Mark J. Roberts and James R. Tybout on Producer Turnover and Productivity Growth

Robert M. Buckley and Eugene N. Gurenko on Housing and Income Distribution in Russia

Shantayanan Devarajan, Lyn Squire, and Sethaput Suthiwark-Nareepat on Reorienting Project Appraisal

Jeffrey S. Hammer on Economic Analysis for Health Projects

Arnold C. Harberger on New Frontiers in Project Evaluation

Edwin G. West on Education Vouchers

Martin Carnoy A Comment on West
The World Bank Research Observer

EDITOR Moshe Syrquin
COEDITORS Shantayanan Devarajan, Shahid Yusuf
CONSULTING EDITOR Elinor Berg

EDITORIAL BOARD Angus Deaton (Princeton University); Barry Eichengreen (University of California, Berkeley); Howard Pack (University of Pennsylvania); Claire Liukula (International Monetary Fund); Martha De Melo, Gershon Feder, Gregory K. Ingram, George Psacharopoulos, Michael Walton (World Bank)

The World Bank Research Observer is intended for anyone who has a professional interest in development. Observer articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the Observer considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the editor. Please direct all editorial correspondence to the Editor, The World Bank Research Observer, at the address in the copyright notice below.

The views and interpretations expressed in this journal 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.

The World Bank Research Observer is published twice a year in February and August. Single copies may be purchased for US$12.95. Subscription rates are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Individuals</th>
<th>Institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-year subscription</td>
<td>US$25</td>
<td>US$40</td>
</tr>
<tr>
<td>2-year subscription</td>
<td>US$45</td>
<td>US$75</td>
</tr>
<tr>
<td>3-year subscription</td>
<td>US$60</td>
<td>US$100</td>
</tr>
</tbody>
</table>

Orders should be sent to: World Bank Publications, Box 7247–7956, Philadelphia, Pennsylvania 19170–7956, U.S.A. Readers with mailing addresses in developing countries and in socialist economies in transition are eligible to receive complementary subscriptions on request. Written confirmation is required every two years to renew such subscriptions.

© 1997 The International Bank for Reconstruction and Development | THE WORLD BANK
1818 H Street, N.W., Washington, D.C. 20433, U.S.A.
All rights reserved
Manufactured in the United States of America
ISBN 0-8213-3784-8; ISSN 0257-3032

Material in this journal is copyrighted. Requests for permission to reproduce portions of it should be sent to the Office of the Publisher at the address in the copyright notice above. The World Bank encourages dissemination of its work and will normally give permission promptly and, when the reproduction is for noncommercial purposes, without asking a fee. Permission to copy portions for classroom use is granted through the Copyright Clearance Center, Suite 910, 222 Rosewood Drive, Danvers, Massachusetts 01923, U.S.A.

The World Bank Research Observer is indexed by the Journal of Economic Literature, the Standard Periodical Directory, Applied Social Science Index, the Index to International Statistics, the Public Affairs Information Service, Inc., the Social Sciences Citation Index, and, online, by the Economic Literature Index and Dialog. It is available in microform through University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.
Producer Turnover and Productivity Growth in Developing Countries
Mark J. Roberts and James R. Tybout

Housing and Income Distribution in Russia: Zhivago’s Legacy
Robert M. Buckley and Eugene N. Gurenko

AN EXCHANGE ON PROJECT EVALUATION
Beyond Rate of Return: Reorienting Project Appraisal
Shantayanan Devarajan, Lyn Squire, Sethaput Suthiwart-Narueput

Economic Analysis for Health Projects
Jeffrey S. Hammer

New Frontiers in Project Evaluation: A Comment on Devarajan, Squire, and Suthiwart-Narueput
Arnold C. Harberger

POINT-COUNTERPOINT
Education Vouchers in Principle and Practice: A Survey
Edwin G. West

Is Privatization through Education Vouchers Really the Answer?
A Comment on West
Martin Carnoy
PRODUCER TURNOVER
AND PRODUCTIVITY GROWTH
IN DEVELOPING COUNTRIES

Mark J. Roberts
James R. Tybout

The reallocation of resources, either across sectors or across producers within a sector, can serve as a potential source of productivity growth. New research findings exploit comprehensive microeconomic data on the manufacturing sectors of Chile, Colombia, and Morocco to document resource shifts as producers enter, expand, contract, and exit operation. The micro-level adjustment is substantial; between 25 and 30 percent of the total number of manufacturing jobs turn over each year. In the short run, the productivity effects of this turnover are modest because the new plants that come on line are only slightly more productive than the ones they replace—and both are typically small. In the longer term, however, the turnover generates more substantial increases in productivity because the new firms that survive record substantial productivity gains in their early years. Moreover, firms that exit are typically on a downward productivity spiral and would probably have dragged down sectoral efficiency farther if they had continued in operation.

Many developing countries in the process of structural transformation are struggling to catch up technologically, and their labor markets face the formidable task of moving workers among diverse occupations. At the same time, various frictions inhibit factor mobility in the industrial sector; these include severance laws that prevent firms from firing workers and regulations that limit the establishment of new firms and the termination of old ones. The movement of capital and labor is further constrained by credit-market imperfections, noncompetitive markets, and limited information about technological advances. Many resources might earn higher social returns if they were redeployed in different activities, and structural changes that work to this end might generate large efficiency gains. Similarly, if rates of productivity growth differ across sectors, resource shifts...
from low- to high-productivity sectors might significantly improve productivity growth throughout the economy.

This notion that growth can be generated by reallocating resources more efficiently has long interested development economists. It is well documented that broad shifts in the composition of output take place as the development process unfolds, typically shifting production away from natural resource-based products and toward manufacturing and service sectors (Chenery 1979; Chenery, Robinson, and Syrquin 1986; and Syrquin 1988). Several studies have found that this structural change is often associated with substantial gains in productivity, but little is known about the resource reallocations that occur within sectors as producers enter and exit and their respective market shares change. As Kuznets (1979) and Syrquin (1984) note, if these processes move resources from less efficient to more efficient plants within the same sector, the gains in productivity may be substantial, but the source of these gains is impossible to identify using aggregate data.1

This paper summarizes recent research on the magnitude and implications of this micro-level reallocation of resources in semi-industrial countries. A common theme of this research is that individual producers within the same industry differ in efficiency and thus the reallocation of resources from less-efficient to more-efficient producers offers the potential for improved economic performance. The research is based on panel data that cover virtually all the manufacturers with at least ten workers in three countries: Chile, Colombia, and Morocco. (Other, more limited data from Mexico, Turkey, and Venezuela are also used.) These data make it possible to track individual producers as they enter, expand or contract production, and exit. Differences in productivity can be measured, and the effect of micro-level resource reallocations on aggregate productivity can be quantified.

Several robust patterns of resource reallocation exist. First, a tremendous amount of adjustment at the micro level is completely masked in aggregate or sectoral data. Each year the entry of new manufacturing plants and the growth of existing plants create new jobs that average between 13 and 19 percent of total employment in the manufacturing sector. At the same time the contraction and closing of other plants are responsible for the simultaneous loss of between 12 and 14 percent of total employment. This high rate of employment reallocation among manufacturing plants is present in all the countries studied and persists in each year throughout the business cycle, reflecting a vigorous process of micro-level adjustment.

Second, the patterns of plant turnover partly reflect differences in productivity across plants. On average, exiting plants are less productive than surviving ones, and entering plants are less productive than more experienced incumbents. As new plants mature, however, their average productivity tends to increase for several years until they reach industry norms. Overall, the empirical results reveal a continual process of resource reallocation that moves resources from less-efficient to more-efficient producers within the same industry and that contributes to long-run improvements in economic performance.
In this article we first review the empirical evidence on patterns of producer turnover, describing the creation and destruction of jobs and the entry, growth, and exit patterns of manufacturing plants. We then summarize the evidence on differences in productivity among plants and the implications of turnover for productivity growth. Finally, we address some of the reasons for the differences in efficiency across producers.

**Producer Turnover and Employment Flows**

It is useful to think of the reallocation of resources as arising from three different forces. The first source, which has received the most attention in the literature on development, is long-run structural shifts in technology, endowments, and demand. These forces generate an expansion of output and attract new producers in some sectors, while inducing a contraction in output and net exit by producers in other sectors. Typically, industrial sectors in developing countries gradually shift out of assembling low-technology manufactured goods and move into more sophisticated products that are relatively intensive in human and physical capital.

The second source of resource reallocations is short-run or cyclical fluctuations in demand that might arise from changing macroeconomic conditions or external market shocks. These cyclical fluctuations may be felt in all sectors, but their effects are likely to be particularly important in industries that have low sunk costs because in such cases short-term, or hit-and-run, entry may be profitable. In industrial countries cyclical patterns of job creation and destruction have been interpreted to have a "cleansing" effect because in recessions resources are released from those activities that produce the lowest return and are subsequently reemployed more productively when the economy expands (Caballero and Hammour 1994).

The third source of resource reallocation stems from market forces that create continual producer turnover within an industry even when macroeconomic conditions are stable. This phenomenon has only recently been formally modeled in the literature (see Jovanovic 1982; Lambson 1991, 1992; Hopenhayn 1992; and Pakes and Ericson 1995). It derives from the fact that sunk entry and exit costs, combined with uncertainty, make it possible for producers at different levels of efficiency to co-exist in the same industry. (Entry costs include licensing fees and irreversible purchases of capital goods; exit costs may include bankruptcy expenses or severance payments to employees.) Differences in efficiency arise from differences in managerial abilities or random variation in the returns on past investments in capital or technology. A given firm is uncertain about its relative efficiency and learns about it through market experience or by observing the outcomes of investment projects. Those companies that find they are relatively inefficient or that invest in unproductive assets eventually exit, while those that find they are efficient or that their investments are productive.
survive and expand. New firms continually enter and try their hand at competing with the incumbents. The speed of this turnover process is affected by market conditions and by the magnitude of the sunk entry and exit costs involved. Institutional frictions such as severance laws and credit constraints also shape turnover patterns.

To study all three types of resource reallocation requires examining all the producers in a sector—not simply large continuing producers. Comprehensive panel data from Chile, Colombia, and Morocco were used to identify entering, incumbent, and exiting producers, impute productivity trajectories for each producer, and calculate market shares. Figure 1 summarizes the magnitude of annual job creation and destruction in the manufacturing sector of each country.²

Figure 1. Job Creation and Destruction

Note: Sample periods are Chile, 1979–86; Colombia, 1977–89; Morocco, 1984–89; Canada and United States, 1973–86. The annual rate of job creation is the number of employment positions added during a year, expressed as a percentage of total manufacturing employment at the start of the year. Similarly, the annual job destruction rate is the number of jobs lost during the year, expressed as a percentage of employment at the beginning of the year. The annual rate of employment growth in the manufacturing sector, referred to as net job creation, is the difference between jobs created and jobs destroyed.

Economic conditions in these three countries were typical of those in most semi-industrial countries in the 1980s, so the findings are probably representative of a broader group of nations. Each of the three countries began the decade with an overvalued currency and was forced to devalue and contract during the debt crisis. By the end of the sample period, each country had undergone some degree of structural adjustment and resumed growing.

There were, however, important differences among the three. Chile suffered a major financial crisis in the early 1980s because the manufacturing sector had become heavily indebted in dollars. Its contraction was severe, with unemployment reaching almost 30 percent and large-scale shutdowns of manufacturing plants. Nonetheless, economic policies remained laissez faire, with low tariffs, almost no nontariff barriers, very little public ownership in manufacturing, and little intervention in the labor market. Colombia’s recession was much milder, but its commercial policy remained more protectionist. The data base for Morocco does not begin until after the recession, so it describes only the prolonged recovery. During that time the government promoted manufactured exports with various tax exemptions but maintained some degree of protection from imports. These differences, as well as variations in the length of the sample periods and the degree of industrialization in each country, probably led to some differences in the reallocation of resources over the long term and in the volume of intra-industry turnover.

**Cyclical Fluctuations**

One way to quantify micro-level resource flows is to look at the creation and destruction of jobs (see figure 1). In Chile, Colombia, and Morocco, new jobs were created at an annual average rate of 13 to 19 percent of total employment in manufacturing, while the average rate at which jobs were eliminated varied from 12 to 14 percent. These rates were remarkably similar across the three countries, considering the very different macroeconomic conditions that prevailed. Together, the average number of new manufacturing positions that were added and existing positions that were lost came to 26 to 30 percent of total manufacturing employment in an average year—somewhat more volatility than one finds in the United States and Canada. One explanation for this result is that macroeconomic shocks were longer in the semi-industrial countries. Nonetheless, annual figures (table 1) show that most job creation and destruction takes place simultaneously at all phases of the business cycle, implying that inter- and intra-industry turnover together are more important than the effects of the business cycle.

If the costs of changing employers are similar for workers in semi-industrial countries and in the United States and Canada, the rapid rate of job turnover implies that the average adjustment burden per worker is relatively high. These costs may be somewhat offset, however, by the high geographic concentration of manufacturing activity in the semi-industrial countries, which makes it less likely that workers will need to move as employment demand shifts.

Mark J. Roberts and James R. Tybout
Table 1. *Job Creation and Job Destruction by Phase of the Business Cycle* (percent)

<table>
<thead>
<tr>
<th>Country</th>
<th>Rate of job growth</th>
<th>Gross job additions</th>
<th>Gross job losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>8.7</td>
<td>17.6</td>
<td>−8.9</td>
</tr>
<tr>
<td>Colombia</td>
<td>2.8</td>
<td>13.7</td>
<td>−11.0</td>
</tr>
<tr>
<td>Morocco</td>
<td>6.5</td>
<td>18.6</td>
<td>−12.1</td>
</tr>
<tr>
<td>Canada</td>
<td>2.6</td>
<td>11.4</td>
<td>−8.8</td>
</tr>
<tr>
<td>United States</td>
<td>3.2</td>
<td>11.1</td>
<td>−7.9</td>
</tr>
</tbody>
</table>

*Average during years of employment expansion*

<table>
<thead>
<tr>
<th>Country</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>−8.2</td>
<td>9.4</td>
<td>−17.6</td>
</tr>
<tr>
<td>Colombia</td>
<td>−2.2</td>
<td>11.2</td>
<td>−13.3</td>
</tr>
<tr>
<td>Morocco</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Canada</td>
<td>−3.0</td>
<td>9.1</td>
<td>−12.1</td>
</tr>
<tr>
<td>United States</td>
<td>−5.5</td>
<td>7.4</td>
<td>−12.9</td>
</tr>
</tbody>
</table>

n.a. = Not available.

*Note:* Sample periods are Chile, 1979–86; Colombia, 1977–89; Morocco, 1984–89; Canada and United States, 1973–86.

*Source:* Roberts and Tybout (1996, ch. 2).

A comparison of the annual job creation and job destruction rates between the semi-industrial countries and the United States and Canada reveals an interesting contrast. As the two industrial countries move from recession to expansion, the reduction in the rate of job destruction is larger than the increase in the rate of job creation. As a result total employment turnover is countercyclical—and consistent with the view noted earlier that recessions “cleanse” the production structure (Caballero and Hammour 1994). In the semi-industrial countries job creation rates are equally—or more—sensitive than job destruction rates to fluctuations in aggregate economic activity. That is, the job creation rate differs more between expansionary and contractionary periods than does the job destruction rate. Thus, the dominant cyclical feature is the large increase in the creation of new job opportunities during expansionary periods. One interpretation is that limited access to financial markets in the semi-industrial countries forces plants to rely more heavily on internal finance, so expansion and entry are more sensitive to demand.

*Structural Shifts*

Despite the high turnover during the sample periods shown in figure 1, relatively little change occurred in the net size of the manufacturing sector in each of the three semi-industrial countries. Over the entire sample period, total manufacturing employment changed by an average of only 0.3 percent a year in Colombia, −1 percent in Chile, and 6.5 percent in Morocco. Especially in Chile and Colombia, therefore, intersectoral labor reallocation between manufacturing and...
the rest of the economy (including the pool of unemployed workers) accounted for a small fraction of gross job reallocation flows. This is presumably because the time periods examined are relatively short and the sample countries were fairly industrialized at the beginning of the sample period. In any case, the data do not show the kind of broad intersectoral reallocation of jobs noted in earlier studies.

Yet the industrial sectors of developing countries continually change character, as labor-intensive, light-manufacturing industries give way to more capital-intensive, durable goods industries (Chenery, Robinson, and Syrquin 1986). One might therefore expect a substantial shift in labor flows across industries within the manufacturing sector, even though aggregate manufacturing employment is not changing much. But we found no more shifting of jobs from one manufacturing industry to another in the semi-industrial countries than in the United States. After controlling for the net expansion or contraction of total manufacturing employment, we found that more than 80 percent of the shift of workers employed in manufacturing occurred within, rather than across, industries. That is, the shift in positions from plants that are contracting or failing to plants that are entering or expanding in the same manufacturing industry accounts for more than 80 percent of the annual change in employment on average. Presuming that worker skills are industry-specific rather than employer-specific, this turnover pattern implies that displaced workers require less retraining than they would if they moved to another sector altogether.

**Intra-Industry Turnover**

The dominance of internal flows of employment within the same industry suggests that industrial-evolution models best describe the data. If this is so, much of the job creation and destruction reflects the continual exit of producers who are relatively inefficient and the continual entrance into the same industry of new producers who are, on average, better. Another implication is that sunk entry and exit costs, which are largely dictated by the capital requirements of production, are fundamental determinants of the speed at which this cleansing process unfolds. Producers are reluctant to enter industries with high sunk costs, so incumbent producers are less likely to be driven from the market, and on average, these industries will purge inefficient plants relatively slowly. Further, in these same industries most of the output is controlled by a few producers, and operating profits are high, even when incumbents have no market power (Jovanovic 1982). This pattern tends to support the conclusion that conditions traditionally associated with monopoly rents may also be consistent with competitive behavior and that antitrust action may not be warranted on welfare grounds.

Several other patterns in the data suggest the importance of sunk entry and exit costs. First, the rate of job turnover, defined as the sum of the rates of job creation and destruction, differs substantially across industries, but the ranking
of industries from low to high turnover tends to be very similar across countries. Industrial technology, which is common to all countries, appears to play a large role in shaping this pattern. Second, the high-turnover industries, such as furniture, apparel, food processing, and wood products, are all ones with relatively small-scale production and low capital intensity, while the low-turnover industries, such as steel, chemicals, glass, and paper, are the opposite (table 2).

Job turnover resulting from entry and exit as well as from expansions and contractions in the size of existing plants reflects managerial reactions to the firm’s success or failure in the market or the outcomes of investments—or both. The resource reallocations that accompany each process, however, are subject to different types of frictions. For example, because sunk costs are associated with opening or closing a business, the decision to enter is forward-looking and reflects expectations about the entire future profit stream. Adjustments by incumbents in the number of workers employed and the volume of materials purchased are driven mainly by current profit considerations, however. Policies such as bankruptcy laws, entry licensing requirements, and severance laws affect both entry and exit decisions and decisions on the scale of operations, but in different ways.

Although entry and exit account for about a third of total job turnover in Chile and Morocco, and nearly half the job turnover in Colombia, they are less important in the United States (table 3). In fact, when business-cycle effects are netted out, the difference between turnover rates in the United States and those in the semi-industrial countries can be attributed entirely to the latter group’s rates of entry and exit. These higher rates, in turn, trace to the relative emphasis in those countries on light manufacturing industries, in which entry and exit costs are small. If low entry costs lead to strong competitive pressures, these

Table 2. Average Annual Employment Turnover Rates by Three-Digit ISIC Industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>Percentage turnover</th>
<th>Industry</th>
<th>Percentage turnover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iron and steel</td>
<td>11</td>
<td>Professional/scientific equipment</td>
<td>19</td>
</tr>
<tr>
<td>Industrial chemicals</td>
<td>12</td>
<td>Printing</td>
<td>20</td>
</tr>
<tr>
<td>Glass</td>
<td>12</td>
<td>Nonmetallic mineral products</td>
<td>20</td>
</tr>
<tr>
<td>Ceramic products</td>
<td>12</td>
<td>Leather</td>
<td>20</td>
</tr>
<tr>
<td>Paper</td>
<td>13</td>
<td>Plastic products</td>
<td>20</td>
</tr>
<tr>
<td>Rubber</td>
<td>14</td>
<td>Footwear</td>
<td>21</td>
</tr>
<tr>
<td>Beverages</td>
<td>14</td>
<td>Fabricated metal products</td>
<td>22</td>
</tr>
<tr>
<td>Nonferrous metal refining</td>
<td>14</td>
<td>Nonelectrical machinery</td>
<td>22</td>
</tr>
<tr>
<td>Electrical machinery</td>
<td>16</td>
<td>Furniture</td>
<td>24</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>16</td>
<td>Apparel</td>
<td>24</td>
</tr>
<tr>
<td>Other chemical products</td>
<td>16</td>
<td>Food processing</td>
<td>24</td>
</tr>
<tr>
<td>Textiles</td>
<td>18</td>
<td>Wood products</td>
<td>28</td>
</tr>
</tbody>
</table>

Note: Definition of industry group based on the International Standard Industrial Classification (ISIC) designed to promote international comparability in statistics of economic activity.

Source: Roberts and Tybout (1996).
Table 3. Entry and Exit as a Source of Job Turnover (percent)

<table>
<thead>
<tr>
<th>Country</th>
<th>Total job turnover</th>
<th>Job turnover due to plant entry and exit</th>
<th>Job turnover among incumbent plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>26.5</td>
<td>8.9</td>
<td>17.6</td>
</tr>
<tr>
<td>Colombia</td>
<td>24.7</td>
<td>11.9</td>
<td>13.0</td>
</tr>
<tr>
<td>Morocco</td>
<td>30.8</td>
<td>8.9</td>
<td>21.8</td>
</tr>
<tr>
<td>United States</td>
<td>19.0</td>
<td>3.5</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Average during years of employment expansion

<table>
<thead>
<tr>
<th>Country</th>
<th>Total job turnover</th>
<th>Job turnover due to plant entry and exit</th>
<th>Job turnover among incumbent plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>27.0</td>
<td>10.4</td>
<td>16.6</td>
</tr>
<tr>
<td>Colombia</td>
<td>25.5</td>
<td>11.3</td>
<td>13.2</td>
</tr>
<tr>
<td>Morocco</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>United States</td>
<td>20.3</td>
<td>4.0</td>
<td>16.3</td>
</tr>
</tbody>
</table>

n.a. = Not available.

Source: Roberts and Tybout (1996).

high turnover rates challenge the popular perception that the high concentration of manufacturing in many developing countries reflects less market competition than is seen in industrial countries (Rodrik 1988, Krugman 1989).

One view of the life cycle of individual plants is that new entrants embody the latest technology and thus are more efficient than older plants. These new plants gradually expand and become significant producers but are eventually replaced, in turn, by producers embodying newer technology. Under this scenario, exiting plants should be among the oldest because they are likely to be relying on outdated technology. An alternative view, however, suggests a contrasting pattern. If plants are born with different efficiencies and only learn their relative efficiency gradually as they gain experience, it follows that the efficient ones survive and grow, while the inefficient ones contract and exit. In this case, exit should be concentrated among younger plants, while older plants, having survived a shakedown process, will be the most efficient.

The data on failures by manufacturing plants in semi-industrial countries support the latter view. Figure 2 shows the share of new manufacturing plants in Chile and Colombia that survived each year during their first four years of operation. As each cohort of new plants ages, the proportion of plants that survive clearly increases. For example, the one-year survival rate in Colombia increased from 79.4 percent for one-year-old plants to approximately 87 percent for plants more than three years old. In Chile, first-year survival rates averaged 73.2 percent, while four-year-old plants had survival rates of 89.2 percent.

Additionally, the average size of the surviving members of the cohort increases over time. For example, in a typical year, one-year-old plants in Chile and Colombia are only 26 percent and 39 percent as large, respectively, as the average incumbent plant, but five-year-old plants are 75 percent and 65 percent as large,
respectively. This increase reflects two effects: the growth of the surviving cohort members and the failure of the smaller plants in the cohort. Both factors, however, indicate that in each year, it is the older plants that are the dominant source of industry output.

Overall, the qualitative patterns of plant turnover are similar to those found in the manufacturing sectors of industrial countries: continual waves of small-scale entrants, many of which exit the market within the first few years of their existence. Theory suggests that heterogeneity in profit or efficiency levels and uncertainty on the part of entrants about their future ranking relative to industry norms lie behind these phenomena. In addition, turnover rates differ across industries, with high-entry industries generally characterized by high exit. At a minimum, the turnover patterns we find in the semi-industrial countries—whether measured in terms of jobs or number of plants—imply an environment with substantial resource mobility, much of it occurring among producers within the same industry.

The Relationship between Productivity and Turnover

Turnover-based productivity gains can come from two basic sources. One is the continual exit of relatively inefficient producers and the simultaneous entry of producers who do better. The other source is a reallocation of market share.
from inefficient toward efficient plants. These gains can be compared with the changes in productivity within plants, which have been the focus of most studies of productivity in semi-industrial countries.

To document these processes, two approaches have been used. One simply amounts to constructing output-to-labor ratios, plant by plant (see Tybout 1992). The other begins by estimating a production function describing the relationship between the output of a good and the inputs required to make that good (see Liu and Tybout 1996). Plant-specific productivity in each year is calculated as the ratio of actual output to the output predicted by the production function, assuming a given level of inputs. Once each plant’s productivity trajectory is calculated, it can be used to show the growth of productivity among incumbent plants that have been operating throughout the sample period and the effects of turnover (entry and exit of plants). The former is simply attributable to improved efficiency; the latter is attributable to productivity gaps among incumbents, entering producers (who are in their first year of operation), and exiting producers (who are in their last year of operation). For example, if incumbents are more productive than entering and exiting plants, net entry dampens productivity growth and net exit boosts productivity. And if new producers are more productive than the producers they replace, ongoing turnover is a steady source of productivity gain. Algebraic details are provided in the appendix.

Some of the growth in efficiency among incumbent producers reflects improved productivity. But gains are also generated when resources shift from low- to high-productivity manufacturers, a shift that is generally accompanied by the creation of new jobs and the destruction of old ones. We consider this turbulence arising from the reallocation of market shares to be one source of the gains in efficiency, and we distinguish it from the increase in productivity that occurs within individual plants (see appendix).

