# THE WORLD BANK ECONOMIC REVIEW

Volume 9 September 1995 Number 3

# THE WORLD BANK ECONOMIC REVIEW



*The World Bank Economic Review* is a professional journal for the dissemination of World Bank–sponsored research that informs policy analyses and choices. It is directed to an international readership among economists and social scientists in government, business, and international agencies, as well as in universities and development research institutions. The *Review* emphasizes policy relevance and operational aspects of economics, rather than primarily theoretical and methodological issues. It is intended for readers familiar with economic theory and analysis but not necessarily proficient in advanced mathematical or econometric techniques. Articles will illustrate how professional research can shed light on policy choices. Inconsistency with Bank policy will not be grounds for rejection of an article.

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

The *Review* may on occasion publish articles on specified topics by non-Bank contributors. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages in length to the Editor.

The views and interpretations 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.

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

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

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

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


ISSN 0258-6770



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

# THE WORLD BANK ECONOMIC REVIEW

Volume 9 September 1995 Number 3

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3: 341-371

# The Soviet Economic Decline

William Easterly and Stanley Fischer

*Soviet growth from 1960 to 1989 was the worst in the world after we control for investment and human capital; the relative performance worsens over time. There is some evidence that the burden of defense spending modestly contributed to the Soviet debacle. The declining Soviet growth rate from 1950 to 1987 can be accounted for by a declining marginal product of capital with a constant rate of growth of total factor productivity. The Soviet reliance on extensive growth (rising capital-to-output ratios) was no greater than that of market economies, such as Japan and the Republic of Korea, but a low elasticity of substitution between capital and labor implied especially acute diminishing returns to capital compared with the case in market economies.*

Why did the per capita economic growth of the former U.S.S.R. decline and then stop, contributing to the final collapse of the Soviet economic and political system? Accounts of the declining Soviet economic growth emphasize different causes: the Soviet reliance on extensive growth, which, given the slow growth of the labor force and the falling marginal productivity of capital, eventually ran out of payoff; the declining rate of productivity growth or technical progress associated with the difficulties of adopting and adapting to the sophisticated technologies being introduced in market economies; the defense burden; and a variety of special factors relating to the absence of appropriate incentives in the Soviet system, including corruption and demoralization (Banerjee and Spagat 1991; Bergson 1987b; Desai 1987; Ofer 1987; and Weitzman 1970).

In this article, we examine alternative explanations with special care to place the Soviet growth performance in an international context. In section I, we start with an overview of the data and of the Soviet growth record and compare it to that of other countries by using a standard growth regression. We also examine the role of the defense burden in the performance of the Soviet economy and then compare the Soviet pattern of extensive growth—rising capital-to-output ratios—to that of other

William Easterly is with the Policy Research Department at the World Bank, and Stanley Fischer is the First Deputy Managing Director at the International Monetary Fund. The authors are grateful to Karen Brooks, Padma Desai, Alan Gelb, Eugene Gravilenkov, Barry Ickes, Barry Kostinsky, Martha de Melo, Gur Ofer, Sergio Rebelo, Bryan Roberts, Marc Rubin, Randi Ryterman, Martin Schrenk, Marcelo Selowsky, Michael Walton, Alwyn Young, and anonymous referees for comments; to seminar participants at the Massachusetts Institute of Technology, the University of Michigan, the Federal Reserve Board, the University of California at Los Angeles, and the World Bank; to Professor Mark Schaffer of the London School of Economics for kindly making his historical Western, official, and Khanin's data available in machine-readable form; and to Elana Gold and Mary Hallward for diligent research assistance.

economies. In section II, we assess the role of extensive growth in the decline of growth by reexamining and updating estimates of the aggregate production function and in particular the elasticity of substitution between labor and capital. We then discuss how Soviet elasticities of substitution compare with those in other economies. In the conclusion, we offer an interpretation of our results.

## I. The Soviet Growth Record

The fundamental problem in evaluating Soviet growth is the quality of data (see Fischer 1992). As a result of both methodological problems—particularly deflating nominal data—and incentives to misreport output within the Soviet system, official data on Soviet output overstate growth. Western analysis of Soviet growth relies on classic studies, including Bergson (1961) and the U.S. Central Intelligence Agency (CIA 1982, various years). The CIA makes the working assumption that physical quantities as presented in the official data were not systematically misreported. Thus the difference between the Western estimate that per capita Soviet gross national product (GNP) increased between 1928 and 1987 by 3.0 percent a year (4.3 percent a year for aggregate GNP) and the official (Soviet) estimate that net material product (NMP) per capita increased by 6 percent a year results mostly from pricing corrections and also from differences in the coverage of NMP and GNP.[1] The classic Western estimates generally assume that data on Soviet investment and capital are more accurate than data on output. This view is supported by Bergson (1987a) and disputed by Wiles (1982). The Western data through 1985 are conveniently summarized in Ofer (1987).

The data set we use in this article is constructed from Western data on output, industrial production, employment, and the capital stock for the U.S.S.R. as a whole. It includes the series on value added and capital stocks in industry by Powell (1963), CIA (1982), and CIA (various years) and the series on GNP, labor input, and capital stock for the entire economy from Moorsteen and Powell (1966), Powell (1968), CIA (1982), CIA (various years), and Kellogg (1989). All these series use 1937 rubles for the period 1928–60, 1970 rubles for the period 1960–80, and 1982 rubles for the 1980s. The direct source of our data sets is Gomulka and Schaffer (1991), who spliced together series from the sources described.

We look at how our preferred data set differs from both the official U.S.S.R.–wide data on real output, industrial production, employment, and the capital stock in the material sector in 1973 rubles, and from Khanin's (1988) data, which are also at U.S.S.R.–wide level, for output, employment, and the capital stock in the material sector.[2] (Note that both the official and

1. This section draws on Fischer (1992). The Soviet concept of NMP omitted from GNP those services not directly related to production, such as passenger transportation, housing, and the output of government employees not producing material output.

2. Ericson (1990) argues that the Khanin data are preferable to the Western series, but Bergson (1987a, 1991a) criticizes the Khanin data for a poorly documented methodology and the use of unweighted averages of physical indicators.

