outside the U.S., include US$8.00 for the first item and US$6.00 for each additional item.

**CHECK METHOD OF PAYMENT**

❑ Enclosed is my check payable to The World Bank.

❑ Charge my ❑ VISA ❑ MasterCard ❑ American Express

_____
Credit card account number

_____ _____
Expiration date Signature

❑ Bill me. (Institutional customers only. Purchase order must be included.)

**PLEASE PRINT CLEARLY**

Name _____

Firm _____

Address _____

City _____ State _____ Postal code _____

Country _____ Telephone _____

1041

# *Distributors of World Bank Publications*

*Prices and terms vary by country*

**ARGENTINA**
Carlos Hirsch, SRL
Galeria Guemes
Florida 165, 4th Floor-Ofc. 453/465
1333 Buenos Aires

**AUSTRALIA, PAPUA NEW GUINEA, FIJI, SOLOMON ISLANDS, VANUATU, AND WESTERN SAMOA**
D.A. Information Services
648 Whitehorse Road
Mitcham 3132
Victoria

**AUSTRIA**
Gerold and Co.
Graben 31
A-1011 Wien

**BANGLADESH**
Micro Industries Development Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi R/Area
Dhaka 1209

*Branch offices:*
Pine View, 1st Floor
100 Agrabad Commercial Area
Chittagong 4100

76, K.D.A. Avenue
Kulna 9100

**BELGIUM**
Jean De Lannoy
Av. du Roi 202
1060 Brussels

**CANADA**
Le Diffuseur
151A Boul. de Mortagne
Boucherville, Québec
J4B 5E6

Renouf Publishing Co.
1294 Algoma Road
Ottawa, Ontario K1B 3W8

**CHILE**
Invertec IGT S.A.
Av. Santa Maria 6400
Edificio INTEC, Of. 201
Santiago

**CHINA**
China Financial & Economic Publishing House
8, Da Fo Si Dong Jie
Beijing

**COLOMBIA**
Infoenlace Ltda.
Apartado Aereo 34270
Bogota D.E.

**COTE D'IVOIRE**
Centre d'Edition et de Diffusion Africaines (CEDA)
04 B.P. 541
Abidjan 04 Plateau

**CYPRUS**
Center of Applied Research
Cyprus College
6, Diogenes Street, Engomi
P.O. Box 2006
Nicosia

**DENMARK**
SamfundsLitteratur
Rosenoerns Allé 11
DK-1970 Frederiksberg C

**DOMINICAN REPUBLIC**
Editora Taller, C. por A.
Restauración e Isabel la Católica 309
Apartado de Correos 2190 Z-1
Santo Domingo