Findings on the effects of turnover in industrial countries are quite mixed. The relevant references include Baldwin and Gorecki (1991); Griliches and Regev 1995; Baily, Hulten, and Campbell 1992; and Olley and Pakes 1996; for a summary see Tybout (1996). For the semi-industrial countries, several basic patterns emerge. First, macroeconomic fluctuations can induce significant turnover with far-reaching effects on productivity. Because entry and exit rates vary during the business cycle, so do the market shares of incumbent producers. During upswings incumbents lose market share because new plants enter more rapidly than incumbents fail. This pattern exerts a countercyclic influence on productivity because new and dying firms are typically less productive than continuing producers. Plant exits during the recession in Chile improved labor productivity more than 1 percent, and the rapid entry of inexperienced firms during Morocco’s boom period reduced labor productivity almost 2 percent (Tybout 1992).

Similarly, the productivity effect of the reallocation of market shares can be substantial in the short run, because firms do not all expand and contract proportionately over the course of the business cycle. In Colombia, inefficient plants shrunk relatively more as the economy went into recession and recovered rela-
tively rapidly when aggregate demand rebounded. This countercyclic productivity effect amounted to several percentage points of efficiency gain or loss in some years and industries. The net market share effect over the course of a full business cycle is not typically large, however.

Second, replacing dying plants with new plants also has a rather small average impact on productivity. The average gains during downturns are roughly offset by the losses from turnover during upswings, so most of the measured productivity growth comes from gains in efficiency by incumbent plants. In part, this is because entering and exiting plants account for only 3 to 5 percent of production in a typical year. It is also because the productivity gap between plants in their first year of operation and those in their last year of operation is small. Entering plants are only about 85 percent as productive as the industrywide average, and exiting plants are roughly 80 percent as productive (figure 3).

Third, by focusing on the efficiency gap between exiting and entering plants, one substantially understates the productivity effects of turnover. As noted earlier, the average productivity of each new cohort of plants rises as it matures, reaching industry norms after about four years of experience (see figure 3). Thus an entering cohort of plants eventually becomes substantially more productive than the cohort of exiting plants it replaced, and this latter group might well have gotten worse if it had not exited (Liu 1993; Griliches and Regev 1995). Although the market share of entering and exiting plants in the transition year is small, 20 to 30 percent of the population of plants typically turns over within four years.

---

**Figure 3.** Cohort-Specific Productivity of Manufacturing Plants in Colombia, 1982–86

Total factor productivity

![Bar chart showing productivity for different age groups of plants.](chart)

**Source:** Roberts and Tybout (1996).
In assessing the gains from plant turnover, it is useful to ask what would have happened without any entering or departing plants? This question can only be answered with a forward-looking model of entry and exit decisions, as well as a counterfactual representation of the productivity trajectory for plants that are prevented from exiting. Developing this framework is an important topic for further research. At present we can only point to anecdotal evidence from countries that distort turnover patterns (by subsidizing or limiting entry, or by propping up inefficient producers that would otherwise exit), which suggests that the costs of such distortions are large (Pursell 1990).

Other Aspects of Productivity

In addition to the sources of productivity gains mentioned above, several others have been investigated in the research program summarized herein, including technology transfers and learning spillovers, ownership structure, scale economies, and international trade.

Technology Transfer

Harrison (1996) finds that plants owned by multinationals are typically closer to the efficient production frontier than domestically owned firms. Contrary to earlier studies based on cross-sectional data, however, foreign direct investment does not appear to generate positive spillover effects for domestic firms in the same industry or region. At least in the short run, it appears that multinationals siphon off demand and high-quality labor from domestic competitors.

Ownership Structure

Foroutan (1996) finds that the distinction between private and public ownership is also relevant to productivity levels. In Turkey publicly owned plants are significantly less productive than privately owned plants, and they exhibit qualitatively different responses to trade liberalization. This finding is consistent with the assumption that public sector managers, lacking the disciplining influence of shareholders, pursue objectives such as job security and compensation.

Scale Economies

If there are economies of scale in production, large plants will be more efficient than small ones, so policies that influence the size of manufacturing plants also affect productivity. For example, trade development strategies may increase the size of export-oriented producers by expanding their potential market. Conversely, to the extent that economies of scale exist, the
same policies may reduce scale efficiencies in those firms that compete with imports, since these producers typically contract when trade liberalization increases import penetration in the domestic market (Rodrik 1988). As a source of productivity growth, however, such changes have probably been overemphasized relative to the other dimensions of performance. The largest plants in most industries typically have attained minimum efficient scale, and these are the plants that dominate industrywide performance (Tybout and Westbrook 1996). One implication is that the computable general equilibrium models that have been used to estimate the gains from trade liberalization do not recognize differences in the size of plants within an industry and thus have often overstated the potential gains from scale economies that accompany trade liberalization (Tybout 1993).

**International Trade**

The degree of exposure of the domestic industry to international markets may affect productivity through other channels. Differences in productivity within an industry are typically greater in industries protected from international competition, suggesting that protection nurtures inefficiency. Higher productivity growth generally is associated with the production of tradable goods. These patterns may reflect limited access to foreign technology and expertise as well as problems acquiring imported intermediate and capital goods under protectionist trade regimes. But there are plausible alternative explanations for the negative association between protection and productivity. For example, economic models suggest that sectors with large start-up costs have relatively little turnover and tend not to sort out firms with low productivity. These sectors may also be relatively protected because they are not sectors in which the semi-industrial countries have a comparative advantage.

**Conclusion**

The turnover patterns we document are difficult to reconcile with the view that entry and exit primarily reflect aggregate demand fluctuations or long-term changes in technology. Instead, they seem most consistent with recent theories that emphasize the heterogeneity of producers, the uncertainty each producer faces about its ability to survive, and the constraints on turnover introduced by the sunk costs of entry and exit.

One implication of the evidence cited here is that artificial impediments that prevent failing businesses from going out of business can be very counterproductive, particularly if they are maintained over long periods of time. Mandated severance payments or prohibitions on plant closings not only inhibit intersectoral reallocations, but also tend to discourage transfers that could lead to a more productive use of resources within an industry. It is not even clear that these
restrictive policies are useful in preserving employment. If they prevent the transfer of production to more efficient producers, they may eventually result in a smaller industrial sector. Similarly, restrictions on access to credit, equity, or other financial markets can reduce the entrance of potentially productive plants or the expansion of incumbents. Policies that reduce entry keep relatively inefficient producers in operation and slow the productive transfer of resources.

Another implication is that industrial concentration may be a very poor measure of market power; it is more likely to reflect the magnitude of sunk costs that constrain entry. Industries with high entry costs tend to have high operating profits and typically remain concentrated for long periods of time, even if firms are behaving competitively. Hence, antitrust policies designed to limit producer concentration may simply reduce efficiency.

Finally, that the cyclical component of job flows is small relative to the average level of reallocation in any year suggests that policymakers who focus on the macroeconomic causes of employment fluctuations may miss the largest source of worker transitions. Attention to the search process and associated market failures may be a much more effective means of reducing the duration and frequency of unemployment spells and the associated efficiency losses they entail.

Appendix. Productivity Decompositions

Define $E_{it} = Y_{it} / f(v_{it}, t_0)$ as the efficiency of the $i$th producers in year $t$, where $Y_{it}$ is realized output, $v_{it}$ is its input vector, and $f(v_{it}, t_0)$ is an estimated production function evaluated at the firm’s period $t$ input vector and the technology prevailing in period $t_0$. (In some instances, capital stocks or intermediate inputs are unobservable, so $f(v_{it}, t_0)$ is replaced with a simple measure of factor use, such as total employment.) Industrywide productivity can then be written as a weighted average of the $n$ plant-specific trajectories, $E_t = \sum_{i=1}^n E_{it} \theta_{it}$, where $\theta_{it} = f(v_{it}, t_0) / \sum_{i=1}^n f(v_{it}, t_0)$ is the period $t$ market share of the $i$th producer in terms of factor use.

To isolate turnover-based productivity growth, one can express growth in this industrywide productivity measure as the sum of three components:

$$\frac{\Delta E_t}{E_{t-1}} = \bar{\theta} \left( \frac{\Delta E_c}{E_{t-1}} \right) + \Delta \bar{\theta} \left[ \bar{\theta} \left( \frac{E_{bt} - E_{bt-1}}{E_{t-1}} \right) + \left( \frac{E_{bt} - E_{bt-1}}{E_{t-1}} \right) [1 - \bar{\theta}] \right].$$

Here overbars denote averages over the periods $t-1$ and $t$, $E_c$ is weighted average efficiency among continuing plants (denoted by $i \in c$), $E_{bt}$ is the weighted average efficiency among plants that enter in year $t$ (denoted by $i \in b$), $E_{bt-1}$ is the weighted average efficiency among plants that enter (die) after year $t-1$ (denoted by $i \in d$), and $\bar{\theta}$ is the market share of continuing plants.

Both the second and the third term in equation A1 pick up turnover-based productivity effects. But efficiency gains attributable to resource reallocations...
among incumbent plants are also potentially important and worth isolating. To this end the first right-hand side element of equation A1 can be further decomposed as:

\[
(A2) \quad \frac{\Delta E_{\ell}}{E_{t-1}} = \frac{\bar{\ell}}{E_{t-1}} \left[ \sum_{i \in c} \frac{\bar{\ell}_i}{\ell} \Delta E_{it} + \sum_{i \in c} \frac{\ell_i}{\bar{\ell}_i} \bar{E}_i \right]
\]

where summations are only over continuing plants \((i \in c)\). The first term on the right can be thought of as measuring the intraplant productivity growth effects that are the focus of representative-plant analysis. The second term—market share reallocation effects—picks up productivity gains or losses due to size adjustments among incumbents.

Notes

Mark J. Roberts is professor of economics at Pennsylvania State University and a research associate of the National Bureau of Economic Research. James R. Tybout is professor of economics at Georgetown University and a consultant to the World Bank. Much of the material presented here is based on a World Bank research project that has been incorporated into a book entitled Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure, which the authors edited.

1. Kuznets (1979) observed that the common three-sector dichotomy—agriculture, manufacturing, and services—neglects all within-sector reallocations, which Kuznets said may have been an important reason why the productivity growth in Taiwan was poorly explained by the structural change methodology. Syrquin (1984, p. 95) summarizes this weakness of the traditional aggregate approach: “The estimated contributions of structural change to growth probably underestimate the impact of resource shifts. The broad definitions of sectors, even in fairly disaggregated studies, hides all [intra-sectoral] factor reallocations. . . . This is important for industrialized countries and for rapidly growing economies.”

2. The unit of observation in the Chilean and Colombian data sets is the manufacturing plant. In the Moroccan data set, it is the firm. This is a minor point because firms tend to be small, single-plant operations. In this paper we refer to the observations for all the countries as manufacturing plants.

3. \(Y_{it} = f(v_{it}, t)\), where \(Y\) represents the amount of output attained by the average plant at the input vector \(v\) in period \(t\). Given \(f(\cdot)\), the efficiency of the \(i^{th}\) plant in year \(t\) is then imputed as \(E_{it} = Y_{it} / f(v_{it}, t)\), where \(Y_{it}\) is the realized output of the \(i^{th}\) plant, \(v_{it}\) is its input vector, and the denominator is a benchmark productivity level in period \(t\).

References


16

The World Bank Research Observer, vol. 12, no. 1 (February 1997)


HOUSING AND INCOME DISTRIBUTION IN RUSSIA: ZHIVAGO’S LEGACY

Robert M. Buckley
Eugene N. Gurenko

“I’m very glad you’ve given up those rooms. . . . We should give up still more.”
Dr. Zhivago, p. 170

The inadequacy of housing and its effect on the quality of life go a long way toward explaining many aspects of Russian life, but little data has been available to determine how housing affected the relative well-being of the citizenry. This paper presents comprehensive data examining for the first time the effect that seventy years of a socialist housing allocation scheme had on the distribution of income. It seems clear that housing provided by the government or by employers has a value that can be measured to yield useful inferences about the distribution of income and wealth.

This article shows that housing allocation had a progressive effect on the distribution of income in Russia. In addition, when the imputed value of housing is added to household income, the increase in income inequality that occurred in recent years is significantly reduced. The analysis concludes with a discussion of how housing policy could be used to address poverty concerns, an important aspect of the transition process.

On his return home from World War I, Boris Pasternak’s fictional character Dr. Zhivago finds that he has to share his family’s Moscow mansion with workers and their families. Since at least that time Russians have been sensitive to the effects of housing on living situations. The housing shortage and the difficulties of living in cramped quarters have been a fundamental aspect of Russian life and writings about it. Unfortunately, aside from figures on the number of square meters of housing space produced and surveys based on interviews with émigrés, little statistical material has been available to determine how housing affected the well-being and social position of Russian
citizens. Unlike Hungary, where analysts have produced a considerable body of statistical work describing housing conditions since the late 1960s (Szelenyi 1983; Hegedus 1987; and Daniel 1985), Russian researchers were not able to sift through the data to determine the role of housing in the economy. The only data available from the former Soviet Union were based on the Berkeley-Duke survey of Soviet émigré households, conducted in the late 1970s. As Alexeev (1988, 1990), who examined these data, observes, this survey is not fully representative of the Soviet population. Atkinson and Micklewright (1992) suggest that data on income distribution in the former Soviet Union appeared next to alcoholism and drug addition on the censor's list of prohibited subjects. Not until 1989 did Russia begin to count as income the overall imputed value of housing in the national income accounts (Ivanov, Rjbuskin, and Homenko 1993). And until now comprehensive data on how housing affected income distribution were not available.

The article relies on data from the Russian Longitudinal Monitoring Household Survey conducted in 1992, the year before the passage of significant reforms in housing policy. The survey, funded by the World Bank (World Bank–Goskomstat 1993), was the first nationally representative household survey conducted in Russia and covered 6,128 households. Our analysis of these data indicates that housing allocation had a progressive effect on the distribution of income. In addition, when the imputed value of housing is included as income, the increase in income inequality recorded in recent years is significantly reduced. When housing is excluded from income, the Gini coefficient (a conventional indicator of social inequality that equals zero in the case of perfect equality and one in the case of total inequality) for Russia more than doubled between 1987–88 and 1993–94, according to a study by Milanovic (1996). That was the largest increase in the eighteen countries in the study. When housing is included in income, the increase in inequality, while still large, is significantly reduced.

Our analysis suggests that when income is adjusted for housing, the distribution of income under the Soviet regime was much more compressed than is commonly thought (Bergson 1984; Gregory and Stuart 1989). Indeed, efforts to fulfill Khrushchev's dream of realizing the communist ideal by 1980, to provide workers with compensation according to need rather than performance, resulted in an extensive housing-based income transfer mechanism—a mechanism that was a fundamental part of socialist wage policy and one that significantly reduced income inequality.

To better understand housing's role in Russian welfare, we describe housing conditions and tenure patterns there and show how the inclusion of housing affects various measures of income distribution. We then consider how housing allocation affected the distribution of income and discuss the methodology used to measure imputed income. Finally, we discuss some of the policy implications of the socialist system's allocating such a large share of income through this in-kind payment, and we then speculate about how this aspect of income distribution might be affecting the transition process.
The Russian Housing Stock

The housing conditions of most of Zhivago's contemporaries in prerevolutionary Russia were dismal. A 1912 census of St. Petersburg and Moscow shows eight occupants per apartment; the comparable figure for Paris was 2.7 (Alexeev, Baker, and Westfall 1991). The average living space per urban dweller was estimated at seven square meters. For workers housing was often a bunk in a rough dormitory attached to a factory.

These substandard conditions are no longer the case for the majority of the Russian population. A massive, post-Stalin construction project—one of the largest government-sponsored investment programs ever undertaken—built nearly 20 million apartments in eight years (Morton 1980). Between the end of World War II and the late 1970s, the average living space per capita more than doubled, reaching 16.1 square meters in 1992 (World Bank–Goskomstat 1993).

Despite this enormous investment, the housing shortage was by no means resolved. In 1992 almost 10 percent of households lived either in overcrowded communal flats or dormitories with less than eleven square meters of floor space for each household member (table 1). More than 20 percent of the residents of St. Petersburg and 17 percent of those in Moscow lived communally. For comparison, in 1974 almost 30 percent of Moscow's residents lived communally, and in 1960, 40 percent of Soviet citizens did so (Morton 1980).

Although the vigorous construction programs substantially reduced the shortage of self-contained housing units, the housing shortage—caused at least in part by the massive destruction of World War II—has never been overcome. In contrast to Western countries, where the ratio of households to the number of housing units—a simple measure of overcrowding—was about 0.9 by the early 1970s, the ratio in Russia stood at 1.17 and has increased since then (Alexeev, Baker, and Westfall 1991). Indeed, by 1992 according to this gross quantitative measure, housing conditions were worse in Russia than in either Poland or Albania, countries noted for housing shortages (Guarda 1993).

A recent comparative study of housing conditions shows that qualitative indicators of housing in Russia are far below the average of those for reforming

<table>
<thead>
<tr>
<th>Share of households living in</th>
<th>Russia</th>
<th>Urban settlements</th>
<th>Moscow</th>
<th>St. Petersburg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rentals or sublets</td>
<td>2.1</td>
<td>1.7</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>Dormitories</td>
<td>3.0</td>
<td>3.9</td>
<td>0.9</td>
<td>0.0</td>
</tr>
<tr>
<td>Communal apartments</td>
<td>6.6</td>
<td>7.3</td>
<td>17.0</td>
<td>20.3</td>
</tr>
<tr>
<td>Self-contained apartments</td>
<td>58.9</td>
<td>69.1</td>
<td>81.8</td>
<td>78.8</td>
</tr>
<tr>
<td>Single-family houses</td>
<td>23.0</td>
<td>14.0</td>
<td>0.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Shared houses</td>
<td>5.9</td>
<td>5.7</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Not known</td>
<td>0.5</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

socialist economies in Europe (Hegedus, Tosics, and Mayo 1996). In 1992 only 44 percent of Russian households had access to running hot water; 66 percent, central heating; and 67 percent, indoor sewerage (World Bank–Goskomstat 1993). Even those who managed to escape from communal flats found their new dwellings poorly planned and constructed. As Ruble (1993, p. 234) noted: “Years of labor by residents are frequently required to correct a familiar litany of irksome deficiencies in the superficially modern Soviet high-rise apartment building of the 1990s; persistent elevator break-downs, plummeting water pressure, electrical surges, and upper-story windows that shatter in ‘high winds’ even though all is calm at ground level; the list goes on.”

These deficiencies were embedded in the tenure structure of Russia’s housing sector. The state owned and maintained almost two-thirds of the housing stock, twice the publicly owned share in Poland, almost triple Hungary’s 25 percent share, and considerably larger than the 17.1 percent share in France and the 2.3 percent in the United States (Buckley, Daniel, and Thalwitz 1996). In large cities such as Moscow and St. Petersburg, the state owned almost 90 percent of the housing, in part because private housing construction was prohibited in cities with more than 100,000 residents.

The state allocated to households this publicly owned housing stock without regard to price. Indeed, for the most part, rents were set by a 1926 law, which remained nearly unchanged until 1992. Because the law held nominal rents and utility costs constant, the high inflation that occurred during the transition meant that by the end of 1992, gross housing services were essentially free.

Our approach to measuring housing income allows us to use the available data to show how this housing subsidy was distributed according to income levels and to examine how this transfer affected the distribution of income. As shown in table 2, income had little or no observable effect on distribution of housing space in 1992. The relationship between income and housing allocation was apparently random: 27.9 percent of households in the lowest income category and 23.4 percent of households in the highest income category lived in overcrowded housing conditions—less than fourteen square meters per capita. A detailed comparison of the housing available to other income groups also fails to reveal any noticeable patterns.

<table>
<thead>
<tr>
<th>Housing space per household member (square meters)</th>
<th>Less than 1,500</th>
<th>1,501 to 3,000</th>
<th>3,001 to 4,500</th>
<th>4,501 to 6,000</th>
<th>6,001 to 7,500</th>
<th>More than 7,500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fewer than 7.0</td>
<td>3.8</td>
<td>1.6</td>
<td>2.7</td>
<td>1.9</td>
<td>1.6</td>
<td>1.1</td>
</tr>
<tr>
<td>7.1 to 14</td>
<td>24.1</td>
<td>19.8</td>
<td>31.1</td>
<td>28.5</td>
<td>27.1</td>
<td>22.3</td>
</tr>
<tr>
<td>14.1 to 25</td>
<td>35.0</td>
<td>37.6</td>
<td>43.1</td>
<td>48.9</td>
<td>49.2</td>
<td>48.0</td>
</tr>
<tr>
<td>More than 25.1</td>
<td>37.1</td>
<td>41.0</td>
<td>23.1</td>
<td>20.7</td>
<td>22.1</td>
<td>28.6</td>
</tr>
<tr>
<td>Total (percent)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The Measurement of Imputed Income from Housing

Despite his caveats about potential biases in his data, the distribution of housing space that Alexeev (1990) gleaned from the 1976–79 survey of emigres was still valid approximately fifteen years later. Table 3 shows that the Gini coefficient for housing space did not change from that survey to the 1992 survey. Conversely—and not surprisingly, given the changes in wage policy under Gorbachev and the economic shocks associated with the transition—the 1992 survey found that Alexeev’s measures of the dispersion and variability of all nonhousing income had changed. No longer was there the same compression and constancy of wages. The Gini coefficients and measures of income dispersion both increased sharply, indicating that inequality had grown.

The data show that a rapid surge in income inequality occurred after 1989. The 1985 Gini coefficient for per capita household income is 90 percent of that for 1989 and only 60 percent of the 1992 figure. Further disaggregation shows that in 1992, 31.7 percent of the country’s total net income—excluding housing-related income—went to the richest 10 percent of the population, while only 2.5

<table>
<thead>
<tr>
<th>Variable</th>
<th>Gini coefficient</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income from all sources</td>
<td>0.34</td>
<td>0.42</td>
</tr>
<tr>
<td>Socialist income</td>
<td>0.29</td>
<td>0.38</td>
</tr>
<tr>
<td>Living space</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>Number of rooms</td>
<td>0.24</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Note: The coefficient of variation provides a standardized measure of income distribution. It is equal to the standard deviation divided by the mean.

The definition of household income per capita is the same as that given by Atkinson and Micklewright (1992). The household is defined as a unit comprising one or more persons, including pensioners, children, and the sick and disabled. Household members may have income from sources other than primary employment, such as a second job or self-employment or income in cash or kind from farming private plots; it also includes capital income and state transfers such as pensions or child benefits. The income is considered shared within a household. Thus, the figure is the sum of household-based and individual-level income. First, the wages, net profits from entrepreneurial and individual economic activities, pensions, and unemployment benefits are aggregated over all individuals in a household. This intermediate total is added to household-level income, which consists of net profits from farm activities, subsidies from enterprises and local authorities, family allowances, income from property sold, scholarships, and private transfers (cash amount plus in-kind valuation by respondent). A more detailed breakdown of the household income components can be obtained from the questionnaire and the Statistical Package for the Social Sciences program written to generate this variable. It should be noted that the definition does not account for imputed rental subsidies on owner-occupied property.

The total household floor space per capita is defined as the total housing space occupied by a family, that is, the total floor space of all bedrooms, living spaces, kitchen, bathrooms, lavatories, entry halls, closets, and storage rooms, including unheated areas. The definition also includes any living space at the disposal of any household member in addition to the residence they share. The per capita levels of housing space are obtained by dividing the total housing space per household by the household size.

percent went to the poorest 10 percent. Cumulatively, these results indicate that by 1992 Russia had, as Atkinson and Micklewright (1992) suggest, come to look like a market economy. The Gini coefficient for 1992 was almost identical to the United States figure for 1991 (U.S. Department of Commerce 1995).

But these figures do not include the value of subsidized housing, which was a significant component of the distribution of income. Rents were approximately zero for all but the poorest households and were only slightly higher for them. Citizens who lived in state or enterprise-owned housing paid low or no rent, in effect receiving substantial in-kind (noncash) transfers. To calculate how the benefits of unpriced housing affected the distribution of income, we derived a measure of the value of the rents and added it to each household’s income to obtain a more accurate measure of total household income. We began with Hicks’ (1946, p. 172) formula: “A person’s income is the maximum value he can consume during a week and still be as well off at the end of the week as he was at the beginning.” It is a short definition, one based on theory, and it nicely links current income from wages (or from a profit and loss statement) to the stock of wealth and the consumption associated with changes in it, such as our measure of imputed housing income.

Analyses of household income must be based on official wage statistics. If we augment that information by including the amount that would otherwise be spent on housing, we can link the notion of income to household wealth. Housing is the largest component of noncash income received by households, and it is one that the United Nations recommended be included in household income statistics.

The approach used here to impute the value of housing follows the United Nations (1968, p. 6.22) definition: “The total of owner-occupied dwellings which is to be included in gross output should, in principle, be valued at rent on the market of the same facilities. It may be necessary to approximate the market rent by an estimate which should cover items such as operating, maintenance and repair outlays... depreciation and mortgage interest in addition to interest on the owner’s investment in the dwelling...” Thus, we must first measure how much rent a household would have to pay to enjoy the same housing services provided by the state or enterprise. One way to make this notion concrete is to consider an individual who sells financial assets to buy a house with cash. Apart from such considerations as risk, liquidity, and administrative costs, if the person were as “well off” after buying the house as he was before, the imputed income (rent that the owner of the house no longer pays) would be equal to the annual interest that would have accrued on the financial assets. In terms of either the United Nations or the Hicks definition, this imputed value must be included as income if income is to be measured correctly.

We used two approaches—the first based on market value and the second on opportunity cost—to estimate the income equivalent of subsidized rent. To account for the possibility that shortages produced by the rationing system might induce households to pay more, we also used a third conceptual approach based
on the "minimum" cash compensation required to obtain adequate housing. If this approach were used, the value of imputed rents could well be higher than that estimated by the methodologies described above.

The Market Value Approach

The most commonly used approach to measuring imputed housing income is to assume it is equal to the market rental value of the housing. This estimate takes the value of imputed income to be equal to the market value of an analogous good. Although a rental housing market did not exist in Russia in 1992, it did exist in industrial nations and several developing countries. Horioka's (1994) careful analysis of household consumption and expenditure patterns in members of the Organization for Economic Cooperation and Development in 1989 found that the average household's expenditure on housing was 18.5 percent of household income, and the median was 19.2 percent. Those averages varied little from country to country.