Table 1. *Growth Rates in the Soviet Economy, Based on Different Data Sources, 1928–87*
(average annual percent)

| | *Western data* | | *Official data* | | *Khanin's data* |
|---|---|---|---|---|---|
| *Indicator and period* | *Total economy* | *Industry* | *Material sectors* | *Industry* | *Material sectors* |
| *Output per worker* | | | | | |
| 1928–39 | 2.9 | 5.0 | 11.4 | 12.5 | 0.9 |
| 1940–49 | 1.9 | 1.5 | 2.1 | 0.1 | 1.0 |
| 1950–59 | 5.8 | 6.2 | 8.3 | 8.9 | 5.3 |
| 1960–69 | 3.0 | 2.8 | 5.4 | 5.7 | 2.7 |
| 1970–79 | 2.1 | 3.4 | 4.1 | 5.2 | 1.2 |
| 1980–87 | 1.4 | 1.5 | 3.0 | 3.4 | 0.2 |
| *Capital per worker* | | | | | |
| 1928–39 | 5.7 | 6.5 | 8.7 | 11.9 | 5.9 |
| 1940–49 | 1.5 | 0.1 | 2.7 | 1.5 | 1.3 |
| 1950–59 | 7.4 | 3.9 | 7.7 | 8.0 | 3.5 |
| 1960–69 | 5.4 | 3.4 | 7.1 | 6.1 | 3.8 |
| 1970–79 | 5.0 | 4.1 | 6.8 | 6.3 | 1.9 |
| 1980–87 | 4.0 | 4.0 | 5.3 | 5.6 | 0.1 |

*Source:* Gomulka and Schaffer (1991).

Khanin's data are presented for the material sectors only and do not include consumer services).

### *Soviet Growth in International Comparison*

Growth rates of the Western series for different periods are presented in table 1, with the official data and Khanin's data given for comparison. The Western growth rates for output per worker and capital per worker are well below the official rates, and Khanin's data are in turn below the Western data. All series show growth declining sharply starting from the 1950s. Although we have greater confidence in the Western data than in the likely exaggerated official estimates and the inadequately documented methodology of Khanin, the disparities suggest some caution in the use of all data sets on aggregate Soviet growth.

How does the Soviet growth record compare with the rest of the world's? We use the Western gross domestic product (GDP) series to compare Soviet per capita growth from 1960 to 1989 with per capita growth rates for 102 economies. (We look here at per capita rather than per-worker growth to enlarge the sample of comparators and make it consistent with the cross-sectional growth literature.) The next-to-last column of table 2 shows that Soviet per capita growth was slightly above the global average in the period 1960–89.

Soviet growth no longer looks respectable, however, once we control for the standard growth determinants from the empirical literature. The last column of table 2 shows the residual from inserting the U.S.S.R. into the core regression of Levine and Renelt (1992), which relates growth to initial income, population

Table 2. *The U.S.S.R. in the Levine-Renelt Growth Regression, 1960–89*

| *Variable* | *Per capita income, 1960*[a] *(1985 U.S. dollars)* | *Population growth, 1960–89 (average annual percent)* | *Secondary school enrollment, 1960 (percent)* | *Investment ratio to GDP, 1960–89*[b] | *Predicted growth (average annual percent)* | *Actual per capita growth, 1960–89 (average annual percent)* | *Growth residual* |
|---|---|---|---|---|---|---|---|
| Average for sample excluding the U.S.S.R. | 1,792 | 2.07 | 21 | 21 | 2.00 | 2.00 | 0.00 |
| Average for the U.S.S.R. | 2,796 | 1.05 | 58 | 29 | 4.70 | 2.36 | −2.34 |
| Predicted effect on the growth differential between the world average and the U.S.S.R. | −0.35 | 0.39 | 1.17 | 1.49 | 2.70 | 0.36 | −2.34 |

*Note:* The sample excluding the U.S.S.R. has 102 countries.

a. Per capita income in 1960 for the U.S.S.R. is calculated based on the Bergson (1991b) purchasing power parity value for 1985 and backcasting to 1960 using the actual per capita growth given in the sixth column.

b. The CIA series for investment is available at five-year intervals for 1960–75; we interpolated to obtain values for the other years.

*Source:* The regression is from Levine and Renelt (1992). For all countries except the U.S.S.R.: Levine and Renelt (1992); for the U.S.S.R.: actual per capita growth is from Gomulka and Schaffer (1991) and Marer and others (1992); per capita income is from Bergson (1991b); population growth is from Feschbach (1983), IMF and others (1991), and Marer and others (1992); secondary school enrollment is from UNESCO (1975) and Marer and others (1992); investment rate is from CIA (various years) and Marer and others (1992).

growth, secondary school enrollment, and the ratio of investment to GDP. The Levine-Renelt regression including the U.S.S.R. is as follows:

(1) Per capita growth 1960–89 = −0.83 + 17.49 investment 1960–89 − 0.35 GDP per capita 1960
(0.85) (2.68) (0.14)

+ 3.16 secondary enrollment 1960 − 0.38 population growth 1960–89 − 2.34 dummy for U.S.S.R.
(1.29) (0.22) (1.43)

103 observations
$R^2 = 0.46$
standard errors in parentheses

Except for population growth, Levine and Renelt showed these variables to be robust to alternative specifications in growth regressions (although concerns about endogeneity remain). The regression results are identical to the Levine-Renelt original, which excludes the U.S.S.R., because we are dummying out the Soviet observation.

Excepting initial income, the values of the Soviet right-hand side variables should have implied very rapid growth.[3] As shown in the third row of table 2, high Soviet human and physical capital accumulation should have increased Soviet growth by 2.7 percentage points above the world average. As it was, growth was only barely above average, hence the large negative residual of 2.3 percentage points for 1960–89. Soviet per capita income in 1989 was only half of what it would have been if the average relationship between growth and the right-hand-side variables had held for the period 1960–89.