**EGYPT, ARAB REPUBLIC OF**
Al Ahram
Al Galaa Street
Cairo

The Middle East Observer
41, Sherif Street
Cairo

**FINLAND**
Akateeminen Kirjakauppa
P.O. Box 128
SF-00101 Helsinki 10

**FRANCE**
World Bank Publications
66, avenue d'Iéna
75116 Paris

**GERMANY**
UNO-Verlag
Poppelsdorfer Allee 55
53115 Bonn

**HONG KONG, MACAO**
Asia 2000 Ltd.
46-48 Wyndham Street
Winning Centre
7th Floor
Central Hong Kong

**HUNGARY**
Foundation for Market Economy
Dombovari Ut 17-19
H-1117 Budapest

**INDIA**
Allied Publishers Private Ltd.
751 Mount Road
Madras - 600 002

*Branch offices:*
15 J.N. Heredia Marg
Ballard Estate
Bombay - 400 038

13/14 Asaf Ali Road
New Delhi - 110 002

17 Chittaranjan Avenue
Calcutta - 700 072

Jayadeva Hostel Building
5th Main Road, Gandhinagar
Bangalore - 560 009

3-5-1129 Kachiguda Cross Road
Hyderabad - 500 027

Prarthana Flats, 2nd Floor
Near Thakore Baug,
Navrangpura
Ahmedabad - 380 009

Patiala House
16-A Ashok Marg
Lucknow - 226 001

Central Bazaar Road
60 Bajaj Nagar
Nagpur 440 010

**INDONESIA**
Pt. Indira Limited
Jalan Borobudur 20
P.O. Box 181
Jakarta 10320

**IRAN**
Kowkab Publishers
P.O. Box 19575-511
Tehran

**IRELAND**
Government Supplies Agency
4-5 Harcourt Road
Dublin 2

**ISRAEL**
Yozmot Literature Ltd.
P.O. Box 56055
Tel Aviv 61560

**ITALY**
Licosa Commissionaria
Sansoni SPA
Via Duca Di Calabria, 1/1
Casella Postale 552
50125 Firenze

**JAPAN**
Eastern Book Service
Hongo 3-Chome, Bunkyo-ku
113
Tokyo

**KENYA**
Africa Book Service (E.A.) Ltd.
Quaran House, Mfangano
Street
P.O. Box 45245
Nairobi

**KOREA, REPUBLIC OF**
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun
Seoul

Korean Stock Book Centre
P.O. Box 34
Yeoeida
Seoul

**MALAYSIA**
University of Malaya Cooperative
Bookshop, Limited
P.O. Box 1127, Jalan Pantai
Baru
59700 Kuala Lumpur

**MEXICO**
INFOTEC
Apartado Postal 22-860
14060 Tlalpan, Mexico D.F.

**NETHERLANDS**
De Lindeboom/InOr-
Publikaties
P.O. Box 202
7480 AE Haaksbergen

**NEW ZEALAND**
EBSCO NZ Ltd.
Private Mail Bag 99914
New Market
Auckland

**NIGERIA**
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan

**NORWAY**
Narvesen Information Center
Book Department
P.O. Box 6125 Etterstad
N-0602 Oslo 6

**PAKISTAN**
Mirza Book Agency
65, Shahrah-e-Quaid-e-Azam
P.O. Box No. 729
Lahore 54000

**PERU**
Editorial Desarrollo SA
Apartado 3824
Lima 1

**PHILIPPINES**
International Book Center
Suite 1703, Cityland 10
Condominium Tower 1
Ayala Avenue, H.V. dela
Costa Extension
Makati, Metro Manila

**POLAND**
International Publishing
Service
Ul. Piekna 31/37
00-677 Warzawa

*For subscription orders:*
IPS Journals
Ul. Okrezna 3
02-916 Warszawa

**PORTUGAL**
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon

**SAUDI ARABIA, QATAR**
Jarir Book Store
P.O. Box 3196
Riyadh 11471

**SINGAPORE, TAIWAN, MYANMAR,BRUNEI**
Gower Asia Pacific Pte Ltd.
Golden Wheel Building
41, Kallang Pudding, #04-03
Singapore 1334

**SOUTH AFRICA, BOTSWANA**
*For single titles:*
Oxford University Press
Southern Africa
P.O. Box 1141
Cape Town 8000

*For subscription orders:*
International Subscription
Service
P.O. Box 41095
Craighall
Johannesburg 2024

**SPAIN**
Mundi-Prensa Libros, S.A.
Castello 37
28001 Madrid

Librería Internacional AEDOS
Consell de Cent, 391
08009 Barcelona

**SRI LANKA AND THE MALDIVES**
Lake House Bookshop
P.O. Box 244
100, Sir Chittampalam A.
Gardiner Mawatha
Colombo 2

**SWEDEN**
*For single titles:*
Fritzes Fackboksforetaget
Regeringsgatan 12, Box 16356
S-103 27 Stockholm

*For subscription orders:*
Wennergren-Williams AB
P. O. Box 1305
S-171 25 Solna

**SWITZERLAND**
*For single titles:*
Librairie Payot
Case postale 3212
CH 1002 Lausanne

*For subscription orders:*
Librairie Payot
Service des Abonnements
Case postale 3312
CH 1002 Lausanne

**THAILAND**
Central Department Store
306 Silom Road
Bangkok

**TRINIDAD & TOBAGO, ANTIGUA BARBUDA, BARBADOS, DOMINICA, GRENADA, GUYANA, JAMAICA, MONTSERRAT, ST. KITTS & NEVIS, ST. LUCIA, ST. VINCENT & GRENADINES**
Systematics Studies Unit
#9 Watts Street
Curepe
Trinidad, West Indies

**UNITED KINGDOM**
Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU34 2PG
England

*Coming in future issues of*

# THE WORLD BANK ECONOMIC REVIEW

Articles on . . .

- **The Political Economy of Growth**
  *by Alberto Alesina and Roberto Perotti*

- **A Presumptive Pigovian Tax: Complementing Regulation to Mimic an Emissions Fee**
  *by Gunnar S. Eskeland*

- **Changing Labor Market Conditions and Economic Development in Hong Kong, Republic of Korea, Singapore, and Taiwan, China**
  *by Gary S. Fields*

- **Capital Mobility in Developing Countries**
  *by Peter J. Montiel*

- **Determinants of the Welfare Costs of Price Controls in Poland**
  *by David G. Tarr*