The World Bank Housing Indicators survey (Mayo and Angel 1993) for 1991 yielded similar results. Covering housing characteristics in fifty-two cities in industrial and developing countries, this survey provides a comprehensive set of data on the housing sector, including the cost and availability of key inputs such as land, infrastructure, building materials; the regulatory environment; demographic variables; finance and subsidies; and the qualitative and quantitative features of the housing stock. These data were used to infer the median value of housing services received by a median-income household in countries whose per capita gross domestic product (GDP) was similar to Russia's. This figure was computed in the form of a rent-to-income ratio and was drawn from a regression of rent-to-income ratios on per capita GDP in terms of purchasing power parity.

The Opportunity Cost Approach

The opportunity cost approach is the one used by Smeeding and others (1993) and Yates (1994) in measuring the imputed net income from owner-occupied housing in several Western countries. As noted above, the estimates of rental value for Russia take into account the provision not only of housing but also maintenance, depreciation, and utilities at essentially no cost. Hence, except for households in cooperatives, residents effectively received a 100 percent subsidy of gross rent. Based on Laidler's (1969) or Poterba's (1992) analysis, these gross rents can be converted into a rate of return that is approximately three times the 2 percent net rate of return used by Smeeding and others (1993). The gross rate of return on housing is much higher than the net real rate of return on other assets because operation and maintenance costs (about 2.2 percent) and depreciation (1 to 1.5 percent) reduce housing returns. Thus to realize a 2 percent net real rate of return requires a 5.7 percent gross rate of return (Alexeev 1990).
Renters will now have to cover the costs of operation and maintenance, multiplying their rental charges by a factor of three. Using the assumption of Smeeding and others (1993) that an average household spends about 7 percent of its income on housing, the rental of a dwelling that had cost a household 7 percent of its income will now cost it approximately 20 percent of its income.

Both approaches, then, suggest that 20 percent of income approximates the amount that the median household would have had to spend to rent the median amount of housing space it currently occupies. Because this figure may understate the amount people would actually be willing to pay, we also assumed that gross imputed housing income accounted for 25 percent of income, and considered a lower rent-to-income ratio of 15 percent (as estimated by Alexeev 1990). The direction of change in the Gini coefficients after including housing income is very similar under all three assumptions. As would be expected, the Gini coefficient is lower—that is, income inequality is more equally distributed—when housing accounts for a larger share of income, and conversely.

Because there are no recorded sales prices for housing in 1992, the technique we used ignored variations in the quality of housing. Kaganova and Malgin (1994) and Pusanov (1993) recently examined these variations for St. Petersburg and Moscow, respectively. Their work shows that the assumption of invariant quality overstates the equalizing effects of housing income on the distribution of income. With variation in quality considered, the Gini coefficient for the distribution of housing in St. Petersburg rises from 23.9 to 30.7. That is, higher-income households had significantly better quality units. To compute the value of imputed gross rent for other households at various income levels, we assumed that the amount of imputed income received was directly proportional to the square meters of housing provided. In other words, a family with twenty square meters of space per capita received 11 percent more in imputed housing income than residents with the eighteen square meters per capita—the median amount of housing space. Because the market rent for the median level of housing consumption per capita is equal to 20 percent of median family income, the imputed rent for a twenty-square-meter space amounts to about 22 percent of median family income. We then added this measure of income to the total income of each household in our data set.

Admittedly, this measurement values housing services solely on the basis of space. Previous estimates by the U.S. Central Intelligence Agency (1982) and Bergson (1961) rely on the same approach, however, although the latter used what he terms a “quite arbitrary” means of discounting the quality of private housing relative to public housing. In contrast Prell (1989) attempts to estimate how qualitative improvements in the housing stock in Russia, of the sort that occurred in the United States in the 1960s, may have affected the growth rate of this capital stock.

We did not attempt to make such quality adjustments for two reasons. First, we focus on one point in time rather than on how the level of investment may have changed over time as a result of unmeasured changes in quality. Second,
without more data on housing prices, we are unable to sort out how housing
was distributed across households. It is by no means obvious that simple mea-

sures, such as the age or size of a building, had an effect on housing quality. Di
Mao (1974), for instance, discusses the decline in quality in the larger build-
ings—contradicting the observations suggested by Prell (1989).

Reconsidering the Distribution of Income

Table 4 shows the cushioning effect that the imputed amount of housing in-

come had on the distribution of income in 1992. In contrast to Yates’ (1994)
findings for Australia, in which inclusion of imputed housing income changed
the Gini coefficient by one percentage point (from 37 to 38), the change for
Russia was almost 6 percentage points. One reason for this was that Russian
households did not pay for maintenance, depreciation, and utilities. The Gini
for combined housing and nonhousing income is about 80 percent of that for
nonhousing-related income. In other words, when these in-kind benefits are in-
cluded in measures of income, there is a substantial reduction in inequality.
Finally, our measure of imputed housing income accounts for about 60 per-
cent of total income for the poorest 20 percent of the population, suggesting that
these households were relatively “house rich.” These households almost cer-
tainly would have preferred to receive their income in a less specific form.

Housing allocation also had significant effects on horizontal equity, that is,
among persons with similar income but different in other respects (Bergson 1984).
For example, we found that gender inequalities in the distribution of individual
earnings were significantly offset by imputed housing subsidies because female-
headed households have more housing space per capita. The inclusion of im-
puted housing subsidies raises the median female-to-male ratio by 7 percentage
points—from 0.54 to 0.61. This effect was even more pronounced for the bot-
tom 10 percent of the population, for which the female-to-male ratio went up
by 11 percentage points (see Buckley and Gurenko 1995).

Such changes in the distributional picture are likely to have affected elements
of the transition process. One possibility is that the relatively high share of hous-
ing income as a proportion of the total income of lower-income families shielded
these households in their adjustment to the postcommunist depression. Milanovic

\begin{table}
\centering
\begin{tabular}{lll}
\hline
\textbf{Income category} & \textbf{Gini coefficient} & \textbf{Coefficient of variation} \\
\hline
Income per capita (including housing) & 35.4 & 83 \\
Income per capita (excluding housing) & 41.7 & 99 \\
\hline
\end{tabular}
\caption{Inequalities in the Distribution of Combined Household Income and Income per Capita, 1985–92}
\end{table}

(1996), who discusses the differences between labor market adjustments during the Great Depression in the United States and other market economies and that of the postcommunist depression, finds that in the former case the adjustment took place through job cuts, while in the latter it was through wage adjustments. In market economies, wages were broadly stable in real terms while unemployment multiplied. In Russia the opposite occurred: real wages declined between 40 and 60 percent, and unemployment remained relatively low.

How much of this difference in labor market adjustment is attributable to the fact that much of the income of lower-wage workers was unaffected by wage cuts? Did enterprises in Russia, which had traditionally provided most of the social safety net, respond in the way they did because employees’ household income would not be affected by such cuts? Conversely, does the illiquidity of the imputed housing income affect the adjustment of labor markets? One recent study of Poland, for example, estimates that as much as 25 percent of its unemployment rate in 1992 occurred because workers could not find housing near the available jobs (Coricelli, Hagemeyer, and Rybinski 1995). Although we do not know the answers to these questions, the adjustment is a part of the economic landscape that should not be ignored.

As shown in table 4, the distribution of combined income (including housing) was substantially more compressed than the distribution of income derived from wage data or total earnings in the techniques used by Atkinson and Micklewright (1992). Moreover, one has the sense that if such information could have been spliced onto the 1989 comparisons, the Soviet Union would not have looked so nearly like the Western economies but instead more like the socialist regime that it was.

Conclusion

The distribution of housing in Soviet Russia reduced income inequality and provided a strong cushion against the consequences of the transition. The effects were similar to those described by Smeeding and others (1993) for the OECD economies, but that study found that housing subsidies represented the smallest share of in-kind compensation. Moreover, in their findings the total share of all nonwage compensation, 21 percent, was only slightly more than our measure of imputed income from housing. In Russia, in contrast, it can be inferred from Alm and Sjoquist (1995) that housing accounted for the largest share of in-kind compensation. They found that expenditures on housing-related maintenance alone, which are just a small share of in-kind housing compensation, accounted for nearly half the financial obligations for social services that firms were now transferring to local governments. Housing is certainly not an income source that can be omitted from consideration (as does Kakwani 1995) on the grounds that the rents charged by the state are low.
An important manifestation of the cushioning provided by housing can also be seen by considering housing’s role in national wealth, particularly in light of the greater dispersion of the distribution of wealth relative to that of income. In France and the United States, for example, housing accounts for one-half and one-third, respectively, of household wealth. In both countries, wealth is far more dispersed than income (Kessler and Wolff 1991). In a society like Russia’s, in which inflation has eliminated most savings of the household sector, imputed income from tangible wealth such as housing is likely to be an even larger component of wealth than it is in market economies. And because of high rates of inflation, nonresidential wealth is likely to have been more widely redistributed than in market economies. As a result, the almost “give-away” privatization of housing that began in Russia in 1993 gave households at least some savings to help cushion the costs of the transition. This program was an important way of permitting households to exploit the distributional benefits of the old system, a step that has not been pursued nearly as aggressively in many other reforming economies. Because the distribution of housing income in Russia is so much less dispersed than is wage income, while at the same time the distribution of nonresidential wealth is almost certainly much more dispersed, housing’s privatization had a strongly progressive effect on the distribution of wealth.

The privatization of housing may also help address poverty concerns. For example, in 1992 the elderly in the lowest-income quartile had less than one-seventh of the income of young households in the highest-income quartile. But they had five square meters more housing space per person. Because up to 30 percent of the elderly were below the poverty line, being able to liquidate this additional housing wealth could go a long way toward addressing their poverty problem.

Devising means by which the poor and the elderly can gain access to this illiquid wealth is by no means a simple task. Complicated questions of property rights, registration, and enforcement as well as such issues as defining condominium rights and responsibilities are fundamental management issues that be-devil simple resolution. Nevertheless, at the very least, this source of income should not be ignored, as it has been in recent analyses (see, for example, Barr 1993). We recognize that our data could be significantly improved, particularly as market transactions permit better calculations, but given the scale of the effects involved, the evidence supports a 1977 United Nations recommendation that housing income be included in analyses of income distribution.

The housing policies that caught Dr. Zhivago’s attention more than seventy years ago continue to affect the basic fabric of life in Russia. Decisions with respect to marriage, children, profession, and job opportunities are no doubt affected by housing considerations. The impact of housing policies on so many aspects of economic behavior makes social compensation systems more difficult to reform. Thus, at a minimum, the former socialist economies should place a greater emphasis on integrating housing decisions into broader measures of economic activity.
Notes

Robert M. Buckley is with the Operations Evaluation Department of the World Bank, and Eugene N. Gurenko is with Planning and Development Collaborative (PADCO), Washington, D.C. This article was initiated as a component of a World Bank research program. The authors would like to thank Michael Alexeev, Patricia Annez, Lev Freinkman, Jozsef Hegedus, Stephen Mayo, and Blair Ruble for their valuable assistance.

1. Pasternak's interest in housing has been a constant theme of Russian literature. See, for example, Brodsky's "In a Room and a Half," in Less than One: Selected Essays (1985) and the chapter called "Housing the People," in Khrushchev Remembers: The Last Testament (1980).

2. Szelenyi was exiled from Hungary when his studies, conducted in the late 1960s and early 1970s, showed that the Hungarian mechanism for allocating housing resulted in a regressive distribution of benefits. (His work was published in English in 1983.)


4. A detailed description of the data set is available upon request.

References

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


AN EXCHANGE ON
PROJECT EVALUATION

Because of the importance of investment to development, the economic appra-isal of projects has long attracted the attention of development theorists and practitioners alike. The subject has recently been enjoying a revival of interest, as several studies revealed a decline in the quality of economic analysis used, both in the World Bank and elsewhere. The following three articles reflect some of the views expressed in the recent resurgence of interest in project analysis. We hope that these articles will stimulate further contributions, as the development community grapples with this important and evolving subject.
BEYOND RATE OF RETURN: REORIENTING PROJECT APPRAISAL

Shantayanan Devarajan
Lyn Squire
Sethaput Suthiwart-Narueput

Traditional approaches to project appraisal fail in practice to address two fundamental questions: whether a project belongs in the public or the private sector; and what effect any external assistance associated with the project has on the country’s development. The first issue is of general interest to both national policymakers and international donors. If the government provides a good or service that would otherwise have been provided by the private sector, the net contribution of the public project could be low. The second issue is of particular concern to donors. If financial resources are fungible, the project being appraised might well have been undertaken without external financing. In this case, donor funds are actually financing some other, unappraised project. Both cases argue for a shift in the emphasis of project evaluation away from a concern with precise rate-of-return calculations and toward broader sectoral analyses and public expenditure reviews. In this context, three areas critical for proper project appraisal include a consideration of the rationale for public intervention, the fiscal impact of the project, and the fungibility of external assistance.

A quarter century ago, several groups of economists (Little and Mirrlees 1969, 1974; Dasgupta, Marglin, and Sen 1972; Harberger 1973; Squire and van der Tak 1975) developed methods for appraising investment projects financed from public revenues. Extending the principles used by the private sector in making investment decisions, these economists advocated the use of “social cost-benefit analysis,” a test that weighs a public project’s costs and benefits in terms of its contribution to national (social) welfare. If the project’s social benefits exceeded its social costs, the recommendation was that the government should undertake the project.

Although there are important differences in the various methods for evaluating the costs and benefits, they have at least three elements in common. First,
they emphasize that a project’s inputs and outputs should not necessarily be valued at current market prices because the market price may not reflect the social opportunity cost of the resource—that is, its cost in forgone benefits to society. Say, for example, that a project hires an unemployed worker. Even though his wage is a cost to the project, it does not represent the social cost, which would be the supply price of labor (that is, the price at which a person would be willing to work). When there is unemployment this usually means a price (or wage) that is below the prevailing—or actual—wage. Thus, rather than use market wages as the cost of labor, for example, the evaluator is instructed to use a set of “shadow prices” that reflect the social opportunity cost of the project’s inputs and outputs.

Second, policymakers should evaluate every component of a project relative to a counterfactual—that is, what would have happened without the project? The example above also illustrates this principle: because the worker in this case was unemployed, no national output is forgone when he is drawn into the project. Finally, the methodology dictates that because the project’s benefits and costs occur at different points in time, they should be combined in some summary statistic, such as the project’s net present value or internal rate of return.

Today, a charitable assessment would have to acknowledge that the practical application of these principles has been limited. Governments and international agencies (some of which developed social cost-benefit analysis) use the techniques only sporadically. In a retrospective paper Little and Mirrlees (1991, p. 376) concluded “that the extent to which [social cost-benefit analysis is] used and [has] real influence is not great, even in the World Bank.”

Although there are many reasons for this decline in the use of social cost-benefit analysis, Little and Mirrlees (1991) suggest four: the growth of nonproject lending; internal incentives in lending institutions; new concerns such as poverty, women, environment, and “sustainability”; and the complexity of the methodology. In addition, the traditional approach may often fail to address the fundamental questions of concern to policymakers and donors today. Among these questions two of the most important are whether a project belongs in the public or private sector and what the effect has been on development of the external assistance (if any) associated with the project?

The first question is of general interest to both national policymakers and international donors. The world has changed substantially in the past twenty-five years. Governments and donor agencies are now reconsidering the role of the state. Instead of asking if the project generates a positive net social benefit, governments and agencies are asking if there is a rationale for public provision of that good or service. Whether the standard approaches to project analysis are suited to answering the latter question is not clear. For instance, a recurring example in Little and Mirrlees’ (1974) book is an industrial project in India—a project whose rationale for public intervention would be dubious today. By contrast, projects with a strong public-sector rationale, such as vector control and immunization (described in Hammer’s companion article in this volume), are
likely to be those with large positive externalities (that is, their value to the community is greater than their price). The resulting improvement in welfare, although apparent, is difficult to quantify in net present value or rate-of-return calculations.\(^1\)

The second issue is of particular concern to donors. If financial resources are fungible, it is unclear what a project's rate of return conveys about the loan's effect on development. The project being appraised might well have been undertaken without external financing. In this case, the donor's funds are actually financing some other project that would not have been carried out otherwise.

What is the role of project analysis in a world in which the public-private boundary is the relevant issue? This article shows that the principles underlying social cost-benefit analysis can incorporate this dimension but that the technique must be modified. Because the projects under consideration by policymakers and donors are public projects, an appropriate counterfactual is what the private sector would have provided in the absence of the project. What is required is a sectoral analysis to identify market failures (that is, instances where the market over- or under-provides goods or services relative to the socially desirable levels) that warrant public intervention. Reorienting project appraisals in this direction also leads policymakers to focus on the project's fiscal impact. The article addresses the problems that arise with standard net present value and rate-of-return calculations when financial resources are fungible and argues that policymakers should not rely on such techniques. A more appropriate analysis is a review of public expenditure programs, both to improve the quality of projects and to ensure that the impact of foreign aid on development is favorable.

The Rationale for—and Cost of—Public Provision

Traditional cost-benefit analysis, as presented by Little and Mirrlees and many others, addressed the following question: Will the project under consideration result in a net benefit to the economy? This is an important question. No one wants to invest in projects that impede overall development. But the answer to this question says nothing about whether the project ought to be in the public or private sector.

Twenty-five years ago that concern may not have been so great. At the time governments were expanding public investment rapidly, and much of that investment was in industrial sectors. Indeed, the first version of Little and Mirrlees' book, published in 1969, was titled *Manual of Industrial Project Analysis*. But now countries worldwide are redrawing the boundary between the public and private sectors and paying more attention to whether or not a project ought to be in the public sector.

The principles underlying the manuals on project evaluation can be applied to the changed circumstances and can help to formulate a new set of questions about the appropriate role of government. In particular, all the approaches to
cost-benefit analysis require the project evaluator to specify the \textit{counterfactual}. This principle is perfectly general; it applies even when the alternative to public provision is a private project.

To illustrate, assume that the government is contemplating an investment that produces private goods—a shoe factory, say. One possible judgment regarding the counterfactual is that, in the absence of the public-sector project, nothing would have happened. In this case, the analyst should focus on a comparison of the costs incurred by the project and the benefits it is expected to yield. Assuming that the evaluation points to a positive net present value (that is, the benefits are higher than the costs), the decision would be made to go ahead with the project. But the appropriate counterfactual might be that the private sector would have produced the shoes anyway (assuming the enterprise is profitable). In this case, the relevant magnitude is the \textit{net} contribution of the government shoe factory. The net present value of the public-sector project over and above that of the private-sector project (evaluated at shadow prices) may well be zero.

Specifying the counterfactual is not always easy. The relevant counterfactual to public provision includes the private market outcome considering taxes, subsidies, and regulations. But some principles can be invoked. First, if the project is producing a private good that is profitable at market prices, there is good reason to believe that the private sector will undertake it. In this case there is no advantage to public provision, and the point can probably be established without a serious evaluation. Second, at the other extreme is the case of pure public goods. Here there is no prospect of private provision and hence no need to worry about a private-sector counterfactual. But pure public goods—defense, for example—are relatively rare.

The third possibility is the need to redistribute income to the poor or correct for market failures, such as externalities and public goods. In such areas private markets will not yield a socially desirable outcome, and a case can be made for government intervention (even though the private-sector counterfactual will not necessarily be zero). Because many projects are likely to fall under this category, we consider it in more detail.

Consider the case of a product with a positive externality—a benefit from producing a good or service that is not fully captured in the product's price. The private sector will provide some of the good, but because producers are not paid the full social value of it, they will produce less than the socially optimal amount. In figure 1, the positive externality is shown as the marginal social benefit, which describes how much society is willing to pay for each unit of the good. That amount is greater than the marginal private benefit—or private demand, which is the price that individuals are willing to pay. The marginal social and private benefit lines are drawn as perfectly elastic for ease of exposition. An example would be secondary education, which many would claim has significant positive externalities, although in many countries private secondary schools exist alongside public ones. In figure 1 the private sector would provide up to $Q$ of the good
on its own. The socially optimal amount, however, is $Q^*$. (In the shoe factory example referred to earlier, the marginal social benefit and marginal private benefit curves coincide, so that $Q = Q^*$.)

In evaluating projects such as these, the analyst needs to establish three points. First, public provision should result in a greater supply of the good than would have occurred with just private-sector provision (that is, overall supply should exceed $Q$). It is also possible that private provision up to $Q$ is not forthcoming because, for example, of capital-market imperfections. But if that were the case, direct intervention in the capital markets would be better than addressing the problem indirectly through public provision. It is important to ascertain precisely why private supply is not forthcoming. The current level produced by the private sector could also represent that amount of output that can be profitably produced and sold given the costs imposed by government-induced distortions. In that case the appropriate comparison should be between the net benefits of the project and the costs of removing those distortions. If the state displaces private investment, it would simply be “crowding out” private providers. The amount produced by the private sector in the absence of the project (the counterfactual) would be equal to the amount produced by the project (assuming costs are the same), and thus the net present value of the public-sector project would be zero.

Second, the project should not result in total supply of the good beyond $Q^*$ (the point at which marginal costs exceed marginal social benefits). Thus, the relevant range for public provision in this sector is between $Q$ and $Q^*$.

**Figure 1. Public Provision and the Private Sector Counterfactual**

![Diagram](Image)

*Note: Marginal cost (benefit) is the cost of (benefit from) an additional unit of output.*

*Source: Authors' calculations.*
Third, although the presence of spillovers and market failure may justify public intervention, it should not be presumed that public provision is required. The relevant counterfactual to public provision could well be the private market outcome under appropriate taxes, subsidies, or regulation. In figure 1, for example, an alternative to public provision up to the optimal quantity $Q^*$ would be a subsidy equal to height $AB$, the size of the positive externality at $Q^*$.

Figure 1 is drawn as though the exact size of the externality were known so that the optimal amount the public sector should provide—$Q^*$—could be determined. In reality, the true value of the externality will rarely be known. Indeed, that is true of most justifications for public intervention, whether the rationale is a market failure or redistribution. Until better estimates of externalities are available, there is no real solution to this problem. Thus the risk of over- or under-supply cannot be eliminated.

The costs of public-sector intervention, however, can be measured much more easily. They can be approximated by calculating the additional burden the project imposes on the budget when consumers are charged the full amount they are willing to pay. In figure 1 the area under the marginal cost between $Q$ and $Q^*$ is the total additional cost, but consumers are paying $BCQQ^*$ to the government; thus the additional cost of the project is shown by the area $ABC$. Of course, this area can be identified only because the size of the externality is known. Assume that the size of the externality is unknown—that is, the marginal social benefit line in figure 1 cannot be located. In fact, that line could lie anywhere above the line depicting the marginal private benefit. If the government still decides to supply $Q^*$, then the area $ABC$ can be interpreted as the additional cost incurred in order to realize the (unknown) benefit implicit in the existence of the externality. Such cost calculations can provide a useful “reality check” on proposed interventions. Whatever the true size of the external benefits, the government must judge that at a minimum the external benefit exceeds this cost for the intervention to be worth undertaking. To illustrate, CNCA (Caisse Nationale du Credit Agricole), a development bank serving rural areas in Morocco, received subsidies that could conceivably be justified on the grounds that the bank operated in an underserved rural credit market and reached poor people. Although these benefits are hard to quantify, assessing the cost of the subsidy is one way to ask whether this subsidy is a good use of scarce public resources and to think about alternative uses. In this case, CNCA's annual subsidy amounted to about 20 percent of the recurrent budget for primary education and 160 percent of the recurrent budget for basic health care. And this in a country where social indicators are quite unsatisfactory—primary enrollment is around 70 percent, and under-five mortality is about eighty deaths per thousand live births.

It should be emphasized that while $Q$ is an unobserved counterfactual, it is routinely assessed implicitly during project appraisal. Most project appraisals include calculations of both the financial net present value (using market prices) and the economic net present value (using shadow prices). A positive financial
net present value strongly indicates that the private sector could undertake the project (because it generates profits).\(^3\) Such a positive value is thus an argument against public-sector investment in the project. (This does not mean that the project should necessarily be undertaken from a social standpoint, simply that it should not be undertaken by the public sector. A private project that generated negative externalities—pollution, for example—could be privately profitable but socially undesirable.) Note that this conclusion is exactly the opposite of current practice where a high and positive financial net present value is used to justify a public-sector project.

An estimate of what consumers are willing to pay provides an indication of how much the private sector would have provided, because a private company should be willing to provide goods as long as it is profitable to do so. An appraisal of the Leyte-Luzon Geothermal Project in the Philippines, a $1.3 billion project, illustrates the idea. Consumers were willing to pay 6.8 cents a kilowatt-hour (kWh) for electricity, based on current bulk energy rates in Luzon, compared with an estimated long-run marginal cost of 5.2 to 5.8 cents/kWh to operate the project. Because consumers were willing to pay more than the cost to produce the electricity, it would have been profitable for a private company to undertake the investment. Public provision here crowded out at least this quantity of private provision.

The same argument holds for projects that aim to redistribute income. For these projects the government charges less than consumers are willing to pay and bears a fiscal cost. Although the incidence of such redistributive transfers may be easier to quantify than externalities are, the government still needs to judge whether the value of the redistribution outweighs its fiscal cost.

As the public sector shifts the composition of projects in favor of those with a clear public-sector rationale—on either market failure or redistributive grounds—these projects will usually imply a subsidy. To put the same point differently, these projects place a burden on the budget. Another change in circumstances or, more accurately, in the appreciation of what conditions support development, is the recognition that macroeconomic stability is an essential prerequisite. Since a prudent fiscal policy is central to macroeconomic stability, it follows that projects that require government funding need to be reviewed with care. Fiscal balance requires that the government recoup the costs of these projects through some other tax instrument, which in turn will introduce distortionary costs somewhere in the economy. These costs should be included in the evaluation of the project. If they are omitted, and public costs and private benefits are treated equally, the net present value of these projects will be systematically overestimated (Squire 1989). Although policymakers do not have a precise measure of the marginal cost of funds for most countries, the net impact of the project on the government’s budget—a minimum measure of the true cost—is still worth showing. Because of the uncertainty surrounding the marginal cost of public funds, it is useful to reestimate the project’s net present value for a range of values of the marginal cost of funds (box 1).
How does the government generate funds to cover the fiscal cost of a project? Even if the funds are borrowed, ultimately the government's only source of revenue is the domestic tax base. Every tax instrument, however, imposes a cost on the economy because it creates a distortion. This additional cost should be applied to the project.