The Soviet residual in this ordinary least squares regression is not actually significant in a two-sided test at the 5 percent level. However, it is notable that the only economies with worse residuals are generally both small and poor: Guinea-Bissau, Jamaica, Liberia, Peru, Suriname, and Zambia. The presence of so many small and poor economies among the large outliers makes us suspect heteroskedasticity. The suspicion is justified. We split the 1960–89 sample into thirds on the basis of total real GDP (that is, population times per capita income valued at purchasing power parity) and reran the above regression for the top and bottom thirds ranked by total GDP. (The U.S.S.R. is included in the top third ranked by total GDP, and we continued to dummy it out.) The Goldfeld-Quandt test statistic for heteroskedasticity indicates that we can reject homoskedasticity.[4] The test results are as follows:

Sum of squared residuals in third of sample with lowest real GDP: 88.3
Sum of squared residuals in third of sample with highest real GDP: 33.9
$F(29, 28) = 2.61$ (significant at 1 percent level)

3. Some would argue that even initial income was more favorable (that is, lower) for subsequent growth than we have depicted it here, because some argue that the Bergson (1991b) estimate of Soviet per capita income that we have used is too high. This would make the Soviet conditional growth performance look even worse.

4. The Goldfeld-Quandt test statistic is equal to the ratio of the sum of squared residuals in these two subsample regressions and is distributed as an *F*-statistic with the number of degrees of freedom of the numerator and denominator corresponding to the degrees of freedom in the subsample regressions.

On the basis of test results, we performed weighted least squares using the log of total real GDP as the weighting series. The results are as follows:

(2) Per capita growth 1960–89 = −0.43 + 15.93 investment 1960–89 − 0.28 GDP per capita 1960
(0.73) (2.19) (0.08)

+ 2.56 secondary enrollment 1960 − 0.24 population growth 1960–89 − 2.28 dummy for U.S.S.R.
(0.73) (0.16) (0.48)

102 observations
$R^2$ (weighted) = 0.84
standard errors in parentheses

The Soviet dummy becomes highly significant with weighted least squares, with a *t*-statistic of 4.8. Taking into account that the only economies doing worse than the U.S.S.R. were small makes the Soviet performance look even worse. After correcting for heteroskedasticity, the Soviet economic performance conditional on investment and human capital accumulation was the worst in the world from 1960 to 1989.

How does the comparative Soviet performance evolve over time? Because the World Bank data used by Levine and Renelt begin only in 1960, we compared the Soviet performance also with the cross-country Summers-Heston (1991) data set that extends back to 1950. We performed a pooled time-series, cross-sectional regression using decade averages for the same specification as before (except that unfortunately we had to omit the secondary school enrollment variable for lack of reliable Soviet data for the 1950s). We used the same Soviet data as in the previous regression, but now broken down by decade. For each decade, we used intercept dummies as well as a separate Soviet dummy. We continued to use weighted least squares with the weighting series being the log of total GDP, because the Goldfeld-Quandt statistic still indicated a significantly larger variance for small economies.[5] The results are as follows:

(3) Per capita growth by decade = 0.022 + 0.120 investment/GDP by decade
(0.005) (0.016)

− 1.5E-06 GDP per capita, initial year of each decade − 0.626 population growth by decade
(3.6E-07) (0.143)

+ 0.005 1960s dummy − 0.005 1970s dummy − 0.015 1980s dummy
(0.004) (0.003) (0.003)

+0.024 dummy for U.S.S.R. 1950s − 0.008 dummy for U.S.S.R. 1960s
(0.011) (0.010)

− 0.017 dummy for U.S.S.R. 1970s −0.023 dummy for U.S.S.R. 1980s
(0.009) (0.009)

391 observations
$R^2$ (weighted) = 0.54
standard errors in parentheses

Although economic growth decelerated worldwide in the 1970s and even more in the 1980s, growth deceleration in the U.S.S.R. is notable even by comparison

5. The *F*-statistic for the ratio of the sum of squared residuals in the bottom third to that in the top third of the sample ranked by total GDP (in purchasing power parity prices from Summers and Heston 1991) is $F(124, 121) = 2.03$, which is significant at the 1 percent level.

with the world pattern: Soviet economic growth was significantly above the world average in the 1950s and significantly below even the poor world growth of the 1980s. Note, especially, the good performance of the U.S.S.R. in the 1950s, even controlling for high investment: it suggests that whatever the weaknesses of Soviet central planning in hindsight, these weaknesses were unlikely to have been apparent before 1960.

### *Possible Explanations for Poor and Declining Soviet Growth*

Before turning to the classic, extensive growth hypothesis, we first consider two other possible factors in the relative Soviet decline: the defense burden and Soviet disincentives for innovation. Could the poor and declining growth performance be explained by the burden of defense on the Soviet economy? Although measurement is problematic, the burden seems to have been high and rising. In table 3, we show some estimates of the Soviet defense burden as a share of GDP. Over the entire period since 1928, Soviet defense spending rose from 2 percent of GDP to the much higher levels of the mid- and late 1980s, around 15 to 16 percent of GDP. Over the period 1960–89 in which the Soviet growth decline occurred, the rise in the defense burden is more modest—from 10 to 13 percent in 1960 to 12 to 16 percent in the 1980s.

The international evidence for adverse effects of defense spending on growth is ambiguous—see Landau (1993) for a recent survey. Landau (1993) himself finds an inverted U relationship: military spending below 9 percent of GDP has a positive effect on growth, but above 9 percent of GDP it has a negative effect on growth. To see whether this affects the Soviet dummy in the growth regressions, we inserted defense spending into the decade-average growth regressions performed earlier (equation 2). We also included a variable measuring war casualties per capita on national territory to ensure that the military spending variable was not simply proxying for wars. Because the military spending data are available only for recent periods, we used data from the 1980s only. Landau covers only developing economies, so we used, instead, data from Hewitt (1993) that covers all economies (including the U.S.S.R. itself). The data for both Landau and Hewitt are mainly from the Stockholm International Peace Research Institute. The data on war casualties are from Easterly and others (1993). The regression including a quadratic function of military spending is as follows:

(4) Per capita growth 1980–88 = −0.003 + 0.127 investment/GDP 1980–88 − 2.7E-06 GDP per capita 1980
(0.017) (0.038) 1.1E-06)

−1.34 population growth 1980–88 +0.007 secondary enrollment 1970 + 0.0081 military spending/GDP 1980–88
(0.38) (0.017) (0.0024)

−0.00041 (military spending/GDP)$^2$ − 0.746 war casualties per capita 1980–88 − 0.0155 dummy for U.S.S.R
(0.0001) (0.343) (0.0268)

least squares weighted by log of total GDP
77 observations
$R^2$ (weighted) = 0.59
standard errors in parentheses

Table 3. *The Soviet Defense Burden as a Share of GDP, Based on Different Data Sources, 1928–87*
(percentage of GDP)

| Year | Ofer | Brada and Graves | | Steinberg |
|---|---|---|---|---|
| | | *High estimate* | *Low estimate* | |
| 1928 | 2 | — | — | — |
| 1950 | 9 | — | — | — |
| 1960 | 12 | 13.34 | 9.90 | — |
| 1961 | — | 13.86 | 10.60 | — |
| 1962 | — | 14.93 | 11.39 | — |
| 1963 | — | 15.49 | 12.32 | — |
| 1964 | — | 15.03 | 12.17 | — |
| 1965 | — | 14.49 | 11.79 | — |
| 1966 | — | 14.11 | 11.54 | — |
| 1967 | — | 14.40 | 11.95 | — |
| 1968 | — | 14.45 | 12.14 | — |
| 1969 | — | 14.61 | 12.08 | — |
| 1970 | 13 | 13.83 | 11.48 | 13.28 |
| 1971 | — | 13.56 | 11.30 | 13.76 |
| 1972 | — | 13.80 | 11.34 | 13.61 |
| 1973 | — | 13.33 | 11.03 | 13.14 |
| 1974 | — | 13.71 | 11.28 | 13.15 |
| 1975 | — | 14.14 | 11.53 | 13.57 |
| 1976 | — | 14.32 | 11.62 | 13.30 |
| 1977 | — | 14.07 | 11.26 | 12.98 |
| 1978 | — | 14.00 | 11.09 | 13.08 |
| 1979 | — | 14.53 | 11.43 | 13.05 |
| 1980 | 16 | 15.06 | 11.82 | 13.91 |
| 1981 | — | 15.48 | 11.75 | 14.03 |
| 1982 | — | 15.36 | 11.70 | 14.58 |
| 1983 | — | 15.51 | 11.63 | 14.36 |
| 1984 | — | 15.55 | 11.57 | 14.37 |
| 1985 | — | — | — | 14.79 |
| 1986 | — | — | — | 14.49 |
| 1987 | — | — | — | 14.63 |

— Not available.

*Note:* The Ofer growth rates are based on current rubles; those of Brada and Graves and Steinberg are based on constant rubles.

*Source:* Brada and Graves (1988), Ofer (1987), and Steinberg (1987, 1990).

We confirmed Landau's result of an inverted U-shaped relationship between growth and defense spending. Military spending reduces the magnitude and significance of the Soviet dummy. However, as Landau also notes, this result is not very robust; omitting Israel and Syria from our sample eliminates the significance of military spending. The defense explanation for the Soviet decline is plausible but not firmly established with cross-sectional data. We will test the defense hypothesis further with the Soviet time series in the production function estimates in the next section.

Another possible explanation for poor and declining Soviet growth could be adverse incentives under central planning for technological innovation (Berliner 1976). Recent theoretical and empirical literature argues that endogenous technological innovation, as measured by resources devoted to research and devel-

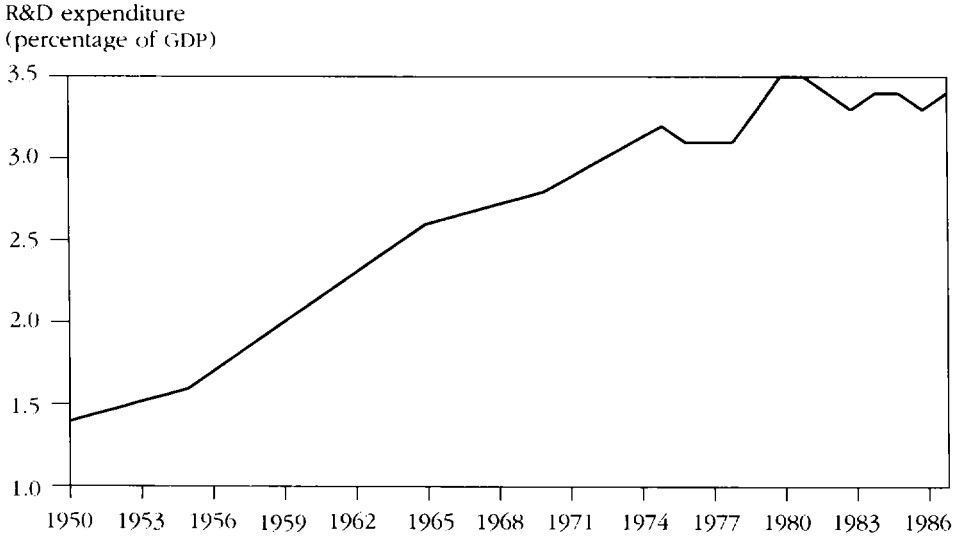
opment (R&D), significantly explains relative growth performance across economies (Coe and Helpman 1993; Lichtenberg 1992; Romer 1989).

Western estimates of the Soviet research effort, presented in figure 1, show R&D spending rising as a share of GDP. The R&D share is above the 2 to 3 percent of GDP in the leading industrialized economies. In 1967, about 1.5 percentage points of Soviet GDP was estimated to be for defense and space R&D (Bergson 1983). The share of defense and space R&D in total R&D is believed to have fallen in the period 1959–84 (Acland-Hood 1987), implying an even steeper rise in civilian R&D. It is well known that the lack of market incentives made the allocation of Soviet R&D inefficient and inhibited the diffusion of innovations. This would explain a low growth payoff for a given amount of R&D; it does not explain why that growth payoff would have fallen over time.