How much does it cost? Ballard, Shoven and Whalley (1985) estimate that it costs seventeen to fifty-six cents in the United States to raise a dollar of extra revenue (yielding a marginal cost of public funds of $1.17-$1.56). This suggests that public projects should produce marginal benefits of more than $1.17 per dollar of cost. In developing countries, the marginal cost of funds is likely to be even higher to the extent that these countries have access to a more limited set of tax instruments (trade taxes, for example), which are highly distorting.

To ensure that public-sector projects recover as much of their costs as possible from the private sector, an appropriate pricing strategy is important. Figure B-1 uses the geothermal project in the Philippines to show how a project's net present value changes with different pricing policies and different levels in the marginal cost of funds. If the electricity is sold at 80 percent of the market value and there is no distortion associated with increased taxation (that is, the marginal cost of funds is one), the net present value for the project is $29.7 million. But with even a small premium of twenty cents to raise each dollar, the project's net present value becomes negative. With a marginal cost of funds of 1.6 (roughly the upper range estimated for the United States), the project's net present value drops to −$93.3 million.

Figure B-1. Leyte-Luzon Geothermal Project: Project Net Present Value at Discounted Price for Electricity

Note: Assumes 80 percent of market rate for electricity charges and 10 percent for discount rate.
Source: Project documents; author calculations.
The Fungibility of Aid

Project-specific appraisals can (at best) only assess the project’s rate of return or its acceptability. This approach is problematical for two reasons, both of which are important for multilateral lending agencies and donors interested in the impact not only of aid-financed projects but also of aid itself. First, as noted earlier, financial resources are fungible at least to some extent (Feyzioglu, Swaroop, and Zhu 1995; Pack and Pack 1990, 1993). It is unlikely that the projects evaluated by the World Bank, for example, are so marginal that they would otherwise not have been carried out. For the ninety-nine projects evaluated in 1993, the World Bank (1994) found an average economic rate of return of 21 percent—a return too high to indicate marginal projects. Second, even if the project would not have been undertaken without external funding, there is no guarantee that it was the best of all the projects under consideration. Yet that is the relevant question.

One practicable approach is to require sectorwide reviews before project-specific appraisals and financing decisions are made. For example, Humphlick and Paterson (1994), who studied the infrastructure sector in Peru, calculated the economic rate of return of expanding each of the major road links in the country. The results varied widely. The report recommended funding only those projects with a rate of return above 12 percent. On this basis several road expansion projects were dropped, cutting $275 million from the government’s road construction program. Donor-financed road projects may have had high appraised rates of return, but if these projects were already included in the government’s prospective investment program, the development impact of donor financing would have allowed funds to be shifted to one of the projects with a rate of return below the 12 percent cutoff rate.

These reviews actually go beyond setting a good foundation for subsequent project appraisal. They also improve the overall quality of the sectoral investment program. Moreover, if the result is satisfactory, the specific project financed and appraised by the donor agency becomes less important. The donor could have more impact on development by associating itself with that project where its knowledge and technical expertise is likely to be of most value. Alternatively, the agency could finance a “time-slice” of a specified expenditure program, an option that is receiving greater attention (Harrold and Associates, 1996). (Time-slice lending finances a certain percentage of public expenditures for a certain number of years rather than the whole project.)

Sectoral expenditure reviews can also shed light on the hypothetical no-program (or counterfactual) state by identifying areas where the private sector already provides, or can provide, the goods and services in question. For instance, a review of public expenditures in Malaysia (World Bank 1992) noted that 62 percent of the Ministry of Health’s funds went to medical care (mostly private goods) and only 23 percent to public health, although the latter clearly had a higher marginal impact on health. Although sectoral expenditure reviews do not provide data on
the marginal cost of public funds, public expenditure reviews regularly conducted by the World Bank that cover a government's entire budget could in principle do so. At the very least, public expenditure reviews should identify those countries in which the gains from applying the marginal cost of funds in project evaluation—or the losses from not doing so—are the greatest.

Conclusion

Altogether governments in developing countries typically spend about 4 percent of gross domestic product, or a total of about $190 billion on public investments every year. Overseas development assistance runs at about $45 billion annually. Even marginal improvements in assessing investment projects could therefore have very high payoffs. At the same time, the analytical skills to undertake careful evaluations are limited and the information required is often lacking.

To some extent, the changed circumstances justify the relatively limited attention that government and donors have paid to the appraisal process (World Bank 1995). In this context, how should the scarce analytical resources that are available be allocated to improve the quality of projects and achieve the maximum impact on development?

Public Expenditure Reviews

Governments and donors should routinely review the entire public expenditure program or its sectoral components before embarking on the appraisal and financing of specific projects. Governments would benefit because such a review would improve the overall quality of the public investment program; donors would benefit because the program would provide some assurance that aid flows were being well utilized even when financial resources were fungible.

Once an expenditure review is in hand, aid programs should be designed to ensure that the recommendations of the review are implemented and that financial assistance is conditioned on a satisfactory program of public spending. The choice of lending instrument—single project, sectoral investment loan, or general budgetary support—should be made according to which vehicle contributes most effectively to the objective. Where the investment program is less than satisfactory, donors would be well advised to focus their efforts on technical assistance and to limit their financial assistance to projects that are likely to be nonfungible.

Project Evaluation

Whether in the context of a single project or a public expenditure review, project evaluation still has a vital role to play. But the traditional tool analysts use—the rate of return—will be less and less relevant given the nature of the
projects that governments and donors will be pursuing. In these new circumstances, two questions should be routinely addressed in evaluating every project, even when it is not possible to measure all the benefits, a characteristic that is likely to describe a growing number of public investment programs.

First, what is the rationale for the public sector's involvement? Since governments and donors are still financing projects that appear to be producing private goods, a greater effort should be made to establish the rationale for public provision. Policymakers should routinely assess private-sector alternatives, including improved regulatory or price (tax-cum-subsidy) policies. Even where the rationale for public ownership is clear, the objective should be to ensure that total supply is greater than would be the case with just private provision.

Second, what is the cost of public provision? If countries and donors systematically implement the preceding recommendation, the net impact on the budget will be negative. Where this fiscal cost arises from an expansion in supply beyond what would have been forthcoming from the private sector, it represents the price that society has to pay to reap the benefits underlying the rationale for public intervention. If the government is not charging the maximum amount the private sector is willing to pay, there is an additional fiscal cost—an *transfer*. Both the expansion and the transfer constitute additional burdens on the budget. To the extent that governments have to rely on (distortionary) taxation, raising the required revenue will entail real costs. These costs, as well as the marginal cost of public funds, need to be incorporated in project appraisal wherever possible.

Are our proposals feasible? Technically, yes. For example, the rationale for public-sector provision can be identified at the same time as the public expenditure review is conducted or early in the project cycle, and without amassing all the information usually associated with a full-blown cost-benefit analysis. Similarly, analysts should be able to provide reasonable estimates of the fiscal costs of projects without knowing all the benefits. Finally, both of these tasks can be undertaken for all projects.

Notes

Shantayanan Devarajan is division chief, Lyn Squire is director, and Sethaput Suthiwart-Narueput is an economist in the World Bank's Policy Research Department. The authors wish to thank Mark Baird, Pedro Belli, Jeffrey Hammer, Anandrup Ray, and several anonymous referees for their insightful comments on an earlier draft.

1. In this article rate of return refers to both rate of return and net present value calculations.
2. Note that this is true for any intervention up to $Q^*$. At $Q^*$, for example, the additional cost $ABC$ is strictly less than the size of external benefits beyond the market outcome, $ABCD$. For interventions beyond $Q^*$, this may no longer be the case.
3. The qualification that capital-market imperfections may prevent the private sector from undertaking the project may apply here, as does the response, namely, that the appropriate intervention would be to remove the capital-market imperfection.
4. Analyses of infant mortality showed that expenditures on safe water and immunization had a much higher effect than expenditures on government-employed doctors (World Bank 1992).

References

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


ECONOMIC ANALYSIS FOR HEALTH PROJECTS

Jeffrey S. Hammer

This paper applies to the health sector a method of project analysis advocated recently by Devarajan, Squire, and Suthiwart-Narueput. A health project evaluation should establish a firm justification for public involvement; establish the counterfactual—what would happen with and without the project; and determine the fiscal effect of the project and the appropriate levels of fees in conjunction with project evaluation. The evaluation should also acknowledge the fungibility of project resources and examine the incentives both for high-level public servants to shift government resources away from project-funded activities to those that have not been evaluated and for lower-level contractors and civil servants to provide good or bad service. Market failures in health services and insurance markets should serve as a starting point for economic analysis, not as a reason to ignore economics in health projects. Project outputs should be predicted after taking into account the reaction of consumers and providers in the private sector as well as market structures of supply, demand, and equilibrium for health services.

In a recent article summarized elsewhere in this volume, Devarajan, Squire, and Suthiwart-Narueput (1996) discuss factors that should be considered to improve the economic analyses of projects. They suggest that project analysis and its supporting sector work should (1) establish a firm rationale for public involvement; (2) determine the counterfactual—what would happen with and without the project, taking into account the reactions of consumers and other suppliers; (3) identify the fiscal impact of the project on public funds; and (4) consider the possibility that loaned funds are fungible and that the real effect of a loan may derive not from the project identified, but from another project, chosen by the government and made possible by the additional funds. This paper examines these recommendations with respect to project evaluation in the health sector.
The Rationale for Public-Sector Involvement in Health

The policy environment for projects today differs considerably from that of the late 1960s, when the basic ideas of contemporary project evaluation were formulated. At that time, economies in the developing world were highly distorted as a result of protectionist policies and government regulation or direct control of industry. It was assumed that governments would take a leading role in industrial projects, and the literature on project evaluation emerged as a way to help governments make socially profitable investment decisions. The question whether or not these activities should be in public hands was not an issue.

The world has changed significantly in the past thirty years. Countries have liberalized their economic policies and become more market oriented. The absolute level of distortions caused by taxes, trade barriers, and regulations has fallen, and many countries, including those that previously had centrally planned economies, have developed active and competitive private sectors. The premise that governments should carry out or rule on all investment projects is no longer accepted as given.

The techniques of project evaluation should adapt to this changed environment. Government investment, like any government intervention, should be justified in terms of the social benefit the project will have over and above that which would occur without public-sector involvement. For any investment opportunity, the focus of analysis should be on the difference between social and private benefits—not on the costs and expected returns to private goods themselves. The standard way to assess this relative benefit is to identify the market failures that characterize the private-sector equilibrium and to quantify the welfare loss from those failures. Priorities for investments should be based on the degree to which investments ameliorate welfare losses. Let me explain.

Policy formation in modern welfare economics usually begins by explaining how private markets allocate resources. In an ideal private market goods will be offered for sale to the extent that prices cover the cost to the seller of the last item sold. They will be purchased to the extent that their value to the buyer is at least as great as their price (and, hence, their cost). When prices, quantities supplied, and quantities demanded are mutually consistent, the market is said to be in equilibrium. In the ideal market the subjective value to the buyer (private) is the same as the value to society (social), and in equilibrium both are equal to the cost to society of producing the commodity, thus ensuring that resources are allocated to their most efficient uses.

Several circumstances interfere with the ability of private markets to operate with such efficiency, causing so-called market failures. These failures may occur, for example, if some goods and services are public goods, which cannot be withheld from persons who do not pay for them and which can be consumed by one person without reducing their availability to other consumers. The social value (total value) of such goods exceeds the private value of production, which is zero, because they will not be sold.
Related to public goods are goods that produce “externalities,” benefits or costs to persons not party to the transaction and not considered by the producer or consumer when deciding how much of the item to sell or buy. The classic example of a negative externality is pollution, whereby a chemical factory does not figure into its costs the damage done by environmental dumping. It will therefore produce too much of its product, reducing the social value of the goods relative to their private value.

Monopolies by one or a few firms may also affect market equilibrium by permitting those firms to raise prices above marginal costs (the unit cost of additional production) and to restrict output, thus driving social costs above social values. Faulty information about the value or cost of products also affects equilibrium.

Associated with each of these market failures is a deviation of social and private value (or social value and social costs) and an associated social welfare loss that is, in principle, quantifiable in money terms.

If a “health project” is any investment in which the improvement of people’s health is an important output, health projects comprise a very mixed assortment of activities. Some of these activities fall under the traditional jurisdiction of ministries of health; others, such as sanitation and safe water, may be directed by other ministries. They are mixed in another sense as well, because the components of health projects range from goods that are almost purely public to those that are almost entirely private, with services showing various degrees of market failure in between. Some principal areas of market failure in the health sector are noted below.

**Infectious Diseases**

Much has been made of the “epidemiological transition,” the shift of the causes of mortality from infectious diseases to the noninfectious, chronic diseases typical of rich countries. Infectious diseases are still responsible, however, for a large proportion of deaths in poor countries, in particular, for deaths within poorer groups in those countries. They make a prima facie case for government intervention for three principal reasons.

First, they have distinct externalities. People with infectious diseases may not seek medical care quickly enough to avoid spreading the disease, and they may fail to complete a full course of treatment, which may lead to a resurgence of their illness, increased transmission, and increased risk of resistance to the drugs available for treatment. Similarly, for diseases for which transmission is decreased by the number of children immunized, an immunization program will confer an external benefit in addition to the benefit gained by the child immunized.

Second, some of the policy options available to fight infectious diseases are almost pure public goods; that is, no one can be excluded from using them, and use by one consumer does not preclude use by others. Goods of this kind will not be provided by the private sector. Control of disease-carrying pests (vector control) is one example. The collection and dissemination of health-sector informa-
Table 1. Female Adult Mortality Rates by Cause of Death and Income

(percentage of females dying between the ages of 15 and 60)

<table>
<thead>
<tr>
<th>Income quartile</th>
<th>Infectious diseases</th>
<th>Noninfectious diseases</th>
<th>Injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richest</td>
<td>0.4</td>
<td>6.7</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>7.9</td>
<td>2.0</td>
</tr>
<tr>
<td>3</td>
<td>0.6</td>
<td>7.6</td>
<td>2.4</td>
</tr>
<tr>
<td>Poorest</td>
<td>1.4</td>
<td>8.9</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Source: Adapted from Murray, Yang, and Qiao (1992).

Third, infectious diseases disproportionately affect the poor. Table 1 shows the distribution of mortality by cause across different income groups of adult women in China. Although it is apparent that the poor suffer higher mortality from all causes, the rate at which they die relative to the nonpoor is very much higher for infectious than for noninfectious diseases (a factor of 3.5 as opposed to 1.3). If table 1 were extended to all age groups (in particular, to children), the relative effects of the kinds of diseases across income groups would be similar, but the incidence of infectious diseases would rise relative to noninfectious diseases. Any reallocation from infectious disease control to noninfectious disease control in China would therefore hurt the poor most.

Uncertainty and Incomplete Information

A second set of market failures relates to the problems of uncertainty and incomplete information that plague the health sector. Although intervention is often suggested to correct these problems, it should be used with a great deal of caution. No market offers complete information about the goods and services it produces. If medical services are underused because consumers have incomplete information, an information campaign might be sufficient to correct the problem. If consumers, once informed, do not change their behavior, they may be expressing opinions about the value of the goods and services, rather than indicating a market failure regarding information.

This being said, imperfect information takes a few specific forms in the health field. First, for preventive health measures with no associated marketable product (the value of washing hands after defecation, for example, or of wearing long sleeves in the evening to protect against malaria), private mechanisms for delivering information may be inadequate. News sources sometimes cover these issues, but the value of news media as purveyors of information may depend on literacy or on the existence of a free press.

Second, the natural asymmetry of information in clinical health services may lead to “supplier-induced demand.” Medical practitioners, who know more about health problems than their patients know, may overtreat for financial gain. Government policy can sometimes improve welfare in such markets, although

50
the appropriate policy depends upon the specific ways in which the market works. Although some aspects of this problem have been modeled for the United States and the United Kingdom, there are few good models for the developing world.

**Health Insurance and Efficiency Loss**

Critical to the functioning of the health sector is the problem of catastrophic loss and the interaction of insurance markets with medical care markets (Griffin 1990). Routine care is not necessarily difficult for consumers to handle out of pocket. It is the infrequent but financially devastating incidents that are of concern. Health expenditures in all countries are skewed toward a small fraction of the population that accounts for a large share of total health expenditure. Insurance markets may fail, and even cease to exist, when those who think they will need care buy insurance, and those who think they will not, do not. The resulting higher costs of coverage will drive out still others who are relatively healthy, and the entire market may unravel. This phenomenon is known as “adverse selection.”

Problems of “moral hazard” may also occur with regard to insurance. The insured may be less careful in preventing an outcome that is covered by insurance or in minimizing the cost of getting service. The service provider, in turn, may overcharge or overtreat insured clients in order to be reimbursed by the insurer. These effects of moral hazard may then lead to suboptimal insurance coverage, in that insurers may refuse to cover certain kinds of illnesses, treatments, or patients.

The absence of a well-functioning insurance market means that large numbers of people who would be willing to pay the actuarially fair rate to protect themselves from catastrophic illnesses are prevented from doing so. The welfare loss associated with the absence of insurance markets is particularly relevant to the case of expensive procedures, for which consumer demand for coverage is highest (Hammer and Berman 1995). In addition, the welfare gain per dollar spent to substitute for insurance will be highest for rare health problems, because these problems will have low expected costs against which to insure. The unpredictable nature of the demand for health care, combined with the widespread absence of insurance is a key feature leading to large discrepancies between the social and private benefits from care. The latent demand for medical insurance (as opposed to medical care), and the efficiency loss induced by its absence in developing countries is an especially neglected element in health economics. The recent Rand Corporation experiment in Indonesia is a notable exception (Gertler and Molyneaux 1995).

**The Alleviation of Poverty**

Although the health sector is frequently called upon to help alleviate poverty, public intervention for this purpose should be treated with care. The kinds of
goods that are the best vehicles for redistribution though subsidized services are
generally those that have very low (preferably negative) income elasticities; that
is, they should be goods or services that poor people consume relatively more of
than others consume. Although the health of the poor is worse than that of
other groups, the poor do not generally demand more health care than others
do. It is more often the case, in fact, that income elasticities for expenditures on
health care are very high—usually greater than 1 and often close to 1.5. In other
words relatively rich people spend a greater proportion of their income on health
care than do the poor. Subsidizing services across the board, therefore, would
transfer money to the wealthy. Van der Gaag (1995), citing examples from China,
Côte d'Ivoire, Indonesia, Peru, and Tanzania, notes that many public health
systems, although justified on the basis of ensuring equity, provide higher subsi-
dies to the relatively affluent. Solon and others (1991), for the Philippines, show
that people with high incomes receive much more of the marginal (additional)
dollar spent on public health facilities than they pay of the marginal dollar col-
lected in taxes.

As noted, however, the prevention of infectious disease usually helps the poor
more than it helps others. The wide variety of possible health services thus has
an equally wide variety of possible effects for different income groups. Before
interventions in the health sector are designed with poverty alleviation as an
objective, their effects on the ultimate beneficiaries need to be carefully calcu-
lated. Many kinds of health subsidies will have a perverse result.

Market Failure and Analysis

The fact that markets fail is not in itself a justification for intervention. It is
also not a reason to ignore economic analysis. When markets work well, the
standard prescription of laissez-faire policies is adequate. It is precisely when
markets fail, and welfare losses occur, that they should be carefully examined to
determine how they have failed and what measures will most effectively im-
prove welfare. More attention needs to be given to the behavior of consumers,
providers, and the markets for medical care and insurance.

Some of the analytical methods proposed for use in the health sector ignore
these essential features. Despite recognizing that the public sector should not
rely on the cost-effectiveness of medical procedures as an allocative criterion,¹
the World Development Report 1993 presents calculations of cost-effectiveness
that are exclusively medical and that include none of the concerns discussed
here (World Bank 1993, pp. 62, 117). Some readers have interpreted these num-
bers to mean that the higher the ratio of clinical benefits is to procedural costs,
the higher will be the priority given to using public money for medical inter-
vention.² Such calculations give no priority to infectious disease (or any other
externality), no regard to the degree to which the private sector might substitute
for the public, and no extra advantage to problems that disproportionately hurt
the poor. With regard to risk and uncertainty, priorities based on the cost-

52

The World Bank Research Observer, vol. 12, no. 1 (February 1997)
effectiveness of medical procedures can get things backwards.\textsuperscript{3} If the principal market failure in a particular context is the faulty insurance market, the items of highest priority for government intervention, from a welfare-improving point of view, should be those that are relatively expensive, holding possible health benefits constant. This preference is inversely related to the criterion of publicly funding the most cost-effective procedures (Hammer and Berman 1995).

\textit{Government Failures}

Just as markets fail, so may government bureaucracies. Just as the behavior of private agents must be examined to judge how serious market failures may be, so must the behavior of public-sector health service providers. The issue of monitoring quality and providing appropriate incentives within the public sector is discussed in more detail at the end of the paper.

\textbf{Establishing the Counterfactual}

The second goal advocated by Devarajan, Squire, and Suthiwart-Narueput is to assess what would happen with and without the project. Three aspects of the health sector make this goal particularly germane. First, health care in most developing countries is characterized by a substantial private sector functioning alongside a large public sector. Second, as a service, health care is largely a nontraded good. Third, health status, a primary output of the health sector, is difficult to value monetarily, leading to a need to account carefully and separately for health status in the net output of the project.

Taking the last first, agreement on a measure of the value of life is unlikely ever to be reached, and keeping separate account of the health effects of a project will thus be necessary. Such accounting requires knowing the actual level of consumption of services (rather than the value of consumption of tradable goods, as in a standard analysis). As a nontraded good, consumption is equal to production and, if a competing private sector exists, nontraded good production can "crowd out" (or possibly "crowd in") private production, leaving net changes in consumption as the focus of analysis.

Establishing the counterfactual therefore requires explicit modeling of demand for, and (nongovernmental) supply of, services. This underscores the value of identifying the market failure motivating the project. In this case, however, determining the behavior of the system will indicate the actual outcome of adding capacity to a market that has an active private sector, rather than simply providing a justification of the project. The standard literature on project evaluation takes nontraded goods into account by modifying the \textit{prices} at which project outputs are valued (with the price capturing the net effect of project output on total market output). In the health sector, the reluctance to use prices on outputs such as lives saved means that the net contribution should be calculated...
explicitly. The behavior of private-sector providers should also be analyzed to see if opportunities exist for improving services through regulation or subsidy that may be less expensive than direct public provision. Some elaboration of these points follows.

The Private Sector in Health

As table 2 shows, a large private sector is the rule in health care. With the likely underestimation of the use of traditional healers, the true size is larger still. This strongly suggests that the reaction of the private sector to public provision will be necessary to assess the net impact of the public intervention.

Valuing Output

The difficulty in valuing outputs that entail extensions to life is one principal reason why health projects have been exempt from formal economic evaluation. Although I shall not contribute to the long and ultimately unsatisfying literature on undertaking this valuation, some judgment is needed to make informed decisions on public interventions in health. A few points relevant to practical project evaluation are noted here.

Valuation is simply a way of aggregating disparate inputs and outputs to get a single number as a measure of a project's profitability. Most of the time prices are the appropriate weights for this measure. The most visible problem in the health sector is the weight to put on life as opposed to money. Many other kinds of outputs within the health sector may be similarly difficult to compare—improving abilities to perform daily functions, for example, or relieving pain and discomfort associated with different diseases—and for which there is no market mechanism for valuation. Beyond an individual’s appraisal of different health problems, the value to society of curing these problems for persons of different ages or functions is often debated in the literature, representing another dimension of aggregation. In addition, many aspects of the output of health systems

Table 2. Public and Private Shares of Health Expenditures
(percentage of total expenditures)

<table>
<thead>
<tr>
<th>Region</th>
<th>Public and foreign aid</th>
<th>Private</th>
</tr>
</thead>
<tbody>
<tr>
<td>Established market economies</td>
<td>61</td>
<td>39</td>
</tr>
<tr>
<td>Middle East</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Former socialist economies</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td>India</td>
<td>22</td>
<td>78</td>
</tr>
<tr>
<td>China</td>
<td>59</td>
<td>41</td>
</tr>
<tr>
<td>Other Asian economies and the Pacific Islands</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>Latin America and the Caribbean</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>33</td>
<td>47</td>
</tr>
</tbody>
</table>

are not specifically related to health. Time spent traveling (or away from work) to reach clinics, waiting time, the courtesy of service providers, and many other facets of a very personal service have been repeatedly shown to be important to consumers.

There is no correct solution to the problem. Any method of valuation chosen must be accepted as arbitrary, treated tentatively, and scrutinized seriously to ensure that policy conclusions are not sensitive to implicit assumptions. It is sometimes proposed, for example, that the present discounted value of a person's income stream be used as the value of life. This "human capital" method of valuation ignores the fact that retirees consider their own lives valuable, and there is no logical or ethical reason why such persons should be ignored in social calculations. This method can be effective, however, if an evaluation is made using lost human capital as a lower bound for the true cost of a disease and still shows a project to be worthwhile (Kim and Benton 1995).

An ideal measure of the value of different kinds of health outcomes would combine the personal preferences of patients (using their own assessment of discomfort, inconvenience, life prospects, and responsibilities) with a more objective appraisal of the medical effectiveness of care. Because this combination of knowledge—personal from the patient, technical from the provider—does not reside in any one person, it is fundamentally unobservable.

A method that most closely approximates this measure is the "quality adjusted life year" used in some industrial countries (Barnum 1995). This measure relies on extensive interviews asking respondents to trade off certain kinds of health problems against others. Even with this method, however, the number is an average and does not allow for individual variation in preferences.

Some of the other methods used have no way of incorporating preferences of patients. Measures such as "healthy life years gained" or "disability adjusted life years" make arbitrary judgments about the relative weights of different kinds of afflictions and the relative social weights of years of life lost at different ages. Anand and Hanson (1995) challenge the underlying logic and ethical judgments implicit in this measure.

Even if a defensible life-value measure were obtained for a specific situation, generalizing it to contexts other than those captured in the measurement exercise would be difficult. A particular consideration is the degree of choice involved in the exposure to risks of death. One way of empirically estimating the value of life, for example, is to estimate wage differentials between safe and risky, but otherwise comparable, professions. Although the results of such studies are frequently interesting, they must be interpreted carefully. The subjects in the samples for the empirical work take risky jobs voluntarily. They may, therefore, be relative risk takers and not representative of the general public. Even if they differ little from others, however, there still remains the (ethical) concern that voluntarily taking risks is not the same as being exposed to risks without consent. Thus, deaths caused by motorcycle racing might be viewed differently from deaths from diseases caused by pollution (Viscusi 1992).
Although it has no solution, this problem cannot be avoided. It has sometimes been suggested that the difficulty in valuation can be circumvented by methods for which no value of life is needed. One such proposal is to use cost-effectiveness, which calculates the ratio of a given health impact from a medical intervention to its cost. Interventions with lower costs per health impact are then said to be preferred, and no explicit value of life is required. This strategy proves illusory in many of its proposed applications, however, in particular, in the choice between alternative treatment options in a clinical setting. It uses rate-of-return calculation to evaluate mutually exclusive options, a practice ruled out by the standard project evaluation literature (Hammer 1993a). An example of the way in which this method may yield unacceptable results is found in a paper by Sudre and others (1992), comparing different treatment options for malaria. Table 3 presents alternative program costs and the expected savings in lives for two different drugs. It is assumed that use of the drugs is mutually exclusive.

The authors note that if cost-effectiveness is used as the criterion for deciding between the two drugs, chloroquine will be selected. If the larger number of lives saved by pyrimethamine-sulfadoxine is considered, however, the implicit value of a life will make the two equivalent ([value of life] x 1,382 – 1,812 = [value of life] x 1,723 – 2,622 or [value of life] = $2.38). They therefore conclude (p. 152) that “chloroquine would be the drug of choice only if the value of a death prevented were less than US$2.38 (but greater than US$1.31).” Not only is there an implicit value to life in the (supposedly value-free) use of cost-effectiveness ratios, but it is absurdly precise and ridiculously low.

Because the problem cannot be avoided, the best advice is to be modest and to examine the logical consequences of alternative valuations. Health effects should be presented separately from other outputs (at whatever level of aggregation satisfies the policy analyst) to allow alternative estimates for the same value.

One way around the valuation of life issue is provided by the National Schistosomiasis Control project for Egypt (World Bank 1992a). For this project the rate of return was calculated under the assumption that the “switching value” that would make the project fail to pass a 10 percent rate-of-return test can be calculated and shown to be unreasonably low (Table 4). This method will not always give clear answers, however. Sometimes the value of life so obtained will be within a reasonable range for such a number. At the least, though, this calculation could give the policymaker something to discuss.

Table 3. Costs and Effects of Alternative Treatments for Malaria

<table>
<thead>
<tr>
<th>Variable</th>
<th>Chloroquine</th>
<th>Pyrimethamine- sulfadoxine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lives saved</td>
<td>1,382</td>
<td>1,723</td>
</tr>
<tr>
<td>Program cost (U.S. dollars)</td>
<td>1,812</td>
<td>2,622</td>
</tr>
<tr>
<td>Cost per life saved (U.S. dollars)</td>
<td>1.31</td>
<td>1.52</td>
</tr>
</tbody>
</table>

Source: Sudre and others (1992).
Table 4. Rate of Return to a Schistosomiasis Control Program: Sensitivity Analysis

<table>
<thead>
<tr>
<th>Implicit value of a year</th>
<th>Deaths averted each year</th>
</tr>
</thead>
<tbody>
<tr>
<td>of life (U.S. dollars)</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Rate of return</td>
</tr>
<tr>
<td>4,600</td>
<td>2,300</td>
</tr>
<tr>
<td>800</td>
<td>40</td>
</tr>
<tr>
<td>600</td>
<td>28</td>
</tr>
<tr>
<td>400</td>
<td>18</td>
</tr>
<tr>
<td>n.a. = Not available</td>
<td></td>
</tr>
</tbody>
</table>

*Source: World Bank (1992a).*

Determining Private Sector Behavior

As argued above, because medical care is a nontraded service, public production or provision (or financing) may displace services in the private sector. Any estimate of improved health status attributable to public expenditure should consequently account for the displacement of private services. The size of the effect is an empirical matter and should constitute a substantial part of the sector work leading up to the project. It can be derived from the overall market structure, which should have been a central focus of that analysis.

Substantial research has been conducted recently on the determinants of demand for health care in developing countries and the substitutability of public and private providers. Although the effort has been largely directed at determining the effect of public-sector pricing on the use of health services (Akin and others 1985, 1986; Gertler and van der Gaag 1990), a growing number of studies have examined other aspects of demand likely to be affected by projects in health. A recent review by Alderman and Lavy (1996) examined the impact of location and quality of public health facilities on use. Table 5 reproduces some of the results.

The policy changes listed in table 5 are likely to be standard project components. Both can be counted on to decrease the numbers of persons who self-treat (do not visit a modern provider). When percent changes are weighted by the share of visits to each type of provider, however, they show that in Ghana 38 percent of new visits to public facilities in response to improved quality are attributable to reductions in visits to private facilities, as are 36 percent of visits in response to better access to public facilities. In Kenya fully 80 percent of the increased use of public facilities attributable to improved drug availability is accounted for by the drop in private facility use. If an evaluation were accurately to predict the increase in public facility use in response to the project but failed to account for the decrease in private sector use, the benefits (as some multiple of persons cured, say) would be overstated by a factor of five.

It is possible that the public sector provides better medical service than the private sector provides, and this quality differential should be examined (Ham-
Table 5. Effects of Public Facility Characteristics on Service Use (percent)

<table>
<thead>
<tr>
<th>Policy simulation</th>
<th>Self care</th>
<th>Public facilities</th>
<th>Private facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghana</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improve quality of care (infrastructure, materials, and staff)</td>
<td>-3.5</td>
<td>127.6</td>
<td>-19.5</td>
</tr>
<tr>
<td>Reduce distance to public facilities by 50 percent</td>
<td>-2.6</td>
<td>95.9</td>
<td>-14.9</td>
</tr>
<tr>
<td>Kenya</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase drug availability</td>
<td>-4.1</td>
<td>3.6</td>
<td>-4.1</td>
</tr>
<tr>
<td>Reduce distance to public facilities by 20 percent</td>
<td>-1.8</td>
<td>1.6</td>
<td>-1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Share of market</th>
<th>Ghana</th>
<th>Kenya</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>51</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>25</td>
</tr>
</tbody>
</table>


It is also possible that improved access to free public facilities (in Kenya) is good for redistributive purposes. Redistribution depends on whether the average clinic user is poorer than the average taxpayer. If public clinics are disproportionately in urban areas and taxes come from agriculture, even this redistributive benefit is unlikely to be realized. In either case, to assess improvements in either health care or equity, the analysis underpinning the project should identify the market structure, the degree of substitutability and differences in the quality of public and private care, and the relevant characteristics of the beneficiaries (consumers).

Although the demand side of the market has been analyzed in some depth, the characteristics of the supply of services have been less well explored, and market analyses combining both supply and demand are rare. Gertler and Molyneaux (1995) provide one study that incorporates information from both the demand and supply sides of the market. They estimate the impact of public facility fees on private sector fees for Indonesia and find a close connection. Net demand changes are dependent on both prices.

Alderman and Gertler (1989) estimate the effect on demand for both public and private services in Pakistan of changing the public sector price of care. Although they have no data available to estimate the private-sector supply response, they explore possible net market effects by means of a sensitivity analysis. They find that the total effect of raising fees in health centers depends on the induced price rise in the private sector, because both prices are determinants of service use. In the context of project evaluation, the same kind of information might be used to examine the effect of making extra services available through the public sector (using quantities provided rather than fees charged).
Because direct information on the supply response of private providers is rare—
the Indonesia study is unusual in that the private supply response was actually
measured—experimentation with different values in a sensitivity analysis, as in
the Pakistan study, is a possible solution. In estimating the net effect of provid-
ing a competing service publicly, a simplifying assumption is that new public
capacity enters the same market and has the same effect as new private capac-
ity. If more detailed information suggests that the new public capacity has some
other effect on the private sector, that effect should be included in sector work.
New public capacity may reduce waiting times, for example, and it is time wasted
waiting for free public services that generates the demand for private services.
Whether or not the parameters (waiting time as determined by capacity, private
sector demand as determined by waiting time) are known with certainty, esti-
mates can be used to approximate the net effect of capacity. Alternatively, new
facilities may decrease travel time, which was the source of private demand.
Estimates of time savings and increased service use could be directly used in the
project evaluation, combining information on demand as a function of distance
with data on the geographical distribution of potential beneficiaries.

As these examples make clear, corrections for the impact of substitution with
the private sector may be considerable. The degree of correction will be larger
the larger the cross-price elasticity between public and private sectors, the larger
the elasticity of supply of the private sector, and the smaller the overall elasticity
of demand for services. Because many projects are long-lived (the expansion of
clinic networks, the establishment of prevention programs), the relevant elasticity
of supply is likely to be the long-run elasticity. Although this is more difficult
to estimate accurately, it will probably be much larger than the short-run elas-
ticity. In the short run, established private-sector practitioners may not move
from their current location or change the number of hours they work. With a
longer time horizon, practitioners may decide to enter or leave a local market
depending on how much the public sector draws potential clients. Similarly,
potential professionals (university students) may choose to enter more profit-
able fields if the medical profession becomes less attractive.

For the consumers elasticity of demand for services may differ greatly. Sev-
eral studies have found that the price elasticity of demand for clinical care is
higher for poor people than for others (Gertler and van der Gaag 1990). Projects
designed to reach the poor may therefore need to adjust less for displacement
effects. Pritchett (1994) finds, with respect to contraception, however, that the
number of children a family has is closely correlated with the number it desires.
The demand for contraceptives is thus likely to be highly inelastic. That is be-
cause the cost of contraceptives is very small compared with the cost of having
children. Contraceptive products are likely to be very elastic in supply (although
methods requiring professional providers will share the supply characteristics of
other medical services); if supply were inelastic as well, the prices of contracep-
tives would fluctuate widely. With elastic supply and inelastic demand, public
programs subsidizing or providing family planning services can be expected to
have very little effect. Pritchett's empirical work confirms this expectation. Pitt, Rosenzweig, and Gibbons (1993) and Gertler and Molyneaux (1994) find similar results for Indonesia, using very different kinds of data.

Another method of determining the net outcome of projects is to estimate the effect of previous expenditures in the public sector on health outcomes. This strategy was followed in a World Bank analysis of Malaysia and repeated for several other countries (World Bank 1992b, 1995). The analysis, using a panel of regions within the countries, estimates the effect of different kinds of public expenditures (usually contrasting primary preventive services with subsidized curative, clinical care) on measures of health status, controlling for income. The results for Malaysia and the Philippines are reported in table 6.

For Malaysia robust results indicated that variations in traditional kinds of public health interventions (immunization and provision of safe water) were highly significant in explaining declines in infant mortality, but that public provision of clinical care services had no effect on health status. For the Philippines the results were less robust, but some specifications suggested the same conclusion. The best fitting specification, however, reproduced in the table, indicates something quite different. The presence of a significant interaction term between regional income and public health subsidies suggests that providing services in poor areas has an important effect on health status, whereas providing services in rich areas has no effect at all. The best explanation for this is that the (large) private sector substitutes closely with the public sector in higher-income areas, resulting in a very elastic supply of private providers. In poor areas public

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>-1.06 (0.97)</td>
<td>-0.223 (0.042)</td>
</tr>
<tr>
<td>Safe water</td>
<td>-0.147 (0.05)</td>
<td>-0.026 (0.027)</td>
</tr>
<tr>
<td>Immunization: diptheria, pertussis, tetanus (DPT)</td>
<td>-0.113 (0.04)</td>
<td>-0.018 (0.013)</td>
</tr>
<tr>
<td>Publicly employed medical personnel per capita</td>
<td>1.03 (0.79)</td>
<td></td>
</tr>
<tr>
<td>Public health expenditure</td>
<td></td>
<td>-0.404 (0.113)</td>
</tr>
<tr>
<td>Public health expenditure × income</td>
<td>0.041 (0.012)</td>
<td>0.55 0.988</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.
a. Instrumental variables estimation, DPT as endogenous.
provision substantially increases access to health care because of a significantly less-elastic private supply response.

For certain kinds of projects, such as some forms of vector control, sanitation, or some kinds of health education and promotion activities, it is impossible to charge for specific services, and there will thus be no private sector at all. For such projects, no correction for the displacement of services need be made. For services in which private sectors compete with the public sector, the private-sector counterfactual should be determined. In terms of substantive changes, this calculation is likely to raise the priority of population-based, public goods projects, which have a substantial overlap with traditional public health interventions, and to lower the priority of clinic-based, patient-initiated services, for which the private sector can, and almost always does, exist.

This will not always be true, however, as the case of the Philippines illustrates. Indeed, consideration of the effect of projects on the private sector will not necessarily reduce the value of the public investment. The World Bank analysis of the health sector in Malaysia (World Bank 1992b), for example, suggests that the existence of a reliable public health service in that country has provided competition to the private sector that has effectively held down fees. This welfare improvement from the public service could not have been estimated solely from the characteristics of the public program; it needed to be understood in the context of the industrial organization of the entire sector. The general lesson is that market structures in poor countries differ substantially, and predicted effects of health projects can go quite wrong if the preparatory sector work is lacking.

Fiscal Impact, Fees, and Projects

The third point made by Devarajan, Squire, and Suthiwart-Narueput is that public funds for investment come at a premium because of the distortionary effects of the taxes needed to collect them. Estimates in the literature indicate losses that are on the order of 30 to 50 percent in industrial countries and that are even greater for developing countries. In many poor countries, which may have underdeveloped tax systems relying heavily on export taxes on agriculture, the distributional effects of higher taxation can make this cost even higher. This situation leads to three main conclusions relevant to health sector projects. First, too many projects are accepted using conventional project evaluation methods. Second, opportunities for recovering costs in the project should be explored. Third, alternative, perhaps cheaper, methods should be examined for correcting the market failure that initially justified the project. Regulations, partial subsidies, or even taxation may be as effective as project funding in improving the market outcome.

As applied to the health sector, the first conclusion suggests that public expenditures should correct only market failures for which the welfare costs are at least as high as the damage caused by the taxes needed to finance them. Because
formal project analysis has not been applied in health and so cannot have led to
a bias in favor of public-sector delivery, ignoring the cost of public funds has
probably led to a similar, if informal, bias.

The second conclusion, that options for cost recovery should be explored,
lands squarely in the middle of a long-standing controversy in the health field
(Creese 1991; Griffin and Shaw 1995). A few points from the analysis in
Devarajan, Squire, and Suthiwart-Narueput might clarify this issue. First, many
of the services offered in health care are private goods, whether or not they are
delivered by the private sector. Fees can be charged for these services and
nonpayers excluded from them, even if charges are not currently levied. If there
is a premium on public funds (or a budget constraint for the health ministry),
the decisions about which projects to support and how much to charge should
be made jointly.

Whether and how much to charge depends on balancing two opposing con-
cerns (Hammer 1993b). Because charging fees will reduce the drain on the gov-
ernment budget for a given project (or allow a fixed budget to be stretched
further), cost recovery will translate into a higher priority for any particular
project. Raising fees will, however, reduce demand for the publicly provided
services. The question to ask is what the consequences are likely to be of failing
to get that care. How many consumers will be dissuaded by higher fees from
seeking publicly provided care, and what difference will it make to their health?

The effect on individual health depends on several factors. First, do people
stop seeking treatment altogether, or do they switch to care provided by the
private sector? Are the health conditions for which they stop seeking care likely
to be serious or not? Second, what is the relative effectiveness of treatment in
the public and private sectors? To put it starkly, if as a result of higher fees,
people are staying away from clinics for treatment of muscle aches and skin
rashes (a large component of demand for local hospital services in Indonesia) or
are buying the same over-the-counter treatments that they would get from the
public facility, that is one thing. If they are sitting home, infecting others, and
dying of tuberculosis, that is quite another.

Who benefits from fee increases, and how, also depends on whether the money
collected is retained by the clinics and used to improve the quality of service.
Jimenez (1987) shows the conditions under which the improvement in quality
can outweigh the financial burden of the fees. Litvack and Bodart (1993) dem-
onstrate just such an effect in Cameroon and argue that poor people in particu-
lar benefit in terms of overall access to services from the combined effects of
fees and improvements in quality.

Fees are least likely to harm the health status of consumers when (1) the
demand for care at public facilities is inelastic, that is, when higher fees do not
dissuade clients from using public clinics; or (2) demand at public facilities is
elastic but clients continue to use the facilities for more serious conditions and
stop using them for minor ailments; or (3) demand at public facilities is elastic
but private facilities are close substitutes, that is, the cross-price elasticity of
demand is high and private supply is also elastic, meaning that clients stop using public facilities but still receive care at private clinics; and (4) private care is effective. If the private sector is characterized by modern providers, such as nongovernmental organizations (NGOs) or public providers in their off hours (as is legal and standard practice in Indonesia), a shift in use from public to private service does no harm to health status. If the private sector consists of traditional healers with no particular skills (not true of all traditional healers), increasing demand for their services is harmful.

To judge the effect of fees, therefore, it is necessary to know a fair amount about the demand for services and the nature of the supply of private services. As mentioned above, market characteristics such as demand elasticities (sometimes with cross-price effects) are known from research but vary substantially from country to country. They cannot be confidently inferred from one country to another and should be investigated in the specific context of the sector work under way. Some market characteristics need to be examined more thoroughly, such as symptom-specific demands for services, which can help explain whether people stop using life-saving care. Note that information about the cost-effectiveness of medical treatments is not part of the essential information needed to determine which services should be provided at subsidized rates in the public sector. Cost-effectiveness shows up only with regard to the difference in effectiveness between public and private care.

Third, the high cost of public funds implies that more effort should be given to looking for policies other than subsidized provision or financing to correct market failures. Monitoring and regulating a private sector may be such an option for ensuring high standards of care at a lower cost to the government. The same might be true for private insurance. Similarly, if imperfect information is a key element in the health market, providing information concerning the quality and effectiveness of private providers can be an important public role (van der Gaag 1995).

The information needed to choose between regulating and providing services is, unfortunately, usually lacking. Similarly, the functioning of insurance markets in the health sector is not sufficiently understood to be able to prescribe confidently the appropriate regulatory framework. Recent experiments in the provision of insurance should shed light on this issue (Griffin and Shaw 1995).

Fungibility and Other Issues of Public Servant Behavior

Devarajan, Squire, and Suthiwart-Narueput note that the true effect of project funds may have little to do with the specifics of the project being evaluated. The government may have intended to start the project anyway, and the extra money simply allows it to finance another project that it may have considered marginal. Donors may be unaware of the project they are actually funding, much less be able to evaluate it. Much of the force of this argument comes from the
significantly larger scope for reallocations between, rather than within, sectors, but it also has relevance to possible reallocations within health ministries. Since the Alma Ata conference in 1977, the international public health community has stressed the need to shift resources toward basic primary care. As a result, much of the public health funding available from donors has been directed toward primary care. When governments’ actual allocations are examined, however, evaluators find that large portions of the budgets are directed toward services that do not conform to the primary care model.

Source of ‘Government Failure’

The fact that actual allocations differ so substantially from the international paradigm suggests that the ministries’ decisions are determined by internal factors, such as political pressure from providers or affluent consumers. It is also entirely possible that governments count on project financing from donors to fund the basic services (immunizations, rural care), leaving the ministries free to satisfy pressures to provide or subsidize urban, tertiary services.

A recent paper by Feyzioglu, Swaroop, and Zhu (1996) analyzes the effect of foreign aid on public expenditure patterns. Within the health sector, the authors find that although foreign aid earmarked for the health sector has reduced infant mortality in recipient countries, the governments’ own resources spent on health have not. This finding suggests that governments’ allocations compensate for the preferences of donors, and it reinforces the point made by Devarajan, Squire, and Suthiwart-Narueput that projects should be evaluated in the context of the overall sectoral strategy or by reviews of public expenditures across the board. To the extent that money is fungible within ministries, this advice seems warranted.

Incentives

Critical to effective project analysis is an understanding of the goals, incentives, and constraints of governments and their workers. Behavior is at the heart of “government failure,” fungibility, and the decision to provide, rather than regulate, services. The incentives facing senior policymakers, as well as their behavior, must be known to understand the fungibility of resources. To understand the true impact of a project, the same information must be gathered about the civil servants responsible for project implementation. To decide between providing and regulating services, the incentives to private providers—as well as to civil servants—their probable behavior, and the ability of policymakers to influence those incentives must be considered.

Although considerable intellectual effort has gone into defining the right prices by which to value outputs of projects, the inputs and outputs themselves are largely treated as given. The guidelines drawn up by UNIDO (United Nations Industrial Development Organization) were one of the original standard texts in
the field; they state that project evaluation is divided into ten steps. The first is
to “ascertain the ‘net output’ of the project and split it into adding to supply and
saving resources” (UNIDO 1972, p. 50); the next nine deal with shadow pricing
(imputed valuations) and the like. In addition, much of the discussion of the first
step concerns splitting up the outputs; almost nothing is written about “ascert-
taining the net output.”

It is naive to assume that project inputs will achieve their intended results
irrespective of the incentive structure facing those responsible for them. These
incentives may be specific to the individuals actually constructing the project
and running the enterprise (if it is kept in public hands) or may affect private
agents (such as farmers in an irrigation project) during the period when project
benefits are generated. The former category raises issues of ownership at the
higher levels of supervision and issues of civil service remuneration or contract-
ing procedures at the lower levels. The latter, which has been the subject of
numerous analyses, depends on the policy framework within which the project
operates. In a recent analysis of incentives, Pritchett (1996) finds that the dis-
crepancy between the value of capital as determined by accumulated costs of
investments and its value as determined by contribution to output (marginal
productivity) varies enormously across countries. This discrepancy is attributed
to the economic environment in which the investments have been made and
argues strongly against using simple input-output relationships that are inde-
pendent of the incentives faced by people working with invested capital.

The incentive structure is especially important for evaluations of health sec-
tor projects, because health care is a service. Its value therefore depends upon
maintaining incentives to service providers for sustained good performance. Just
because a health clinic is built does not mean the providers will show up for
work. And if they do come to work, there is no guarantee they will devote
themselves to the care of their patients. The actual output of the investment
depends on policies concerning pay and other incentives for good performance
in public employment. Just as incentives facing people in the private sector should
be examined for evidence of market failures, so should incentives in the public
sector be examined for evidence of government failures.

Recent analyses of public health systems point to serious problems in this
regard. One indication of these problems is the common bypassing of local pub-
lic health facilities for private (or higher-level public) facilities even when the
public service is free (Kloos 1990; Korte and others 1992). Among the reasons
given for this are lack of concern shown by the public provider, social distance
between the medical practitioner and his or her clients (exacerbated by assign-
ing doctors to ethnic areas different from their own), and other aspects of the
behavior and degree of commitment of the civil service doctor.

Lewis, Sulvetta, and LaForgia (1991, 1996) point to profound problems in
the technical efficiency of public hospitals and clinics. In one study in the Do-
minican Republic, the proportion of expenditures actually reaching patients as
services is estimated to be as low as 12 percent. Once again, the incentive struc-
ture, in this case for hospital administrators, is central to the problem. If the financial viability of the enterprise has no impact on pay and promotion, the quality of management is likely to suffer.

In sector work on Indonesia (World Bank 1994), a particular dilemma was identified. Regional variation in epidemiological conditions and the variety of tasks expected of public health employees argued for increased local discretion in the allocation of resources. The incentive system in place, however, which allowed doctors to maintain private practices in public facilities (in the afternoon when the public facilities were closed), raised the possibility that this discretion would be used perversely, leading to heavier reliance on clinic-based activities to the detriment of outreach and population-based public health concerns. Reforms were recommended to compensate public providers who better served public priorities.

It is widely believed that a consequence of fee-for-service private care is the tendency of doctors to overtreat so that they gain more income from more extensive service. A fair question is whether the same doctor on salary will undertreat or badly treat if there are no financial consequences for failing to satisfy customers. Both problems can, in principle, be addressed by careful monitoring and appropriate sanctions. As important as the issue is, however, little is known about the relative abilities of governments to manage public systems or regulate private systems.

Conclusion

This article has tried to show how the issues raised by Devarajan, Squire, and Suthiwart-Narueput might be applied to project evaluation in health. Although the health sector is characterized by several market failures that may justify public-sector involvement, the existence of imperfectly competitive markets should not be used to ward off economists and to justify simply any intervention. Specific market failures should be identified, and the analyses identifying them should give guidance on ways to correct for them.

The inadequacy of insurance markets is one characteristic of the health sector. This market failure should lead to an analysis of, and attempts to measure, the value of the reduced insecurity that a project might offer to beneficiaries. Also characteristic of the health sector are services that either are pure public goods (pest control) or have distinct external effects (infectious disease control). In addition, health projects are often promoted to alleviate poverty. Different kinds of services, however, have very different distributional characteristics. Basic sanitation, hygiene, and even the education of girls from poor families may have a greater impact on the health of the poor than will general subsidies to clinical services. It should be noted that none of these three areas—risk, external effects, and poverty alleviation—are handled by current applications of cost-effectiveness analysis.
With regard to establishing the counterfactual, the trick is to give a full account of what will happen with the project taking into account the reaction of other actors in the health system: the consumers, private providers (including traditional healers), NGOs, and insurers (including informal credit or private transfers). Because health is largely a nontraded good, the participation of the government will affect the overall consumption of health services. Being aware of such features of the medical service market as demand elasticities (with respect to prices as well as to project characteristics such as service location or quality) and supply elasticities of private providers (particularly with regard to price or the existence of a competitive public sector) is critical to being able to establish this reaction.

The third element, determining the fiscal impact of the project, suggests that the premium on public money should lead to a search for alternatives to subsidized provision. In health the setting of fees for clinical services is the most important area in which this issue arises. This paper argues that higher subsidies should be directed toward services that have higher social returns relative to private returns and, of those that warrant subsidy, to those that have more elastic demand. It is important to know when raising prices will be counterproductive (for example, for services with high price elasticities) and when money can be saved for other high-priority needs by charging fees (for example, for services with low price elasticities).

The fourth component of good project analysis, acknowledging the fungibility of project resources, relates to the much broader issue of understanding the behavior (motives and incentives) of public servants. The motives of high-level public servants are likely to be an issue in the health sector if donor funds earmarked for primary health care free up domestic funds for services on which those donors would place low priority. More fundamental, however, is understanding the incentives that will encourage civil servants or government contractors to provide high quality, responsible care. Virtually all projects assume that project inputs are used appropriately and that the output is known (at least in terms of numbers of patients seen at some assumed level of quality). This assumption is not valid for analyzing services, and the incentive structure of service providers should be a far more important topic for analysis and research than it has been.