### *The Extensive Growth Hypothesis*

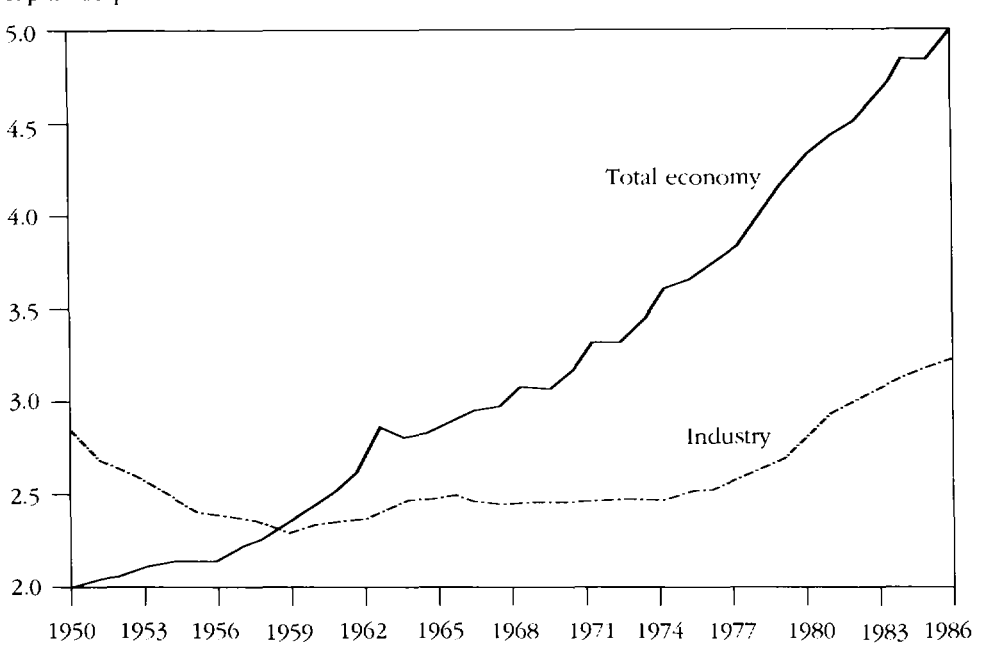
As noted in the introduction, the conventional hypothesis for the Soviet decline in growth is the pattern of extensive growth, defined by Ofer (1987) as growth mainly through input accumulation rather than through productivity growth. Ofer notes that a key feature of extensive growth is a rising capital-output ratio. Figure 2 shows the evolution of the capital-output ratios implied by the alternative data series for 1950–87. We begin the graphs in 1950 because of the extreme volatility of all of the capital-output series before 1950, which no doubt is related to shocks such as collectivization and World War II as well as questionable data quality. The total economy (GDP) series in figure 2 shows the capital-output ratio increasing two and a half times between 1950 and 1987.

Figure 1. *Research and Development Expenditures as a Percentage of GDP in the U.S.S.R., 1950–87*




*Source:* Joint Economic Committee (1990).

Figure 2. *Capital–Output Ratios in the U.S.S.R., 1950–86*




*Source:* Authors' calculations based on Western data.

The capital-output ratios industry first decline in the 1950s and then rise sharply after 1960. The capital-output ratio in the official series also rises steadily beginning at the end of the 1950s, more than doubling between 1958 and 1987. Khanin's data, by contrast with the other two series, show only a small increase in the capital-output ratio between the early 1950s and 1987.

In the neoclassical model, a rising capital-output ratio implies capital deepening during the transition to a higher steady state, but this capital deepening will sooner or later run into diminishing returns that will cause growth to slow or stop. With technical change, steady-state growth is feasible with a constant capital-output ratio. If the Soviet reliance on capital deepening is to explain its growth decline in relation to market economies, then market economies must have relatively constant capital-output ratios. Capital-output ratios in market economies have indeed long been thought to remain relatively stable, according to the famous Kaldor stylized fact (see, for example, the recent statement by Romer 1990).

However, recent research on capital accumulation in market economies contradicts the Kaldor stylized fact, shedding a new perspective on the Soviet experience of extensive growth. Appendix table A-1 lists the annual growth rates of

the capital-output ratios in a selection of recent growth-accounting studies and a few older ones. All studies agree that the capital-output ratio in the United States has remained remarkably constant, which perhaps accounts for the conventional wisdom that Kaldor's stylized fact holds. But several recent studies point to capital-output ratios rising at Soviet-style rates in Japan and in some of the East Asian newly industrializing economies (NICs), such as Korea (Young 1994b; Kim and Lau 1994; King and Levine 1994; Benhabib and Spiegel 1994; Nehru and Dhareshwar 1993).[6] These studies present a new fact for the literature on Soviet extensive growth to explain: why did the extensive growth strategy lead to eventual stagnation in the U.S.S.R., whereas the same strategy sustained rapid growth in Japan and Korea?

Moreover, King and Levine (1994), Benhabib and Spiegel (1994), and Nehru and Dhareshwar (1993) show that rising capital-output ratios are a feature of growth for many economies. (See also Judson 1994, who shows the capital-output ratio rising systematically with income.) The three studies compute capital stocks for a large sample of economies, using a variety of data sources (such as Summers and Heston 1991 and World Bank data) and a variety of assumptions about initial capital stocks and depreciation rates. The three concur that rising capital-output ratios are by no means rare: the median annual rate of growth in the capital-output ratio in their respective samples is around 1 percent, and in a full quarter of the samples the annual rate of growth in the capital-output ratio is greater than 1.7 percent.[7] And it is not only developing economies that are shown to have rapid capital deepening. For example, the studies concur that capital-output ratios in Austria and France increased at more than 1.5 percent a year.

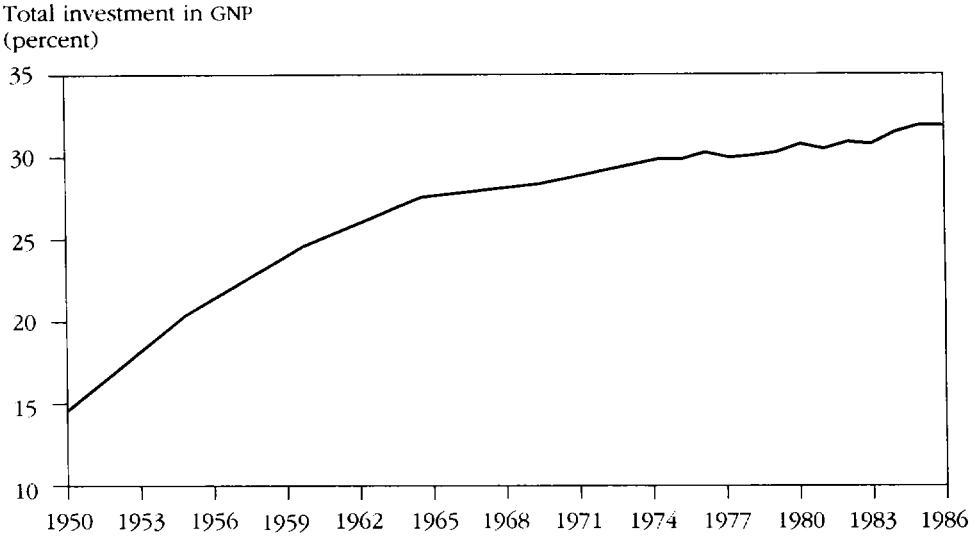
As either a cause or a consequence of falling output growth, the level of the Soviet capital-output (K-Y) ratio had become extreme by the 1980s. The K-Y ratio as measured by the Western GDP and total capital stock series was 4.9 in 1985, which is higher than any of the 1985 K-Y ratios in the Benhabib-Spiegel and King-Levine exercises. In the Nehru-Dhareshwar sample there are only four economies with a K-Y ratio above that of the U.S.S.R. in 1985, none of which seems especially relevant as a comparator—Guyana, Jamaica, Mozambique, and Zambia.