Much of the analysis relevant to projects should be done before the project evaluation stage. Indeed, given the issues of fungibility and incentives, the best form of intervention by donors may not be through traditional projects at all but, rather, through general loans conditioned on overall sector strategy and reform. If a standard project is proposed, however, considerable information from supporting sector work will be needed before evaluation. In particular, if clinical or other services that require public participation are part of the project, the supply and demand for substitute services will need to be known. Just as shadow (imputed) exchange or wage rates for project evaluation should be derived from supporting economic work done before

*Jeffrey S. Hammer*
the specific project, so should the market structure of health be part of the background investigation.

Adequate market analyses are rare in developing countries. What, then, should project evaluation look like before such analyses become available? First, evaluators should be serious about collecting information. In the interim, projects may be designed to focus on the principal market failures that such research is likely to uncover. Evaluations of health activities that are pure public goods or that address genuine externalities will be little affected by more detailed market analysis and so may proceed. For projects in which poverty alleviation is central, it should be established that the project beneficiaries are, in fact, the poor. Health care as a redistributive device should be used with caution, unless by geographic placement or other means, the poor can be encouraged to take advantage of the health services and others discouraged. For other projects, with large clinical components, more scrutiny is needed. Without knowing much about the market the government will enter or the incentive structure for public providers, it is difficult to know what health effects (or even service use) will result from a project. Estimates of the elasticity of demand for services specific to the country and guesses concerning the elasticity of competitive supply should be incorporated as weights on predicted project outputs. Identifying these parameters should be the focus of future analyses of the sector, and the absence of information on these parameters should encourage caution on the part of project evaluators.

Notes

Jeffrey Hammer is principal economist with the Public Economics Division of the Policy Research Department at the World Bank. The author thanks Shanta Devarajan, Fred Golladay, Maureen Lewis, Sandy Lieberman, Phil Musgrove, Lant Pritchett, Lyn Squire, and Sethaput Suthiwart-Narueput for discussions on the issues considered in this article and for comments on an earlier draft.

1. "Using cost-effectiveness to select health interventions for public financing does not necessarily mean spending the most resources where the burden of disease is greatest. Instead, it means concentrating on the interventions that offer the greatest possible gain in health per public dollar spent. The relevant comparison is usually not with a situation in which nothing is done but with the situation created by privately financed health interventions" (World Bank 1993, p. 65).

2. Bobadilla and Saxenian (1993, p. 10), for example, precede their reference to the medical intervention numbers by stating: "That is why the first step in designing a country’s essential health package is to determine the cost-effectiveness of a health intervention—the net gain in health compared with doing nothing, divided by the cost.”

3. The contribution a particular health problem makes to the overall burden of disease has also been suggested as a criterion for setting priorities for public intervention [Murray and Lopez 1994]. This is also incorrect for contexts in which the main market failure is in insurance. The health conditions that will be most seriously underfinanced (in welfare terms) in the absence of an insurance market are the rare ones, rather than the common ones. With a few exceptions, the total impact of a problem is irrelevant to decisionmaking, the pertinent criterion being the marginal impact a policy will have in correcting the problem. The exceptions are pure research in which neither the potential cost nor the likeli-
hood of success is known and cases in which strong economies of scale can be demonstrated (Murray and Lopez 1994).

4. A shortcut for estimating the net effect of providing a competing service publicly is available from the author.

References

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


NEW FRONTIERS IN PROJECT EVALUATION?
A COMMENT ON DEVARAJAN, SQUIRE, AND SUTHIWART-NARUEPUT

Arnold C. Harberger

Readers might infer from the paper by Devarajan, Squire, and Suthiwart-Narueput and from Hammer’s applications of that methodology to health projects, that the authors are proposing fundamental modifications of the standard techniques and of received theory of cost-benefit analysis (otherwise known as applied welfare economics and in some uses as social or economic project evaluation). Such an inference is not warranted, however. So far as I can see, nearly everything the authors propose fits quite easily within the inherited corpus of applied welfare economics. The steps that they advocate are modifications not of standard cost-benefit analysis, but of habits that have developed over the years and decades both in the World Bank and quite generally among practitioners of economic project evaluation.

Hammer nicely summarizes Devarajan, Squire, and Suthiwart-Narueput’s main prescription: there should be a firm rationale for public involvement if a project is to be done in the public sector; the project should be compared with a clear “counterfactual”; the fiscal impact of the project should be clearly identified and should be assigned a properly estimated cost; and the issue of fungibility of funds should be clearly addressed in weighing the economic consequences of project loans.

Establishing the Counterfactual

Perhaps just as professors in introductory economics courses harp on the difference between “average” and “marginal” costs, revenues, productivity, and so on, professors of project evaluation try to hammer home the dichotomy of “with” versus “without” a project. Ideally, we build a running scenario of the world “without” the project, a moving picture going ahead ten or twenty years, depending on the project’s economic life. Into this scenario we then insert the project, duly noting all the relevant differences between the two scenarios. These differences are the basic inputs into project evaluation. Usually we start with the
project's own purchases and sales or other flows of direct costs and benefits. But we really should go on to assess differences that arise outside the project. This is what we are doing when we use, say, the economic opportunity cost of foreign exchange, or of capital, in project evaluation. These opportunity costs capture the fact that the use of foreign exchange in the project displaces other imports. The tariff revenue lost on these imports is captured in the economic opportunity cost of foreign exchange. If the project pays tariffs on some of its own imports, those tariffs are captured as part of the difference between the project's financial benefits and its economic benefits.

Economists as a profession are pretty good at capturing the external effects that are summarized in the opportunity costs of capital and foreign exchange because these are visualized as being generated when any new demand is inserted into the corresponding (foreign exchange or capital) market. Economists are not nearly so good at capturing types of displacement that are unique to each project. Not all evaluators even look for them; others may look in a perfunctory way and turn up some, but not all, of the relevant externalities. I know of no one, even among those who have made the most serious efforts, who would state flatly that no important externality has escaped his or her net (in a real-world project evaluation).

Economists know what they should do in principle, and they recognize this when they give lectures and ask or answer examination questions. But when they are out in the field, evaluating real projects with limited time and resources, they practice far less than they preach.

Issues of Public-Sector Involvement

Traditional project evaluation and traditional applied welfare economics have been color-blind about whether a project is located in the public or the private sector. All they are interested in is the profile of the project including all relevant externalities. If profiles A and B are the same, the projects get the same evaluations, independent of the sector in which they are situated.

I do not believe that Devarajan, Squire, and Suthiwart-Narueput would disagree with this statement, but they might say several other things. First, that many public-sector projects have been evaluated without properly tracking the counterfactual. If a public-sector textile project displaces a certain amount of private-sector textile capacity, then such things as the taxes that would have been paid by the private sector but are not paid with the public-sector project must be counted as costs.

Second, they may impute to the public sector a lower degree of efficiency or a slower adaptation to economic change—or both—than the private sector. The assumption here is that private firms will adopt cost-reducing technologies more quickly and will abandon nonviable lines of production more rapidly than will public sector firms. Economists used to find this hard to say, but they now un-
derstand that the public sector's sluggishness in these respects arises out of real political pressures. Recognition of this fact has been one of the factors behind the recent worldwide wave of privatizations. The current response to pressures that prevent state-owned firms from being economically efficient is to privatize those firms and thus get the government out from under.

Thus if the project is in the public sector and can be expected to supply goods to the public at a higher real cost, on average, than would a private producer, that should be enough to condemn the project to the discard pile, for one of the most fundamental precepts of project evaluation is that one should not ascribe to a given project a benefit that is greater than the alternative cost of providing that (or an equivalent) benefit by another route.

I believe that in these and similar cases, old-fashioned project evaluators would have followed a path very similar to the one recommended by Devarajan, Squire, and Suthiwart-Narueput. The nuance of differences that might exist (and I am far from sure that it does) could be that Devarajan, Squire, and Suthiwart-Narueput might really prefer the private-sector project to the identical public-sector project, which comes close to implying that they might opt for the private-sector project even if its true benefits were somewhat less than the public-sector project.

A Shadow Price for Fiscal Funds

On this point I have a confession to make. For some three decades I have ardently defended the convention, in cost-benefit analysis, of assuming that the marginal source of project funds was (government) borrowing in the capital market. I did so (and still do) because this is where marginal increments of money typically come from (as governments increase debt to cover deficits, both foreseen and unexpected) and go to (when either foreseen or unexpected surpluses appear). This convention is useful because the capital market (the banking system in some developing countries) is the actual source to which governments turn. It is also useful in that the ultimate sources of funds are determined by elasticities of supply and demand in the capital market. These are likely to be far less capricious than the weights that would apply to successive tax changes, which tend to be very different in their impact from one tax law (eliminating investment credits) to another (an alcoholic beverages tax or an increase in a value added tax) to yet another (a reduction in income tax rates and a partial integration with the corporate income tax).

I thought I had made a very strong case for a capital market "convention." And I felt little pressure to make a stronger one, because nearly every framework for project evaluation uses the capital market convention anyway. So there was no vocal group out there clamoring to be convinced.

My complacency was interrupted, about a year ago, when I was writing a sort of current status report on economic project evaluation. One of the topics I
was asked specifically to address was that of shadow prices for fiscal funds. My first reaction to the idea of such a shadow price was negative, because I thought that it meant abandoning the widely accepted convention of sourcing in the capital market. But further reflection revealed to me that it meant no such thing. Indeed, it was not only fully compatible with that convention, but in fact was arguably a natural complement to it.

First, consider a public-sector project that just pays for itself in budgetary funds, discounted at the economic opportunity cost of capital. For such a project one can either use a shadow price for fiscal funds or not. If fiscal costs equal fiscal benefits, then fiscal costs multiplied by \((1 + \lambda)\) equal fiscal benefits multiplied by \((1 + \lambda)\). \([1 + \lambda]\) is the shadow price of a dollar of fiscal funds.\]

Next consider a project that does not generate enough fiscal revenue to cover its costs. One option would be to assume, following the convention, that the funds (say, $10 million) initially came from the capital market and that project revenues were sufficient to pay back, say, $6 million (in present value terms). What should be done about the remaining $4 million? My old assumption seemed to imply that this $4 million of present value gets incorporated into the debt that never will be repaid. Seen in this light, the assumption does not seem too wise.

The next step was to assume that at some point during or after the life of the project, the government would use fiscal revenues to pay off the project's remaining debt. Raising those revenues would carry an excess burden because all real-world taxes are distortionary. It is difficult to measure that excess burden because the pattern of tax rates that will be used to cover the project’s deficit cannot be predicted. This particular dilemma does not disappear when one moves from a “fiscal finance convention” (rejected by me and by most) to a “shadow price of fiscal funds” (proposed by Devarajan, Squire, and Suthiwart-Narueput and accepted at least by me). My suggestion at this point is that one should work with the lower end of the plausible values of \(\lambda\) (the excess cost per dollar of fiscal funds). That would mean leaning toward the 0.17 end of Devarajan, Squire, and Suthiwart-Narueput’s reported range of 0.17 to 0.56 (for the United States).

In any case, if one assumes that the project’s fiscal deficit is covered by fiscal funds raised by taxes, one might think of applying the factor \((1 + \lambda)\) just to the fiscal deficit. But once again, simple accounting identities come to the rescue of economists. One way to apply this factor to the fiscal deficit of the project is to apply it to all the fiscal outflows and all the fiscal inflows, that is, to treat it as a shadow price of fiscal funds generally. I take this to mean all the cash outlays on the project that are done with government money (own or borrowed) together with all the inflows from the project into government hands.

My final conclusion is that I am not at all troubled, indeed I am pleased, with the idea of using a shadow price of, say, 1.20 or 1.25 for all fiscal flows on a project. I believe this order of magnitude would be on the low side (as I think it should be) not only for the United States, but for nearly every other country.
On the Fungibility of Funds

Standard project evaluation does not deal with the fungibility of funds because it focuses on the project rather than on the means of financing. Put somewhat differently, the capital-market-sourcing convention says that the money spent on the project could always simply be dumped back into the capital market, where, by definition, its true economic return (counting all externalities) would be the opportunity cost of capital. That is why the economic opportunity cost is used as the discount rate. (There is some literature that advocates treating "soft loans" differently from regular sources of funds, but I believe that most serious thinking on this subject comes to the conclusion that all projects should pass the test, using the opportunity cost of funds as the discount rate, no matter how much soft money happens to be put at the project's disposal.)

I believe that these observations imply an implicit recognition of the fungibility of funds, but it is from a different angle. Devarajan, Squire, and Suthiwart-Narueput look at the problem from the standpoint of the lending agency, and when push comes to shove, they seem to be as interested in evaluating the loan as in evaluating the project. This is perfectly sensible and straightforward. In fact, life becomes very simple once one recognizes that two different evaluations are involved.

If one focuses on evaluating the loan, one can bypass the project that it ostensibly finances and try to track down the project that it really ends up financing. But trackers will find the jungle very murky here. Rarely will it be possible to identify the specific projects that the loan will finance.

What is the solution? My opinion has always been that a good alternative for international lending agencies as well as foreign aid agencies is to insist that not only must the project be evaluated by sensible criteria, but also that the host country use similar criteria to evaluate its other projects.

This may be close to what Devarajan, Squire, and Suthiwart-Narueput prescribe, but I think there is some element of difference. I would emphasize the importance of a country's developing—and institutionalizing—its own project evaluation capability. I have participated in a few exercises in this direction and can assure readers that it is a difficult row to hoe. It is critical to have technically sound (and morally upright) people in charge of such an effort. One does not often have that at the beginning. So the row turns out to be a long one, with the lending and donor agencies planting the seeds and nurturing the seedlings as they grow.

This sort of work is a prior step to what Devarajan, Squire, and Suthiwart-Narueput recommend for sectoral public expenditure reviews. But certainly we are all on the same side of the fence when it comes to prescribing how to deal with the problem of fungibility. My position is fully compatible with their call for periodic sectorwide reviews of the results of past projects, the likely need for new ones, and public expenditure allocations in general. A great deal of attention has been focused on limiting the size of government and enhancing its effi-
ciency, but I suspect that old-fashioned project evaluation theory, if well applied, would not only lead to a similar conclusion, but to one that was based on considerations of economic efficiency.

One way to limit the government's reach would be to impose, in evaluations of public projects and programs, a penalty on the use of funds by the government. This is where we come out in the end, but the penalty of $\lambda$ is motivated in both our cases by standard efficiency considerations—the excess burden associated with all real-world forms of taxation—rather than stemming from a bias against the public sector as such.

The well-known problem of the fungibility of funds remains just as difficult as ever. In the end, evaluators must despair of knowing where the money really went when loan funds are used to finance a project that would have been built anyway. Devarajan, Squire, and Suthiwart-Narueput and I respond to this problem by urging greater vigilance in all projects and programs. If all of them pass the cost-benefit test, fungibility entails no serious cost. Because the actual level of vigilance is so bad in nearly every country, there are many paths that would lead to its improvement. The paths urged by Devarajan, Squire, and Suthiwart-Narueput (public expenditures reviews) and by me (institutionalizing the widespread employment of cost-benefit analysis) are both sensible directions.

**Observations**

The four issues examined here go to the core of applied welfare economics. As one delves into them, one discovers conceptual subtleties that often go unrecognized. One also realizes how rudimentary are many of the practices that have become habitual among project evaluators.

It all goes back to the mantra of project evaluation courses—comparing the situation "with" and "without" the project. If only evaluators could do this really well, all or even most of the time! Part of the concern over whether a project should be undertaken by the public sector may stem from the ethos of our time—a time in which budgetary problems have been a prime concern, in which issues of efficiency in government have come to the fore, and in which privatizations and other forms of downsizing have helped make at least some governments significantly smaller than they were in the recent past. I believe that evaluators as well as donors can be sensitive to these considerations without appealing to any antigovernment bias simply by specifying the relevant alternatives much more carefully than they have in the past, by using in their discounting a properly calculated economic opportunity cost of capital, and by greatly improving the quality of the work they do, using tools of applied welfare economics that have been available all along. In short, the fundamental conceptual framework need not be modified, but there is plenty of room to improve what is actually done within that framework. Devarajan, Squire, and Suthiwart-
Narueput have performed a useful service by pointing out several important directions in which such improvement is needed.

Note

Arnold C. Harberger is professor of economics at the University of California, Los Angeles, and Gustavus F. and Ann M. Swift Distinguished Service Professor Emeritus at the University of Chicago.

1. For at least a couple of decades, it was common to speak of social project evaluation and of social opportunity costs of labor, foreign exchange, and capital. In the 1970s or early 1980s, economists began referring to economic project evaluation and to the economic opportunity costs of productive inputs and outputs. This change in terminology involved no change in meaning that I can detect. I believe it was motivated by a desire to avoid people's misinterpreting the concepts via an association of the term "social" with things like "social policy," "social services," "social problems," and so on. That linkage was never really present, and the use of the adjective "economic" better reflects this fact.
POINT-COUNTERPOINT

An occasional series of articles and comment, presenting differing views.
EDUCATION VOUCHERS IN PRINCIPLE AND PRACTICE: A SURVEY

Edwin G. West

An education voucher system exists when governments make payments to families that enable their children to enter public or private schools of their choice. The tax-funded payments can be made directly to parents or indirectly to the selected schools; their purpose is to increase parental choice, to promote school competition, and to allow low-income families access to private schools. Some opponents predict that vouchers will destroy the public system, aggravate poverty, and foster segregation. Others fear that voucher-receiving independent schools will be regulated out of recognition.

The main purpose of this article is to examine the recent emergence of voucher systems as an interesting phenomenon in its own right. The evidence summarized relates to voucher systems operating in twenty countries, provinces, and states. The typical "funds-follow-the-child" voucher system, in which governments subsidize "schools of choice" in strict proportion to enrollment, appears to be the favorite form. This type of voucher has been adopted by developing countries—notably Bangladesh, Belize, Chile, Colombia, Guatemala, and Lesotho—as well as by industrial countries such as Poland, Sweden, the United Kingdom, and the United States. Much of the recorded experience with such programs is pertinent to the longstanding theoretical debates on the desirability of voucher systems.

A tax-funded education voucher in the broadest sense is a payment made by the government to a school chosen by the parent of the child being educated; the voucher finances all or most of the tuition charged. The system introduces competition among public schools and between public and private schools; and it enables schools to offer diverse educational packages to meet the different preferences of parents.

The voucher systems discussed here apply to education up to and including high school and are funded through tax revenues (for a discussion of vouchers in higher education and privately funded voucher systems, see West 1996). First,
however, it is important to understand the rationale for the basic intervention that calls upon taxpayers to finance education.

The Rationale for State Intervention

In economics the three most quoted normative reasons for state intervention in education are to protect children against negligent parents, to internalize beneficial "externalities," and to ensure equality of opportunity. Compulsory education laws are generally regarded as satisfying the first argument for state intervention. The externalities argument, to be completely persuasive, needs the support of evidence that externalities really exist and are positive at the margin—that is, that people outside the family unit are willing to pay for extra units of education beyond what parents would purchase. In the absence of formal or systematic evidence, most writers simply assume, explicitly or implicitly, that positive marginal external benefits do exist.

The third argument for intervention—the need to ensure equality of opportunity—reflects concern about the distributional implications of purely private provision. Richer parents are likely to spend more than poorer parents to educate their children, just as they spend more on cars, homes, and clothes. The view that children's life chances should not depend on the wealth of their parents or the fortuitous circumstances of the community in which they live is widely accepted. The prospect of upward mobility, of ensuring that one's children will be better off, has been a key-stone of political support for the public school system in the past.

This "equality" argument for intervention depends on the assumption that governments are best equipped to supply the appropriate institutions. But a public system that confines children to schools nearest their home or within administratively determined attendance zones can actually reduce mobility. And where the quality of public education is better in middle-class zones than elsewhere, upward mobility is obviously blocked. In other words, the public system can often narrow a child's options, forcing the child to attend an inferior school when a superior one may be physically within reach. One of the arguments for vouchers is that they enable families to break through these obstacles to give equal opportunity a genuine chance.

The Rationale for Voucher Systems

The goal of all voucher plans—to provide families with maximum choice within a decentralized and competitive system of schools—embodies four principles: consumer choice, personal advancement, the promotion of competition, and equal opportunity. Consumer choice, in education, equals parental choice: parents choose schools for their children by virtue of their parental authority and are thus, in a fundamental sense, the real consumers of education. Under a
voucher plan, government serves the consumers of education—parents—rather than the suppliers of education—schools.

The second principle, that of personal advancement, is rooted in the conviction that people want to shape their own destinies. The opportunity to choose and to decide stimulates interest, participation, enthusiasm, and dedication. Many government programs—for example, Social Security, welfare, health programs, student loans—directly subsidize the individual recipients with funding for services among which they can select. Social security recipients, for example, can spend their checks however they choose. The goal of educational vouchers is to extend this principle to education.

The third principle, the stimulation of competition applies here because public schools are usually monopolies. The objective of vouchers is to challenge them to compete—with each other and with private schools—through reducing costs, increasing quality, and introducing dynamic innovation.

The fourth principle—the goal of equality of opportunity—underlying the rationale for vouchers is a logical outcome of the other three and is expressed in the objective of increasing access to private schools. This goal is embodied particularly in those “selective,” or targeted, voucher schemes that give low-income families greater access to private schools, schemes that have been advocated by Oakland (1994) and Becker (1995). Oakland concludes that a case can be made for some redistribution in the provision of social services generally but suggests that redistribution is better accomplished by extending the welfare system to provide the poor with vouchers for selective government services such as education. This is in preference to the usual system whereby higher levels of government supply lower levels with grants that vary with the levels of local wealth and income. Although fiscal considerations are a factor in Becker’s recommendation, he advocates a targeted system primarily “because the bottom quarter or so of the population are most in need of better education” (p. 11). He quotes studies that not only demonstrate the superior performance of private over public schools in the United States, but also show that “students from disadvantaged backgrounds tend to gain the most from attending private schools.” This fact, he observes, is not surprising “in light of the more extensive choices available to middle class and rich students” (p. 12).

Studies comparing the performance of public with private schools in developing countries generally appear to match those in the United States. Analysis, for instance, by Lockheed and Jimenez (1994) of private and public secondary schools in five developing countries revealed that private schools have a significant advantage both in student achievement and in unit costs.

Different Applications of the Voucher Principle

Under most tax-funded voucher systems, education is compulsory up to a legal school-leaving age, but parents are free to choose among alternative sup-
pliers of the compulsory service. Compared with an education tax rebate, vouchers help even those who pay little in direct taxation.

With vouchers children are not assigned to schools by attendance zones or any other criterion of the school system. Instead, vouchers enable parents to select a school for their children among any eligible and participating schools, public or private. In the most common application of the voucher principle, known as “funds follow the child,” government funding is directed straight to the school chosen by the parent. Because it has no other direct government subsidy, each school is thus in competition with every other school for students. Good schools attract many students, redeem many vouchers, and prosper. Inferior schools, avoided by parents, are stimulated to improve or must close down.

In practice, tax-funded voucher systems operate under many different regulatory rubrics. They may include government inspection of schools receiving the vouchers. They may also operate only under the condition that the teachers are licensed by the government. Vouchers may be available to all families or to low-income families. The value of the vouchers can also be made to vary inversely with income, so that poorer families receive vouchers worth more than those received by richer families. A variant of the funds-follow-the-child arrangement is a system of chits, given to each parent, cashable only by appropriately designated schools, who then return their vouchers to the relevant government authority and receive the cash value, which they use to pay expenses such as staff salaries. The value of the chit could be equal to, or somewhat less than, per student government expenditure in public schools. Finally, vouchers might provide access to private schools only, public (government) schools only, or to both public and private schools.

Selective Vouchers

Selective vouchers can be restricted to families receiving less than a given income level. Such vouchers can of course be found outside the context of education. They have been used for housing, for health, and—perhaps the best example for these purposes—for food, in the United States federal government’s food stamp program. The federal government uses an income test to determine eligibility for food stamps. Recipients use the stamps instead of cash to buy groceries. The grocery stores then return the stamps to the federal government and receive cash in return. This method is similar to the “chits” version of education vouchers described above. But whereas black market operations seriously threaten the food stamp system, the school voucher largely avoids this problem because it is quite difficult to transfer (sell) the rights to the education obtained.

Selective vouchers can be allocated on the basis of gender as well as income. In Bangladesh, for instance, vouchers are supplied exclusively to females in grades six through ten.
Open Enrollment and Charter Schools

It is sometimes contended that the objectives of vouchers can largely be achieved exclusively within the public sector. This argument involves the so-called "open enrollment system," wherein the family can choose public schools across extensive geographic areas. In practice, however, disproportionate applications to enroll in a popular school lead administrators to declare it to be full. Unpopular schools, therefore, are not faced with serious costs of undercapacity and typically continue to survive such weak competition.

Another potentially interesting scheme is the relatively new phenomenon of charter schools. These are decentralized and fairly autonomous institutions that operate under contract or charter to an authorized public body. If a charter school does not attract and keep its students, it will go out of business and its charter will be revoked.

Because government subsidizes the charter school in direct proportion to its enrollments, the voucher principle is at least partially respected because "funds follow the child"; for the principle to be fully respected, private schools would also have to be eligible to receive the grants. Nevertheless the charter school provides some alternative to the one public school in a child's administration zone to which he or she is usually assigned. In urban areas, moreover, parents may be able to choose between charter schools themselves. Further details of these institutions in the United States and Europe are provided by West (1996).

Voucher Systems in Operation

Table 1 summarizes voucher systems for primary and secondary education that have been implemented in twenty countries, states, or provinces around the world in the 1990s. Typically these voucher systems are the funds-follow-the-child kind, in which governments subsidize schools in strict proportion to enrollments.

Space does not allow extended discussion of each entry in table 1. Five countries have therefore been selected for brief comment here; as case studies they may cast some light on the arguments for and against vouchers reviewed in the next section. The countries are Chile, Colombia, Puerto Rico, the United States (Milwaukee), and the United Kingdom.

Chile

Following the introduction of subsidized ("voucherized") private education in Chile in 1980, the number of students attending private schools increased considerably. By 1988 private schools accommodated 30.4 percent of the elementary school population (compared with 14 percent in 1980) and 40.8 percent of total secondary school registration (compared with 15.9 percent in 1980).
<table>
<thead>
<tr>
<th>Country</th>
<th>Qualifying population</th>
<th>Coverage</th>
<th>Regulations and practices</th>
<th>Monetary value of voucher per student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>Females grades 6–10</td>
<td>Selected localities</td>
<td>Public or private schools, minimum attendance and progress required</td>
<td>From $12 in grade 6 to $36.25 in grade 10</td>
</tr>
<tr>
<td>Belize</td>
<td>Elementary and secondary school attendees</td>
<td>75 percent of primary,</td>
<td>Strong government partnership with the churches</td>
<td>Not available</td>
</tr>
<tr>
<td>Canada: Province of</td>
<td>Families patronizing independent schools</td>
<td>Denominational and</td>
<td>Schools receiving vouchers have to have been established for 3 years minimum</td>
<td>30 percent of public school costs per student $500 in 1978</td>
</tr>
<tr>
<td>British Columbia</td>
<td></td>
<td>secular private schools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada: Province of</td>
<td>Families patronizing independent schools</td>
<td>Mainly private secondary schools</td>
<td>Public inspection; teachers must have same qualifications as in public schools; same curriculum</td>
<td>60 percent of the costs of public schooling (80 percent for schools “in the public interest”)</td>
</tr>
<tr>
<td>Québec</td>
<td></td>
<td></td>
<td></td>
<td>Full-time equivalent capitation grants</td>
</tr>
<tr>
<td>Canada: Province of</td>
<td>Families patronizing independent schools</td>
<td>Private schools</td>
<td>Public inspection; teachers must have same qualifications as in public schools; same curriculum</td>
<td>55 percent of public school per capita cost</td>
</tr>
<tr>
<td>Manitoba</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada: Province of</td>
<td>Families patronizing independent schools</td>
<td>Private schools</td>
<td>Curriculum, teacher qualifications, enrollment</td>
<td>50 percent of public school cost</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada: Province of</td>
<td>Families patronizing independent schools</td>
<td>Private schools</td>
<td>Curriculum, teacher qualifications, language requirements</td>
<td>Average value in 1991: $143 a year</td>
</tr>
<tr>
<td>Alberta</td>
<td></td>
<td></td>
<td>Receiving schools can also charge fees</td>
<td></td>
</tr>
<tr>
<td>Chile</td>
<td>Low-income elementary and secondary school</td>
<td>More than one-third of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>attendees</td>
<td></td>
<td>total enrollments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>Low-income students</td>
<td>Operational in 216</td>
<td>Program participation renewable if student performance satisfactory</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>municipalities; vouchers usable in private schools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Eligibility</td>
<td>Required Attendance</td>
<td>Funding Details</td>
<td>Costs/Requirements</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Selected rural communities with children 7-14 years old</td>
<td>330 communities; 30,000 children</td>
<td>Minimum attendance required</td>
<td>About $65 a year</td>
</tr>
<tr>
<td>Japan</td>
<td>Children over 15 years old (not covered by compulsory education laws)</td>
<td>Public and private high schools</td>
<td>Private schools must submit financial statements to the foundation for the promotion of private schools</td>
<td>40 percent of the cost in private high schools covered by government</td>
</tr>
<tr>
<td>Lesotho</td>
<td>Elementary and secondary school attendees</td>
<td>Most schools</td>
<td>Government trains and appoints teachers; strong partnership with the churches</td>
<td>Not available</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Children subject to compulsory education</td>
<td>All municipal areas</td>
<td>State finance of schools for each religion where local demand demonstrated; secular private schools also state financed</td>
<td>Public and private schools are financed on a completely equal basis</td>
</tr>
<tr>
<td>New Zealand</td>
<td>All school-age children</td>
<td>All public schools, selected independent schools</td>
<td>Open enrollment system in a decentralized public sector; school autonomy strengthened via local parent-elected boards</td>
<td>Teacher salary grants to independent schools amounting to 20 percent in 1993 with expressed intentions to raise it eventually to 50 percent</td>
</tr>
<tr>
<td>Poland</td>
<td>Families associated with one of the 36 sponsoring organizations, including the University of Warsaw</td>
<td>Private, mainly nonsectarian schools</td>
<td>Government approval required to open independent schools; a wide variety of curricula allowed in practice</td>
<td>Per capita subsidy level at 50 percent expenditure</td>
</tr>
<tr>
<td>Puerto Rico (until 1995)</td>
<td>Families with school-age children and incomes below $18,000</td>
<td>Public and private schools</td>
<td>Use of a lottery when demand for vouchers exceeds supply</td>
<td>$1,500</td>
</tr>
<tr>
<td>Sweden</td>
<td>Children subject to compulsory education</td>
<td>All municipal areas</td>
<td>Schools must follow national curriculum; supervision by the National Assembly of Education</td>
<td>At least 85 percent of per pupil cost in municipal schools</td>
</tr>
</tbody>
</table>

(Table continues on the following page.)
<table>
<thead>
<tr>
<th>Country</th>
<th>Qualifying population</th>
<th>Coverage</th>
<th>Regulations and practices</th>
<th>Monetary value of voucher per student</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>Low-income students with above-average ability</td>
<td>“Assisted Places” in private schools only</td>
<td>Participating schools must be approved by Education Department</td>
<td>$3,500 (approx.) a year on average (1992)</td>
</tr>
<tr>
<td>United States (1)</td>
<td>Low-income students in Milwaukee, Wisconsin; maximum 1,500 students</td>
<td>Private nonsectarian schools</td>
<td>Participating schools must limit voucher students to 65 percent of the student body</td>
<td>$2,900 a year (1994)</td>
</tr>
<tr>
<td>United States (2)</td>
<td>Low-income students in Cleveland City School District (K to 3rd grade), 11,864 students in the current pilot scheme</td>
<td>Public and private schools from K-3rd grade up through 8th grade; religious schools are included</td>
<td>Vouchers in the form of checks payable to the parents of “scholarship” students</td>
<td>Based on tuition charged by private schools of choice; normally around $3,000 a student</td>
</tr>
</tbody>
</table>

The Chilean reforms were described by the government as a move toward decentralization. Public schools were transferred to the municipalities, and a new subsidy law provided for the allocation of resources on a per-pupil basis and on equal conditions to both private and municipal schools. A “student performance examination” called Programa de Rendimiento Escolar (PER) operated between 1982 and 1984. This program lasted only two years because it encountered political difficulties. The Sistema de Medición de Calidad de Educación (SIMCE) national test followed in 1988. It indicated that the quality of education was significantly higher in the subsidized private educational establishments than in the municipal schools (with the exception of one group). The reforms were followed by an increase in the average number of years of schooling among the Chilean population, including the lower-income groups.

Economic recession has brought some setbacks in recent years, notably a reduction in the real value of the voucher, but to offset this partially, the new private schools have been allowed, since 1993, to charge fees for their services. This provision enables parents voluntarily to pay additional sums to their school with the object of trying to maintain or increase educational quality. Municipal primary (elementary) schools are not allowed to charge fees.

Colombia

A voucher system was introduced in 1992 and by 1994 was operating in 216 municipalities, serving 90,807 low-income students in 1,789 schools. The vouchers, worth on average about $143, were issued to students entering the sixth grade. An early examination of the program confirmed that, as intended, the vouchers were being successfully allocated exclusively to poor families.

The voucher system was introduced primarily to respond to the shortage of places in public secondary schools in Colombia, where 40 percent of the secondary schools are privately owned. The vouchers help poor students gain access to the private schools; simultaneously, the vouchers benefit the public secondary schools by reducing overcrowding.

The Colombian experience recalls that of Vermont in the United States, where approximately 95 percent of the state’s 246 communities have no public secondary schools. The communities choose instead to pay tuition for their students to attend either private high schools or public high schools in another town. This program has been in place for more than a century “to enable small and geographically distant communities around the state to provide high school education for students without incurring the expense of building their own public schools” (Walberg and Bast 1993, p. 109).

Puerto Rico

Puerto Rico’s governor, Pedro Rosello, signed a voucher plan into law in September 1993, which was limited to families earning below a given income.
The vouchers, worth $1,500, have been portable between public schools, as well as from private to public and public to private schools; religious schools were also included.

Preliminary evidence appears to refute opponents' predictions that a voucher program would ruin the public school system. Of the 1,809 vouchers awarded in the fall of 1993, 1,181 were used by students to transfer from one public school to another, 317 to move from private to public schools, and 311 to shift from public to private schools.

Following opposition and litigation from the teachers' unions, who argued that it was unconstitutional to spend vouchers at schools affiliated with religions, the Supreme Court of Puerto Rico on November 30, 1994, ruled (5–2) that the scholarship program allowing low-income students to attend the school of their choice violated Puerto Rico's constitution. The court, however, permitted the program to continue until the end of the school year (1995). Meanwhile Governor Rosello and other supporters have promised to try to find a way to continue the program.

Milwaukee

One of the most striking examples of a successful voucher system for the poor is found in Milwaukee, Wisconsin, in the United States. Pioneered largely by Democrat representative "Polly" Williams in 1990, the plan originally permitted up to 1,000 low-income students to use state funds ($2,967 for the 1994–95 school year; the amount is adjusted annually) to attend a private, nonsectarian school of their choice.

The Milwaukee program began operation in 1990 with 300 children using vouchers at six private schools. Five years later (1995) 832 students attended one of eleven participating private schools. The Milwaukee plan has been opposed by various educational establishment groups, including the State School Board Association and the Wisconsin Congress of Parents and Teachers, Inc. This opposition has probably influenced the administrative restrictions that have accumulated recently. Thus in 1994 the state legislature set a ceiling on the program of 1.5 percent of Milwaukee's 100,000 school-aged population, or 1,500 students. The private schools participating in the program must limit voucher students to 49 percent of their student body, which limits the number of places available. Since the program's inception, the lack of space has meant that more students have been turned away than have been accepted into the program. In consequence, spaces are apportioned by lottery (McGroarty 1994).

The Milwaukee scheme, though small, warrants attention because it is the only source of hard evidence on the effects of vouchers in the United States. Comments on the program's performance have been based on the annual reports of Professor John F. Witte, the state-selected outside evaluator. His first reports led some critics to complain that the participating schools suffered excessive attrition (dropouts) and that achievement tests were biased because the
mothers of the families using vouchers had a higher average high-school completion rate than mothers of students who did not use vouchers. These complaints were later rebutted by McGroaty (1994).

Those findings of Witte's evaluations that are unambiguously positive, meanwhile, combat three of the popular fears or predictions about the voucher program, discussed in greater detail in the next section. The first is the suspicion that vouchers will help individuals who are not poor and who therefore need help least. Witte's evidence shows, on the contrary, that "choice families" are among the poorest of the poor. Their average income in 1994 was $11,625—half the income level of the average family with children in Milwaukee's public schools.

The second commonly expressed fear is that vouchers will lead to segregated and antisocial schools. Evidence supplied by Witte shows instead that the Milwaukee program fosters diversity and that no participating school has been teaching cultural supremacy or separation. "The student bodies of participating [voucher] schools vary from schools that are almost all one minority race, to racially integrated schools, to schools that have used the Choice program to diversify their almost all-white student bodies" (Witte and others 1995, p. 15).

The third fear—that voucher schools will skim off the "cream" of the student "crop"—is countered by Witte's finding that "the program is offering opportunities for a private school alternative to poor families whose children were not succeeding in school. This is a positive outcome of the program" (Witte and others 1995, p. 16).

Other positive conclusions from Witte's reports include the finding of high parental involvement, once in the system, and high parental satisfaction with the program—in particular, that it increased learning and discipline. "Respondents almost unanimously agreed the program should continue" (Witte and others 1995, p. 17).

The case for vouchers rests also on the argued need to weaken the public school monopoly or, in other words, to promote competition. But when competition is introduced, those suppliers who initially lose, or expect to lose, customers will, in self-defense, act to lift the quality of their services. Applied to our education context, five years of the Milwaukee plan is more than enough time for the threatened public schools to have improved under the pressure of new voucher competition. And insofar as vouchers can take some credit for inducing the improvements in tested achievement that have in fact occurred over the years 1990–96 (in public and private schools), findings of no current difference in achievement growth between public and voucher (choice) schools do not unambiguously imply that vouchers have failed to improve efficiency.

The future of vouchers in the United States obviously will be influenced not only by official annual reports, but also by the assessments and responses of the parents. The fact that demand for voucher places in Milwaukee currently well exceeds supply could already be pressuring politicians to allow more families to participate.
In 1981 the Assisted Places Scheme was established in the United Kingdom with the aim of providing a ladder of opportunity for able but poor students. Under the scheme today, low-income parents can obtain assistance with tuition fees for an independent school if the school has been approved by the Department of Education and Science.

By 1995 about 29,800 students were using these selective vouchers at 294 specified independent schools in England (there is a separate system for Scotland). About 5,000 new pupils enter the program every year, mostly at the ages of eleven or thirteen.

The English experience raises two questions that have implications for the general debate on vouchers discussed in the next section. First, why—in view of the government's stated wish to encourage competition and "market discipline"—is the Assisted Places method so limited in coverage (see U.K. Department of Education 1992)? Second, why are the places limited mainly to able pupils who exhibit the potential for high academic achievement, when such pupils can expect a higher than average lifetime income whether or not they are in Assisted Places? The contrast with Milwaukee's selective voucher, where the low-income students designated for help have not been succeeding in school, is striking.

The voucher principle has also been extended in the United Kingdom to further education and (prospectively) to nursery schooling. Further education colleges (similar to community colleges in the United States) have recently been reestablished as autonomous institutions independent of their former local governments. A new system of "learning agreements"—effectively, individual contracts between a college and a student, specifying the precise qualifications aimed for—enables government funding to follow the student to the college of his or her choice. At the preschool level, the Department of Education declared in 1995 that it was about to extend free entitlement for all four-year-olds to good quality private, as well as public, nursery education (World Bank 1995, p. 4.1). Currently the initiative has been limited to two pilot schemes in East Anglia. The plan, however, is to be extended to all four-year-olds in April 1997 (The Economist 1996).

The Current Debate on Vouchers

As the case for parental choice and competition has gained in popularity, the criticism of those antipathetic to vouchers has increased in intensity. Debate has focused on the potential effects of vouchers on the public benefits connected with education; the possibilities for damage to the quality of public schools on the one hand or to the identity and autonomy of private schools on the other; their impact, if any, on poverty; the issue of windfall gains for the middle class; and the possible effect of a voucher system on the government's administrative costs.
Vouchers and the Market Place

Some view vouchers primarily in terms of a free market that vouchers would encourage. They then see this as a prime example of "economic man" sacrificing social welfare to his selfish pursuit of individual material gain. But economists have long abandoned narrow assumptions about self-interest. As Becker (1993, p. 385–6) observes: "Behavior is driven by a much richer set of values and preferences. [My] analysis assumes that individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, spiteful, or masochistic." A pertinent example is the objections made by the members of the Polish Civic Educational Association in the late 1980s to the national school system inherited from the collapsed Communist regime. Their position was that they wanted to maximize welfare as they as individuals saw it, as a welcome change from having welfare defined and imposed by totalitarian authorities or highly centralized bureaucracies. The type of institutions they demanded were nonstate (including religious) private or independent schools (Glenn 1995, p. 127).

A related argument by opponents of vouchers is that a free market would lead to discrimination on grounds of race or disability. Krashinsky (1986, p. 143) argues that vouchers could lead to racial segregation. The usual reply here is to quote Coleman's (1990) findings that segregation is in fact greater in public than in private schools. Shanker and Rosenberg (1992) suggest in the same vein as Krashinsky that profit-making schools would reject difficult-to-educate children under a voucher system. Lieberman (1991a) found, on the contrary, that the single largest U.S. group of for-profit schools serves the disabled. Blum (1985), meanwhile, provides evidence that urban private schools maintain a higher level of discipline than do public schools.

Another common argument against vouchers is that parents cannot be expected to make sound choices for their children (Bridge 1978, Carnegie Foundation 1992, Levin 1991, Wells and Crain 1992). Others reply that parents simply need some initial experience (hitherto denied them) with making such choices in order to become more adept. A second response is that, in a democracy, any serious impediments to decisionmaking by parents will show up also at the ballot box when they choose political representatives to make decisions on education. A third response is to quote empirical studies demonstrating rational choice for their children by parents who themselves have only modest amounts of education (Fossey 1994).

A further concern—that vouchers (or tax credits) for education might introduce fraudulent practices—is put forward by Murnane (1983), who draws an analogy with food stamps in the United States. Experience there, he observes, shows that unscrupulous parties make claims for fictitious individuals. Schmidt (1995) shows that serious shortcomings of fraud and dishonesty are already present in the public school system. Moreover, and to reiterate, the school voucher largely avoids the black market problem because it is difficult to transfer the rights to education.
Public versus Private Benefits from Education

It is generally accepted that a child's education provides not only private benefits to the family (mainly by prospectively increasing income), but also public benefits (positive externalities). The latter include poverty reduction, economic growth, and the pursuit of common values (see Krashinsky 1986). The economic model supporting this argument is that of "joint supply." One classical example of it is wool production: wool cannot be produced without simultaneously producing meat, and vice versa. Furthermore, a switch from one breed of sheep to another is likely to improve the wool production at the expense of meat, or the converse. Similarly, so the argument goes, the cost of more or improved public benefits from education shows up in fewer, or worse quality, private benefits, introducing an interesting trade-off problem. The public benefits are quite distinct from the private. Thus the inducement to an orderly society that educated citizens bring is one example of a public benefit. The increase in expected lifetime income that education bestows on students is, in contrast, an example of a purely private benefit.

Some economists object to free choice of schooling through a voucher system because they believe families will not trade off private for public benefits but will allocate their expenditures on the basis of their private benefits exclusively. In other words, the valuation that others in society place upon the education of one's child will be neglected, and public benefits will suffer relative to private benefits—the well-known "public good problem."

Proponents of that view, such as Krashinsky (1986) and Levin (1991), claim that public schools have a unique ability to produce the "common values" just mentioned. But this claim also is now contested. Cohn (1979), for instance, observes that, in practice, public schools in the United States have successfully resisted attempts to homogenize their procedures, so that "a student in one school district might receive an entirely different set of common values than his counterpart in another school district" (p. 262). Nevertheless the belief that public schools possess an absolute advantage in producing the "public good" benefits remains strongly entrenched among educationists as well as among some economists.

Private schools are direct producers of externalities (Hettich 1969), and they also generate them indirectly (West 1991). It is generally agreed that private schools are more efficient at producing private benefits, through more effective teaching of the basics, such as literacy. This is so partly because public schools are monopolies, while private schools have greater output per dollar because they experience competition. But literacy is also a public benefit, a necessary condition for communicating common values and fostering economic growth. This indirect assistance by private schooling to the production of such public benefits is at least as important as the direct production.

Krashinsky (1986) focuses on what he calls transaction costs, such as the costs of communication in obtaining the public benefits of education. His position is that these costs are too high for the government to contract out to private
suppliers because the public benefits “are so subtle” (p. 155). Even if this were the case, family consumers of private benefits from education could equally claim, bearing in mind the variety of cultural aspirations, that the education quality they seek is so inarticulable that the transaction costs of delegating the task to governments are prohibitive. In any case a central government still faces similar transaction costs in issuing instructions to thousands of school districts, which in turn face even higher costs in supervising tens of thousands of individual public schools.

The public good argument, as employed by Krashinsky, contains a serious theoretical flaw. The classical example of a public good is that of the fishermen who need a lighthouse. Even though all the fishermen in a given area would benefit from the beam of light generated, each one will conceal his true preferences and wait for others to provide it. But because each fisherman in turn will believe in the same way and try to “free ride,” the lighthouse will not be built. Because there is no mechanism parallel to the usual market system leading to the revelation of sincere (true) preferences, so it is argued, we have a case of “market failure.” In the context of education the preferences that are not revealed are those of the “neighbors” who value the education for separate reasons. Krashinsky’s assumption is that the problem will be solved by government intervention. But he assumes unjustifiably that the government possesses all knowledge of the relevant preferences of each and every neighbor. And even if government were to consult everyone individually, individual neighbors would have no more incentive to reveal their true preferences to government than they would on the conventional market. Government failure therefore matches the market failure.

Potential Damage to the Public School System

Unions of public school teachers and administrators frequently contend that a voucher system will destroy the public school system. Krashinsky (1986), for example, argues that middle- and upper-class parents would desert the public system in favor of private schools that discriminate in various ways against poor, disadvantaged, or minority applicants. The poor would be left in gutted, underfunded, and decaying public schools. But this argument rests on the questionable assumption that the public system will refuse to adjust in the face of competition from private schools (Wilkinson 1994). Holmes (1988, p. 23) maintains that “there is no reason why inner city schools of the future, where alternatives are available [with vouchers], will be worse than the ones at the moment where there is no choice.” In addition, Krashinsky’s fear that middle-income parents will desert the public school system with the aid of vouchers has no basis where they are allotted exclusively to low-income families, as they are today in such widely different countries as Bangladesh, Chile, Colombia, Puerto Rico, the United Kingdom, and the United States. By most reports, such systems are improving the condition of the poor relative to those in the rest of society.
Vouchers and Poverty Reduction

Krashinsky’s implicit assumption is that the public school system benefits the poor in a way that is superior to any alternative. But low-income families are segregated residentially, and their children are typically allocated to the schools nearest their homes. If they want to choose a better public school in a middle-class area, they must purchase a home there, and the housing prices are usually beyond their means. Middle-class families, by contrast, can move more easily because they are less restricted financially. The result is that the public provision of schooling becomes heterogeneous, with the poor, on average, receiving the worst quality. Vouchers would help remove the barriers to mobility.

Friedman and Friedman (1980) insist that they too favor reducing poverty and promoting equal opportunity but argue that in both respects the voucher system would unmistakably improve things. They contend that liberty, equality of opportunity, and the reduction of poverty are complementary and not competitive goals of the voucher system. Their main argument is that lower-income families, trapped in large city ghetto schools, would benefit most from vouchers. “Are the supermarkets available to different economic groups anything like so divergent in quality as the schools?” they ask. “Vouchers would improve the quality of the public schooling available to the rich hardly at all; to the middle class, moderately; to the low-income class, enormously” (p. 169).

Windfall Gains for the Middle Class

Some opponents of vouchers focus on what they call the inequitable windfall gains for families (usually well-to-do) that customarily purchase private education. In other words costs to governments would increase if vouchers (or tax credits) are extended to rich private school clients not now financed by government (Gemello and Osman 1983). Seldon (1986) points out, however, that total costs to government could fall depending on the value of the voucher as a proportion of per capita public school costs. The government savings would occur, according to Friedman (in Seldon 1986, p. 20), if the voucher value was 75 percent of public school costs. The reasoning is that the economies effected by migrants from public to private schools, who would now cost the government 25 percent less than before, would offset the cost of the windfall gain to accustomed users of private schools. (Clearly, because a strong argument put forward by voucher supporters is that private schools can deliver at lower costs than public, their case looks more consistent if they demand vouchers at values less than 100 percent of average per pupil costs in public schools.)

The windfall gains problem could also be handled by making vouchers subject to tax. But selective voucher systems, restricted to low-income families would be even more effective—indeed, such selective vouchers would automatically prevent high-income families now patronizing private schools from enjoying the windfall gains.
Regulatory Threats to Private School Identity

A potential drawback to vouchers has recently been suggested by strong believers in the philosophy of freedom who want to see more competition in schooling but fear that voucher systems would seriously threaten the autonomy of independent schools. Currently the most articulate and influential spokesman for this point of view in the United States is Sheldon Richman (1994). In his words: “It is likely that before schools could accept vouchers, they would be required to meet a raft of standards that before long would make the private schools virtually indistinguishable from public schools” (p. 83). Voucher initiatives that insisted on zero regulation would stand no chance of acceptance, Richman says, because, “as the opposition would inevitably point out, the voucher plan would appear to authorize appropriation of ‘public’ money to institutions not accountable to ‘public authorities’” (p. 84). In the same vein, Gary North (1993) argues: “We will have federal guidelines operating in every voucher-using school, equal opportunity policies and quota systems of every kind, teaching hiring and firing policies, racially and religiously mixed student bodies. There will be a whole army of federal bureaucrats, not to mention state bureaucrats policing every ‘private’ school” (p. 149).

Friedman has always separated three levels of issues: first, whether schooling should be compulsory; second, whether it should be financed privately or by the government; and third, how it should be organized. His position has been that whatever one’s views may be on the first two issues, a voucher scheme would produce a better and a more effective organization than the present one—that is, vouchers remain a superior alternative to a system of schools run and financed by government. Like North and Richman, Friedman sees benefits also in eventually removing compulsion and government finance, but he is primarily concerned with the question of how to get there from here. Vouchers, he believes, are still a practical transitional measure (Friedman 1993).

As for the threat of a government regulatory take-over of private schools, Henderson (1993) points out that these institutions do not have to accept vouchers with all their strings. Others argue also that the recipients of vouchers can and will lobby their government against heavy regulation. Lieberman (1991a, p. 6), meanwhile, argues that the more likely cause of increased regulation will be the political objections to funding both public and private schools while closely regulating only the former. Consequently, Lieberman observes, supporters of vouchers must argue that to approach parity, what is needed is the reduction of the regulation of public schools, not an increase in the regulation of private schools.

The Costs of Implementation

A common concern about the administrative costs of implementing a voucher system is whether the size of the bureaucracy necessary to oversee the total system will have to expand significantly. Wilkinson (1994) finds no reason to
believe that costs such as those associated with monitoring student attendance and quality of education should be any higher for private than for public schools; school quality can be overseen by periodic inspections in the same way as it is in public schools. Even in the unlikely event that administrative costs did rise, such an increase would be more than offset by the savings realized given the evidence cited above that private schooling generally costs less than public. Tax-funded vouchers in the countries described in table 1 are typically valued at considerably less than the public school per capita cost; the Milwaukee plan, for instance, supplies students with vouchers worth about half of the public school cost. It is highly improbable that additional administrative costs could equal such a huge differential. Indeed, a strong argument for governments to use vouchers is the need in these days of budgetary stringency to economize on public spending.

**Final Comments**

The main purpose of this article has been to provide information on the theory and practice of education vouchers throughout the world and to summarize briefly the principal points raised in current academic and political debates on the issue. Absence of real world evidence has until recently hampered discussion—indeed, until recently, has been adduced by several writers to demonstrate that vouchers were not desirable. But emerging evidence (see table 1) suggests otherwise, and this may well be due to changing circumstances.