One other implication of the extensive growth model—also pointed out by Ofer (1987)—is that investment ratios have to rise over time if growth is to be maintained while the capital-output ratio rises. The Soviet investment share doubled between 1950 and 1975, as can be seen in the Western estimates presented in figure 3. After 1975 the investment share continued to increase, but more slowly.

6. The literature has less of a consensus on whether the East Asian economies have high TFP growth. But even those who argue for high productivity growth acknowledge rising capital-output ratios in East Asia. See, for example, World Bank (1993) and Pack and Page (1994).

7. For the two studies that use Summers and Heston data (Benhabib and Spiegel 1994 and King and Levine 1994), we omitted Africa from the sample because ratios of investment to GDP are implausibly extreme (both high and low) for the 1950s.

Figure 3. *Share of Total Investment in GNP in the U.S.S.R., 1950–86*




*Source:* Joint Economic Committee (1990).

How unusual is the doubling of the investment rate over a twenty-five-year period? In the Summers and Heston (1991) international data base for 1950–75, eight of fifty-two economies—most notably Japan and Taiwan (China)—had a doubling or more of investment rates.[8] When the sample period is shifted forward by ten years to expand the sample, six of seventy-two economies had a more than doubling of investment during 1960–85, among which fast-growing Korea and Singapore are of particular interest. Again, Soviet extensive growth was not so different from East Asian extensive growth. What was unusual was the fact that Soviet growth was declining while its investment was rising. The standby of Soviet industrialization—investment in machinery and equipment—also increased sharply as growth declined. The importance of machinery to growth has been emphasized by de Long and Summers (1991, 1992, 1993), but the Soviet data suggest that a high ratio of machinery investment to GNP is not sufficient to generate growth.

## II. Production Functions and Extensive Growth

If extensive growth explains the Soviet growth slowdown, we should see it reflected in a falling marginal return to capital accumulation. The alternative explanation for falling growth would be a slowdown in total factor productivity (TFP) growth for some reason other than excessive reliance on capital accumula-

8. We continue to exclude African economies from this and the following sample.

tion. In this section, we assess the consistency of the data with these alternative explanations.

### *Productivity Growth*

In table 4 we show TFP growth for the total economy and the industrial sector in the U.S.S.R. TFP growth is calculated assuming a Cobb-Douglas production function with labor's share equal to 0.6 and the share of capital equal to 0.4 (slightly above that used by Bergson (1979, 1983) and the CIA (1982), but within the conventional range for developing economies).[9] With the assumption of Cobb-Douglas production (unit elasticity of substitution between capital and labor), we see a strongly declining trend in TFP growth after the 1950s.

The most interesting aspect of table 4 is that the 1950s once more stand out as an exceptional period in Soviet growth. It is especially striking that even conservative Western estimates for the industrial sector imply productivity growth in that decade of more than 4 percent a year. The Western GNP data give a more pessimistic assessment of Soviet productivity performance, implying that productivity growth in the U.S.S.R. started being poor in 1960. By contrast, Khanin's data, which uniformly exhibit lower overall growth than Western GNP data, imply positive post-1950 productivity growth, a result of the lower rates of growth of capital in the Khanin series. There is only a moderate difference between the official and Western data on factor input growth, but Khanin shows substantially lower rates of growth of capital (table 1). The lower capital growth rates reflect Khanin's view that hidden inflation is as serious in capital goods industries as in consumer goods, a view shared by the "British school" of Hanson (1984), Nove (1981), and Wiles (1982).

How do the Soviet TFP growth rates look in a comparative context? We calculated TFP growth rates for the East Asian extensive growers assuming the same

Table 4. *Total Factor Productivity Growth in the Total Economy and in Industry in the U.S.S.R., 1928–87*
(average annual percent)

| *Period* | *Total economy* | *Industry* |
|---|---|---|
| 1928–39 | 0.6 | 2.4 |
| 1940–49 | 1.3 | 1.5 |
| 1950–59 | 2.8 | 4.6 |
| 1960–69 | 0.8 | 1.4 |
| 1970–79 | 0.1 | 1.8 |
| 1980–87 | –0.2 | –0.1 |

*Note:* TFP calculated as (growth of output per worker) – 0.4 (growth of capital per worker). Growth rates are taken from table 1.
*Source:* Authors' calculations.

9. It has long been a stylized fact in the development literature that capital shares are higher in developing than in industrial economies (see, for example, De Gregorio's 1992 estimate that the capital share is between 0.4 and 0.55 for Latin America). Western estimates of Soviet per capita income suggest it was a developing rather than an industrial country. The Benhabib-Spiegel, Nehru-Dhareshwar, and King-Levine cross-economy data sets all assume a capital share of 0.4.

Cobb-Douglas production function with a capital share of 0.4 across economies, using the alternative estimates for East Asian accumulation and growth by the various authors. Table 5 shows how Soviet TFP growth rates compare with the East Asian cases in such a comparison. Soviet TFP growth, measured assuming a common production function, was worse than that of the East Asian extensive growers. There is a lively controversy about whether East Asian TFP growth in fact was above average or whether it was low, as argued by Young (1992, 1994a, 1994b). Young's estimates are lower than those shown in table 5 because he estimates a higher capital share, as discussed in section I. The important point for this section is that East Asia seemed to get more output growth and less diminishing returns out of its extensive growth strategy than the U.S.S.R. did.

How does Soviet TFP growth compare with TFP of other economies within the sample of all economies? To answer that question, we performed an exercise similar to our growth regression in section I. We used TFP estimates from Nehru and Dhareshwar (1993), who also assume a Cobb-Douglas production function with a capital share of 0.4, for a sample of seventy-eight developing and industrial economies. As with per capita growth rates, TFP growth rates of small economies seem to have much higher variance than those of large economies.[10] Hence we use weighted least squares again with total GDP in 1960 as the weights for a regression of TFP growth rates in 1960–90 on a constant and a dummy for the Soviet observation. The results are as follows:

(5) Annual TFP growth 1960–90 = 1.32 − 1.12 Soviet dummy
(0.065)(0.214)

least squares weighted by total GDP
78 observations
$R^2$ (weighted) = 0.83
standard errors in parentheses

Assuming the same Cobb-Douglas production function for all economies, Soviet TFP growth is a little more than 1 percentage point below the GDP-weighted average of productivity growth in all economies, and the difference is significant.

### *The Low Elasticity of Substitution Hypothesis*

Table 4 suggests that a slowdown in TFP growth corresponded to the overall growth slowdown. Such a productivity slowdown is puzzling when, as we have seen, R&D expenditures were increasing. However, the TFP calculations in table 4 imposed the Cobb-Douglas production function with a unit elasticity of substitution between capital and labor. Following the pioneering work of Weitzman (1970, 1983) and later contributions (including Desai 1976 and 1987, and Bergson 1979), we investigated whether a constant elasticity of substitution (CES) func-

10. Perkins and Syrquin (1989) note some of the extreme negative TFP observations for small economies; these extreme negative values give a smaller mean TFP growth for small economies than for large ones. We find only a larger variance for small economies, not a significantly smaller mean.

Table 5. *Productivity Growth Assuming the Same Production Function Structure in the U.S.S.R. and in East Asia, Based on Different Data Sources, 1950–88*
(average annual percent)

| *Data source and economy* | *Period* | *Change in capital-output ratio* | *Productivity growth* |
|---|---|---|---|
| *Western data* | | | |
| U.S.S.R. | 1950–87 | 2.5 | 0.8 |
| | 1960–87 | 2.6 | 0.2 |
| | 1970–87 | 3.0 | –0.3 |
| *Young (1994b)* | | | |
| Singapore | 1970–90 | 2.8 | 0.8 |
| Korea, Rep. of (excluding agriculture) | 1966–90 | 3.6 | 1.5 |
| Taiwan, China (excluding agriculture) | 1966–90 | 2.6 | 2.3 |
| *Kim and Lau (1994)* | | | |
| Singapore | 1964–90 | 1.4 | 2.1 |
| Korea, Rep. of | 1960–90 | 3.5 | 1.7 |
| Taiwan, China | 1953–90 | 3.1 | 2.2 |
| Japan | 1957–90 | 3.2 | 1.9 |
| *Benhabib and Spiegel (1994)* | | | |
| Japan | 1965–85 | 2.6 | 1.9 |
| Korea, Rep. of | 1965–85 | 2.8 | 2.1 |
| Singapore | 1965–85 | 2.4 | 2.1 |
| Taiwan (China) | 1965–85 | 3.0 | 1.9 |
| *Nehru and Dhareshwar (1993)* | | | |
| Japan | 1950–90 | 2.7 | 1.7 |
| Korea, Rep. of | 1950–90 | 3.7 | 2.8 |
| *King and Levine (1994)* | | | |
| Japan | 1950–88 | 2.3 | 2.5 |
| Korea, Rep. of | 1950–88 | 3.0 | 2.0 |
| Singapore | 1950–88 | 2.9 | 2.8 |
| Taiwan, China | 1950–88 | 2.6 | 2.4 |

*Note:* Productivity growth is calculated assuming a constant Cobb-Douglas capital share of 0.4 across all economies, using data on output, capital stocks, and labor.

tion provides a better representation of the data than the Cobb-Douglas production function does.

Weitzman's basic finding was that a CES production function with a low elasticity of substitution of 0.4 fit the data better than the Cobb-Douglas and that the hypothesis that the elasticity of substitution was one could be rejected. With his specification, declining marginal product of capital, rather than declining TFP growth, explained the growth slowdown. This result has been hotly debated. Bergson (1983) criticized this result on the grounds that it implied implausibly high estimates of the marginal product of capital in earlier years. Desai (1987)

concurred with Weitzman's finding for aggregate industry but argued that Cobb-Douglas was an adequate representation for some branches of industry.

In general, there will be multiple specifications of production functions consistent with the data. A classic article by Diamond, McFadden, and Rodriguez (1967) showed that it is in general impossible to identify separately a time-varying elasticity of substitution parameter and the bias of technical change (neutral, labor-augmenting, or capital-augmenting). However, following Weitzman, we examined whether some plausible alternatives fit the data better than others. We first identified the substitution parameter by presuming it to be constant over time. We also assumed technical change to be neutral, but we did allow it to vary over time. We then tested the alternative of a low elasticity of substitution against a hypothesis of falling TFP growth.

Estimation of production functions in industrial economies is the subject of a large literature. The usual method is to estimate parameters of factor demands derived from the cost function, the dual of the production function (see Jorgenson 1983 for a survey). This is obviously inappropriate for a nonmarket economy, such as the U.S.S.R. Direct estimation of production functions is usually thought to be tainted by endogeneity of the factor supplies, particularly capital; we believed this would be much less of a problem in the nonmarket system of the U.S.S.R.

Table 6 shows elasticities of substitution estimated by nonlinear least squares, and recalculated TFP growth rates for 1950–87 (assuming Hicks-neutral technical progress) for subperiods with the CES form, where $t$ denotes time:

$$(6) \qquad \ln(Y/L) = c_1 t_{1950-59} + c_2 t_{1960-69} + c_3 t_{1970-79} + c_4 t_{1980-87} + [c_5/(1+c_5)] \ln[c_6 (K/L)^{(1+c_5)/c_5} + (1-c_6)] + c_7.$$

We find indeed that the Western data on Soviet output and capital growth per worker lend themselves to the CES form with low elasticities of substitution between capital and labor (significantly below one) and roughly constant rates of TFP growth.