During the last two decades governments have become increasingly unwilling or unable to continue to raise the share of public expenditure spent on education. The prime focus has switched accordingly to attempts to obtain higher output from given expenditure levels. The use of vouchers valued at much less than 100 percent of the cost per pupil in public schools has already been successful in Sweden, Milwaukee (United States), and Poland, and may become a popular way of economizing. Economists, meanwhile, see the key role in such efficiency gains to be the gradual removal of the current monopoly structure in education.

Some have argued (for example, Carnoy in this issue) that a central difficulty with the voucher plan is that, even if it is limited to the poor, it “diverts attention from an even greater problem: the much larger investment that societies need to make in low-income children if they hope to overcome the effects of poverty on learning.” Carnoy does not indicate, however, how this larger investment is to be financed. The voucher system could help him, because substantial resources could be made available for his program by switching (by means of vouchers) to the significantly lower cost of education supply that private schools offer.

It is too early to reach firm general conclusions about the effectiveness of vouchers. There are only twenty entries in the table, and these show a wide
variety of design. Those who fear that government regulations associated with vouchers will ultimately strangle the individuality of private schools will insist that this may yet happen. Nonetheless significant numbers of families are now obtaining positive firsthand experience with private schooling through voucher systems. This phenomenon alone could well alter the political climate in their favor.

Notes

Edwin G. West is professor of economics (emeritus) at Carleton University, Ottawa.

References

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


National Program of Self-Administration for Educational Development (PRONADE), October 1996.


The World Bank Research Observer, vol. 12, no. 1 (February 1997)


IS PRIVATIZATION THROUGH EDUCATION VOUCHERS REALLY THE ANSWER?
A COMMENT ON WEST

Martin Carnoy

Government-financed school vouchers promise to improve consumer choice while still providing the public funding needed for families to invest adequately in their children's education. But politically and in practice, the choice that vouchers offer can mean many things, from eliminating neighborhood residency as a condition for attending a public school to allowing groups of parents and teachers to form their own public schools (charter schools in the United States, for example). And it can mean an educational system that is in part publicly funded but that is privately owned and operated with no public regulation. Choice through vouchers can also produce undesirable side effects that more than offset their optimistic promises.

In this context, E. G. West presents a particular version of the case for vouchers. Although he allows for a range of possibilities, he puts primary emphasis on the privatization of education. This version of choice has to be distinguished from one that limits increased choice to public schools and from the use of vouchers to improve education for the poor. The three emphases are different, mainly in what they contend are the social benefits and costs of various degrees of choice. Whereas the case presented by West may appeal at an abstract level, I argue that the evidence supporting it is at best mixed and possibly negative.

The Foundation of West's Claims

West's argument for vouchers rests on a foundation of three claims that bear careful scrutiny. The first is that privately managed education is inherently more effective and more cost-effective than is public education. The second is that a privatized education system is more efficient than public education in improving the social mobility of youngsters from low-income families. The third is that the social costs of privatizing a public educational system are minimal.

If private education is not both more effective and more cost-effective than public education, extending vouchers to unregulated private schools might in-
crease the competition in false claims or symbols of success, but it would not support West’s claim of “reducing costs, increasing quality, and introducing dynamic innovation” (West, p. 85). Further, if the meaning of effectiveness is redefined to include the production of outputs other than measured learning and not in the public “charter,” private schools could be more effective than public schools and could increase some families’ private welfare but might not necessarily enhance the public good. For example, in the United States and many other countries, private schools can offer religious education, which is not legal in public schools. Religious training for their children is considered a benefit by some families. Yet the use of the tax funds for religious education could be contrary to the political principles of the vast majority of the national community and could increase the level of conflict over community goals.

And unless private education is more effective, its relevance to the principles of choice and personal advancement put forth by West is also less clear. West (p. 87) argues that a publicly monopolized education system restricts “good” schooling—that, in practice, even under an “open enrollment system” of public school choice, popular, more effective public schools declare themselves “full,” and poor schools continue to operate without threat of further competition. Thus extending vouchers to private schools would increase choice by expanding the number of good school places available. This implies that private schools are both more effective and more cost-effective than public schools, and that with private schools free to attract students, either they would not declare themselves full or, seeing a chance to make money, additional effective private schools would enter the market and expand the number of places available. The argument assumes the existence of (potentially) large numbers of private schools that are more effective than public schools. (Otherwise, new private schools would not emerge or would be unable to compete in the longer, or even medium, run with public education.)

It is possible that the inclusion of private schools in the voucher plan would expand the choice options, even with equally effective public and private schools. That possibility occurs, as noted above, if private schools can offer outputs that public schools cannot, such as religious education or guaranteed admission to a university or a job in a particular company.

West contends that a privatized system of education has even greater positive consequences for poor families than for wealthier families, because it would equalize the distribution of educational quality and educational outcomes among families with different incomes. He also argues that extending vouchers to private schooling would have few and only minimal negative consequences. In West’s discussion, these two claims are related, because potential downsides to privatization are closely linked to negative distributional effects, namely, harm to the public school system and windfall gains to middle- and upper-income groups. West contends that existing voucher plans have resulted in little, if any, negative impact on public education and perhaps have even had a positive result because they have increased competition. Further, he suggests that negative dis-
tributive effects could be handled by limiting vouchers to low-income families, as is the case in some countries. But this caveat is inconsistent with his more general claims for privatization. Limiting vouchers to low-income families restricts choice, which contradicts the choice principle for vouchers. In any case, the main reason to argue for a voucher plan limited to low-income families is to protect against the likelihood that a free market in education for all families would produce greater educational inequality than a public school system.

If the main issue is improving the education of low-income children, privatization through vouchers has both positive and negative aspects. On the positive side, it draws attention to the poor quality of many public schools in low-income areas and may benefit more highly motivated families. Gary Becker (1995) argues that "the best voucher system is limited to poor families" because "the bottom quarter or so of the population are most in need of better education, and the poor are most likely to benefit from competition by private schools." I agree that if a voucher system were limited to the poor, the poor could benefit, although not necessarily because access to private education would raise their test scores. Studies comparing Catholic with public education in the United States suggest that Catholic schools give students from low-income families a better self-concept, but not much higher test scores (Bryk, Lee, and Holland 1993). The objective of such a voucher system, however, would be to alleviate poverty, not to increase school choice or privatize schooling. Politically, this distinction is important. On the negative side, such a voucher plan diverts attention from the much larger investment that societies need to make in low-income children to overcome the effects of poverty on school achievement.

Is Private Education More Effective?

As West has noted, a number of voucher plans are in operation, but only two—Chile and Milwaukee, Wisconsin—have data that attempt to compare academic results for voucher pupils attending public and private schools.

Evidence from Chile

The Chilean plan, implemented by the military regime in 1980 as part of an overall free-market package, provided for fully subsidized deregulated private schools, which competed head-on for pupils with deregulated public schools in all metropolitan neighborhoods. But because few rural areas have any private schools, children in these areas (more than a fourth of Chile's municipalities) attend only public schools (Winkler and Rounds 1993). The Chilean plan revoked teachers' contracts and eliminated the teachers' union as a bargaining unit, so by 1983 public schools could hire and fire teachers without regard to tenure or union contracts. At the same time, all schools were released from the rigid structure defining the national curriculum and national standards.
What were the results? The first was that even when parents’ contributions are included, total real spending on education fell in the 1980s. In 1985 the federal contribution was 80 percent of total spending on education, and total spending was 5.3 percent of gross national product (GNP). Five years later the federal portion was 68 percent of the total, which had fallen to 3.7 percent of GNP. So municipal and private spending rose but not quickly enough to offset an 18 percent drop in real federal contributions. Most of the decline came at the secondary and university levels, where public spending per pupil dropped drastically, but by the late 1980s spending cuts had also affected primary schools. The middle- and upper-income groups made up the shortfall in two ways: first, more prosperous districts were able to spend more on public schools than were poorer areas; and second, parents in wealthier areas whose children attended private schools could supplement their vouchers with higher add-on fees than could parents in poorer districts.

The second result of Chile’s voucher program was that higher-income families were more likely to use the vouchers for private schools. In 1990, nine years after the voucher program was established, 72 percent of families in the lowest 40 percent of the income distribution attended public schools. Among the next-highest 40 percent, 51 percent of families sent their children to public schools (and 43 percent to subsidized private schools). Of families in the top 20 percent of the income distribution, only 25 percent sent their children to public schools, 32 percent used subsidized private schools, and 43 percent sent their children to (nonvoucher) private schools.

The third finding was that there was no improvement in student achievement, contrary to predictions by voucher proponents. For example, an evaluation based on nationally standardized (although not strictly comparable) achievement tests in Spanish and mathematics found that the scores of fourth-graders declined between 1982 and 1988. Beyond these general averages, all the results reported by socioeconomic class have to be considered tentative, because test scores are reported only as school averages, and the socioeconomic level of the school is based on a questionnaire given to school administrators. Keeping such limitations in mind, comparisons across socioeconomic levels suggest that students in lower-income public schools recorded the sharpest drop in test scores, but scores of students in lower-income subsidized private schools also fell (Prawda 1993). Students in the middle—socioeconomic level schools had small increases in test scores whether they were in public or subsidized private schools.

One estimate of the means of the 1990 national test showed an increase of 9 percent in Spanish and 11 percent in math, bringing these scores to about the same level as they were in 1982 (Rounds Parry 1994). Another estimate showed declining test scores for all groups between 1988 and 1990 (Espinola 1993). Both Rounds Parry and Espinola agree that pupils in middle—socioeconomic level schools averaged higher scores in private than in public schools, but that pupils in the lowest-level schools tended to do better in public schools. There is general agreement that private subsidized schools tend to “cream” (select) the
better pupils. Thus these results may show only that private schools in middle-
and even lower-middle-income areas are more successful at selecting academi-
cally capable pupils than are private schools in low-income areas. The way the
data are collected makes correcting for such selection bias difficult, if not
impossible.

In 1990 Chile’s newly elected government increased spending on education
and raised the value of the voucher. Low socioeconomic-level schools were singled
out for targeted spending and technical assistance (Comite Tecnico 1994). Ac-


Comite Tecnico 1994). According to Ministry of Education figures, test scores rose across socioeconomic
levels in both public and private schools between 1990 and 1992 and leveled off
over the next two years. So after thirteen years (1981–94) of a national voucher
plan, student achievement levels are probably back to where they were in the
early 1980s, or perhaps slightly higher. Pupils in private schools that draw from
higher socioeconomic levels tend to have higher scores than those in public schools
of the same socioeconomic level. Yet, absent any correction for selection bias,
the higher scores for these private schools do not constitute evidence that they
are more effective.

It is generally agreed that public schools in Chile cost more per pupil in gov-
ernment funds than do subsidized private schools (Winkler and Rounds 1993),
because class sizes are somewhat higher in private schools (Comite Tecnico 1994).
Based on public spending per pupil, the only attempt to estimate the relative
cost-effectiveness of Chilean public and private schools shows a small advan-
tage for private education. But subsidized private schools—particularly those in
middle-class areas—charge fees in addition to vouchers, and all private schools
require uniforms and other family expenses not demanded by municipal schools.
Estimates of these additional costs are not available for Chile, but based on
studies in other countries, they could be as high as a third of the total costs of
schooling, even at the primary level (Carnoy and Torres 1994). This raises the
complex issue of biased estimates of private versus public schools costs (see

The Milwaukee Case

In Milwaukee qualifying low-income pupils received a $2,500 voucher (by
1995–96 the amount had risen to $4,375) to attend nonreligious private schools.
More qualified pupils applied than could be accommodated, so acceptances were
issued by lot. An evaluation of the academic performance of private-school
voucher pupils and public school students of the same socioeconomic background
(corrected for selection bias) found no significant difference in outcomes be-
tween the two groups, whether the students had been in private school for one,
two, three, or four years (Witte, Thorn, and Prichard 1996). This result is espe-
cially strong because the sample of those who had been in private schools for
more than a year was increasingly self-selected. Several students who began in
the private schools in the first year of the program later either moved or chose to
return to public schools (the attrition rate was 30 percent). In addition, several of the original private schools closed, forcing voucher students to return to public schools. As a result only the most motivated students in the successful private schools remained after four years.

Another recent study, by Greene, Peterson, and Du (1996), estimates significantly higher test scores for third- and fourth-year voucher students compared with students who had applied to the voucher program but had not been accepted in the lottery; this was true, however, only in the case that socioeconomic background differences are not corrected for. They claim that their study makes a true correction for selection bias because it compares voucher students with an identical group of students who had qualified for the voucher program and thus were equally motivated and met the low socioeconomic requirements of the voucher plan. Witte (1996) argues, however, that the Greene-Peterson-Du methodology is never fully specified and appears to be flawed and that once socioeconomic status is controlled for, the differences in test scores are no longer significant.

On average, the private elementary schools in Milwaukee's choice program seem to have provided instruction at lower cost than did the public schools, at least in the first years of the program. The cost data are not strictly comparable, however, because private school tuition fees do not reflect the real cost per pupil. But as teachers in the private schools were paid about a third less than those in public schools and class size was approximately the same, the implication is that the unit costs for private schools are about a third less than those for public schools. Because test scores do not differ across comparable groups, this means that private schools were more effective than public schools. However, salaries of teachers in private schools have apparently risen more rapidly than those of public-school teachers. At today's $4,375 voucher, per-pupil costs in private schools appear to be approximately the same as those in Milwaukee's public elementary schools.

Is Private Education Better for Low-Income Students?

In a review of the European situation, John Ambler (1994, p. 470) argues that in Britain, France, and the Netherlands "the primary negative effect of school choice is its natural tendency to increase the educational gap between the privileged and the underprivileged." The Netherlands is especially interesting because about 70 percent of pupils attend fully subsidized private schools. There are waiting lists to get into the "better," more selective schools, which charge fees on top of the voucher subsidy. To make lower-income, "disadvantaged" pupils more attractive, the government provides them with a voucher that is 25 percent more than the average. Even so, the students who end up in municipal (public) schools are mainly those lower-income pupils (with the larger vouchers) who still cannot attend the more desirable schools.
In Chile private schooling is widely available to lower-income families in urban municipalities, but, as argued above, analysis of test scores suggests that pupils from such families may have gained least from the privatization process. Although West claims that privatization helped expand Chile’s secondary education (secondary education did expand rapidly in the 1980s), and therefore increased social mobility for low-income Chileans, there is no evidence that secondary education would not have expanded even more rapidly through public education, as it did in many countries of Latin America and Asia in the 1970s and 1980s. After Haiti, Brazil has the most privatized secondary education system in Latin America, and access by low-income pupils is highly limited.

Colombia’s voucher plan is also targeted at students from low-income families. It serves those who could not qualify academically for free public secondary schooling, but whose families are willing to pay the tuition at private schools over and above the voucher. There appears to be little “substitution” effect, which means that the plan really does expand the number of places available to low-income students, and in its first years, the plan is allowing most of these otherwise academically excluded students to attend traditional private parochial schools that have excess capacity and relatively high academic standards. Those students undoubtedly benefit from the voucher plan. But once the capacity of existing schools is reached, private schools of much lower quality may open for business. This is already a problem, and the Colombian government has stepped in to regulate these schools (Calderon 1996). Is expanded low-quality private education paid for by government funds superior in any way to expanded (and possibly low-quality) public education costing about the same? Vouchers could have a high payoff for the “best” group of excluded low-income students up to the point at which they can be absorbed by existing higher quality private schools, but beyond that, payoffs could decline rapidly.

Is There a Downside to Privatization?

West argues specifically that there is little to lose from privatizing education and much to gain. Although there is no convincing evidence that the gains exist, that does not mean privatization should not be tried. That is, unless it could also lead to a net social loss.

One possible loss is greater inequality. Evidence suggests that privatizing education through vouchers for all pupils may produce a more unequal distribution of educational quality and choice than under a predominantly public system. If the choice to attend a private school is not equally distributed across the population (as is the case in Chile and in Europe), private (and even public-school choice plans) might widen the divide between urban and rural and between lower- and higher-income students (Levin 1995). Chile’s voucher plan appears to have widened the gap between high- and low-income students in terms of test scores without increasing the overall level of academic achievement. It also re-
duced government efforts to improve schooling, in part because the free market was expected to increase achievement. Only after the central government intervened with increased in-service training, new curriculum standards, and technical assistance to low-income schools did achievement levels appear to increase across the board.

Better private schools for middle-class pupils may produce higher test scores than do public schools, but mainly by keeping out hard-to-manage or less-motivated pupils. As Winkler and Rounds (1993) report, parents in Chile select schools mainly on the basis of the characteristics of the pupils that already attend those schools. Willms and Echols (1992) found the same behavior in Scotland, where a choice plan allowed parents to select from among public schools. But private schools are not uniformly exclusive or effective; some accept (and keep) less-motivated or less able youths who score poorly. Why would parents send their children to such schools? One answer to this question is that because some private schools are associated with a higher social stratum and higher test scores (not necessarily because they are more effective, but because they limit enrollment), their reputation could rub off on all private schools. Winkler and Rounds (1993) cite an unpublished experiment in which changing the school name from Spanish to English and introducing school uniforms in public schools (techniques used by private subsidized schools in recruiting pupils) increased enrollments in public schools.

West could claim that well-informed parents would soon withdraw their children from less effective private schools and reenroll them in public schools. But that might not be the case if vouchers had a negative effect on public schools. If, for example, private schools attract the better students from public schools, the latter will increasingly be seen as suitable only for low-income or slow students. West argues that there is no such effect, but he falls back on evidence from voucher plans that cater only to less able, low-income youths, such as those in Milwaukee and Colombia. West cites Witte’s assertion that private schools in Milwaukee did not draw the “cream of the crop” or help those who needed help the least as evidence that private schools would not do either under a generalized voucher plan. But the voucher plans in Milwaukee and Colombia are restricted to those very low-income pupils who need help the most. Other students could not qualify for vouchers. Witte was simply reporting that the objectives of this restricted plan were being met and that those low-income students who were dissatisfied with the public schools and so participated in and stayed with the program were, by and large, satisfied. But recent attempts by the State of Wisconsin to offer vouchers to the school population as a whole have prompted concern in Milwaukee’s African-American community, which backed the targeted voucher plan, that the broader eligibility will favor higher-income families (Miner 1995).

School choice, after all, is also affected by the costs of transportation and by how much parents know about various schools. West implicitly assumes that neither is a factor in determining the alternatives available. Yet studies of choice
plans in both education and health care suggest that the location of alternatives is crucial and that the higher the income of the consumer, the more quality, rather than location, plays the key role (Elmore 1990). Those who are better off can provide or pay for private transportation for their children to a wider range of school locations, so the same voucher for all means more restricted access for the poor.

Similarly, less-educated families are less likely to search out and use information on the quality of educational alternatives. This does not mean that less-educated families are not aware of the quality of their children’s education, particularly if that education is especially poor or especially good. Rather it suggests that choice systems “require substantial information on alternatives in order to support effective decision-making . . .” and that “education is a complex service in which even those with more sophistication must rely heavily on word-of-mouth and interpretation of incomplete data on school quality” (Levin 1995, p. 10). Proponents of privatization sometimes characterize the view that less-educated parents have less information as “elitist,” but all the empirical evidence suggests that it reflects reality. “Studies of parental information on choice show that less educated, minority, and lower-income parents are often unaware of their choice options and are not cognizant of differences among schools of choice . . . [and] studies have consistently shown that . . . the lower the socioeconomic status of a family, the less likely that it will ‘choose’ the school attended by its offspring . . .” (Levin 1995, pp. 11-12).

Thus introducing competition between private and public schools (and even among public schools) may actually exacerbate inequalities in the educational system. By increasing the incentive to rely on public relations techniques (such as changing to an English name) that have nothing to do with the school’s effectiveness, the free market preys on a less-educated clientele. The idea that increased competition in the production of complex, difficult-to-define products, such as education, automatically increases social welfare is simply wrong. This is especially true where private profit is the driving force behind the competition. That is why governments have been called upon to regulate private education and protect the public interest in the voucher plans cited here.

It is also possible that privatizing education could reduce the incentives of the politically powerful middle class to push for more resources to education. Although these parents would allegedly still have an interest in raising the average amount of the voucher to offset the cost of private (or public) education, it is also possible that private schools could raise their fees as the vouchers increased in value. When vouchers were introduced in Sweden, for example, the private schools that previously charged tuition raised their fees by 9 percent (Miron 1993).

Furthermore, a privatized system could affect the distribution of public money, shifting funds from teachers’ compensation to private school operators. Between 1983 and 1990, teachers’ salaries in Chile dropped sharply, especially in the more privatized secondary school system. The decline was partly the result of a
steadily lower voucher value, but it was also related to the profits marked for
the owners of the schools. Although many would applaud lower salaries for
teachers as one way to make education less costly, this approach is not necessar-
ily more cost-effective. Teachers are crucial to improving education. If labor
markets work properly, lower salaries for teachers will attract less talented people.
Because teaching conditions are usually more difficult in schools with low-
income pupils, the effect of reducing teacher pay could have its greatest impact
on such schools.

In an analysis of Great Britain's 1980 public-school choice law, Willms and
Echols (1992, p. 347) concluded that the moderate gains that may have been
achieved through contextual effects when children went to better schools may
have had high costs for the educational system as a whole:

Schools serving pupils in disadvantaged areas will be receiving incorrect
signals; many of the them will lose pupils to higher SES [socioeconomic
status] schools despite effective teaching practices. Some high SES
schools will also receive incorrect signals, because many parents are
choosing these schools even though their performance is mediocre or
poor when compared with schools with similar social class intakes.

Privatization or Public Improvement?

There is no persuasive evidence that private schools are more effective than
public schools and the evidence that they are more cost-effective is mixed. More-
over, not every income group's welfare is raised through increased choice. Even
if satisfaction and educational attainment were raised in absolute terms, a loss
in relative position could make lower-income groups feel worse off than before.
Unlike income, the absolute level of education or even quality of education at-
tained is probably less important than the relative amount and quality of educa-
tion an individual or group attains, since the payoff associated with a given level
of education declines rapidly as a large fraction of the population attains it
(Carnoy 1972).

West would limit vouchers to low-income families and restrict privatization
to the poor as a means of equalizing an otherwise unequal system. This is a good
answer to the distribution of choice problem, but it does not address the ques-
tion of effectiveness. Private schools do not seem to improve the achievement of
pupils from low-income families any more than they do for the pupil population
as a whole. Even so, data from Milwaukee and Colombia suggest that a voucher
program limited to the poor could make low-income families feel better off than
they would if the present system of education were left untouched. Yet limiting
vouchers to poor families contradicts the broader argument of privatization
advocates.

The voucher question boils down to how societies feel politically about the
responsibilities and role of the state in education. The political judgment that
the public sector is not capable of adequately providing education but should finance it supports the view that vouchers for private education are appropriate even if they are unlikely to improve the average level of learning. Many of the arguments against privatizing education can be applied to public choice plans, such as Scotland’s. So implementing a voucher plan that allows parents to send their children to public schools outside of their neighborhoods or to regulated “charter” schools that serve niche markets within public school districts is also a political decision, one that is bound to favor certain groups more than others. If a society wants to offset the effects of family poverty on children’s education, it could use vouchers limited to the poor as one element of an antipoverty program. If a society is convinced that both dissatisfied and more inert parents are better served by efforts to improve neighborhood public schools than by efforts to provide exit options to the dissatisfied, it should focus on improving public education.

I would like to believe, with Professor West, in a panacea that could make everyone learn more without investing enormous time and effort in improving children’s nutrition, home lives, and the way all schools deliver knowledge. Our task as educators and social reformers would be that much simpler. Unfortunately, vouchers tend to divert attention from the overall complexity of the learning problem rather than providing a real solution.

Note

Martin Carnoy is professor of education and economics at Stanford University.

References

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


The best in development statistics just got better!

World Development Indicators 1997

The World Bank introduces its most comprehensive, statistics-packed product to date—World Development Indicators 1997. Formerly the statistical appendix to the World Development Report, these comprehensive data are now available in their own volume and have been enlarged to include more than 80 tables with some 400 indicators. This major new publication provides an expanded view of the world economy for more than 130 countries—with chapters focusing on people, economy, environment, states and markets, and global links. Concise, insightful commentary tells the story of how people live and work, and how countries are expanding and changing.

April 1997 350 Pages Stock no. 13701 (ISBN 0-8213-3701-7) $60.00

World Development Indicators on CD-ROM

This comprehensive database, which replaces World Data, contains most of the underlying time-series data for the World Development Indicators and World Bank Atlas. We've added powerful new features—now you can generate maps and charts, and download your results to other software programs. Requires Windows 3.1.


World Bank Atlas 1997

One of the Bank's most popular offerings, the Atlas has been redesigned as a companion to the World Development Indicators. Tables, charts, and 21 colorful maps address the development themes of people, economy, environment, and states and markets. This easy-to-use, inexpensive book is an international standard in statistical compilations and an ideal reference for office or classroom. Text, maps, and references appear in English, French, and Spanish.

April 1997 48 Pages Stock no. 13576 (ISBN 0-8213-3 576-6) $15.00

World Bank Publications

For US customers, contact The World Bank, P.O. Box 7247-8619, Philadelphia, PA 19170-8619. Phone: (202) 473-1155, Fax: (202) 522-2627. Shipping and handling: US$5.00. For airmail delivery outside the US, add US$8.00 for one item plus US$6.00 for each additional item. Payment by US$ check drawn on a US bank payable to the World Bank or by VISA, MasterCard, or American Express. Customers outside the US, please contact your World Bank distributor.
# Order Form

**CUSTOMERS IN THE UNITED STATES**

Complete this coupon and return to:
The World Bank
Box 7247-8619
Philadelphia, PA 19170-8619
USA.

To have your order shipped faster, charge by credit card by calling (202) 473-1155 or send this competed order coupon by facsimile by dialing (202) 522-2627.

**CUSTOMERS OUTSIDE THE UNITED STATES**

Contact your local Bank publications distributor for information on prices in local currency and payment terms. 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.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Title</th>
<th>Stock #</th>
<th>Price</th>
<th>Total price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*SHIPPING AND HANDLING* charges are $5.00 per order. If a purchase order is used, actual shipping will be charged. For air mail delivery outside the United States, add $8.00 for one item plus $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 (required to validate all orders)

- Bill me. (Institutional customers only. Purchase order must be included.)

**PLEASE PRINT CLEARLY**

Name

Address

City  State  Postal code

Country  Telephone

Subtotal cost US$ __________

Shipping and handling* US$ __________

Total US$ __________