We regard these results as very suggestive, but we note again the caveats. First, these results are conditional on assumptions that we need to identify the production function. Second, the results were much less sharp when we used the entire 1928–87 sample, where, as indicated earlier, the data before 1950 are volatile. Finally, when we tried similar estimates on Khanin's and official data, we found some support in the official data for the low-elasticity hypothesis but none in Khanin's data.

The results with the Khanin data are intriguing because they support a story of unit elasticity of substitution, a high capital share (higher than was imposed for table 4), and poor (although not strongly declining) productivity growth. According to Khanin's data, growth declined mainly because capital growth slowed (see table 1 again). Given Bergson's (1987a, 1991a) criticisms and the

Table 6. *Elasticities of Substitution and Total Factor Productivity Growth with Estimated Constant Elasticity of Substitution Functions, U.S.S.R, 1950–87*

| *Variable or statistic* | *Total economy* | *Industrial sector* |
|---|---|---|
| Elasticity of substitution ($c_5$) | −0.37 | −0.21 |
| | (−9.8) | (−3.1) |
| TFP growth | | |
| 1950–59 ($c_1$) | 1.09 | −0.19 |
| | (3.4) | (−2.3) |
| 1960–69 ($c_2$) | 1.10 | −0.25 |
| | (3.1) | (−2.9) |
| 1970–79 ($c_3$) | 1.16 | −0.21 |
| | (3.2) | (−2.3) |
| 1980–87 ($c_4$) | 1.09 | −0.22 |
| | (3.2) | (−2.3) |
| Intercept ($c_7$) | −0.82 | −0.95 |
| | (−9.8) | (−16.7) |
| Capital share parameter ($c_6$) | 0.96 | 0.47 |
| | (71.2) | (6.2) |
| Autocorrelation coefficient | n.a. | 0.81 |
| | | (15.8) |
| $R^2$ | 0.9987 | 0.9995 |
| Durbin-Watson | 1.9228 | 1.2968 |

n.a. Not applicable
*Note:* Estimates are based on Western data. *t*-statistics are in parentheses.
*Source:* Authors' calculations.

limited information about the methodology behind Khanin's data, these differing results can only point to the need for further research into Khanin's approach to see whether his work represents a valid criticism of the Western estimates. For the moment, we are forced to regard conclusions based on Khanin's data as unproven.

We would have liked to examine the implications of the "British school" of Hanson (1984), Nove (1981), and Wiles (1982), who made somewhat similar claims to Khanin's. However, we could not do so because those researchers did not provide alternative time series for output and capital. Note that a lower estimate for the growth rate of capital over the entire period, as implied by these authors' arguments, would imply higher TFP growth but does not necessarily imply a higher estimated elasticity of substitution.

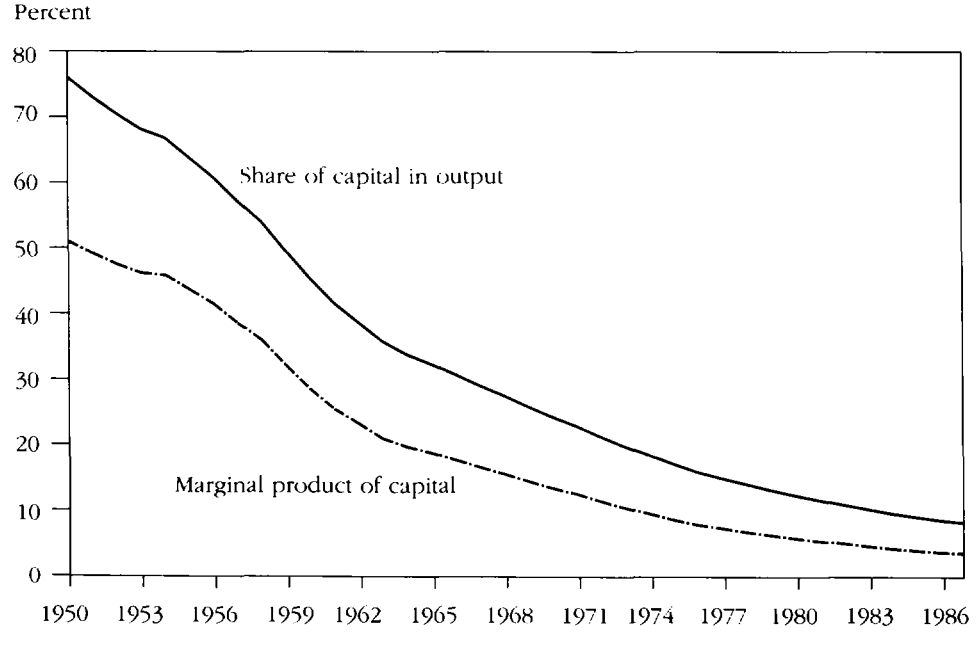
For our purposes here, the most striking feature of table 6 is that the implied rates of TFP growth show no significant decline between the 1950s and 1980s. Thus, once we free up the functional form of the production function, we find no evidence for a slowdown in TFP growth. In table 4, both extensive growth and declining productivity growth account for the overall fall in growth; in table 6, extensive growth—diminishing returns to capital—accounts entirely for the growth slowdown. Of course, as Ofer (1987) points out, extensive growth and low productivity growth are not necessarily independent of each other. The Soviets may have felt compelled to pursue extensive growth because productivity

growth was low; there may have been a mechanism in the Soviet system by which rapid growth in capital leads to poor TFP performance.

The Western GDP estimates yield a constant TFP growth rate of 1 percent a year, in contrast to the sharply falling TFP growth implied by the Cobb-Douglas estimates in table 4 for the 1970s and 1980s; the industrial TFP growth is slightly negative and also constant. Because the U.S.S.R. was spending increasing R&D resources and borrowing Western technology in the 1970s and 1980s, we find a constant rate of TFP growth more plausible than a falling rate. We cannot see a compelling story for the sort of worsening efficiency in resource use and technology adaptation (in the sense of Nishimizu and Page 1982) that would be required to explain falling TFP growth.

In figure 4 we examine a second implication of the estimates in table 6: these are estimates of the "share of capital" implied by the production function parameters for the Western estimate of GDP for 1950–87, under the hypothetical assumption of marginal productivity pricing. When the elasticity of substitution is less than one and the capital-output share is rising, the share of capital will fall over time. Diminishing returns will not be very severe with a high capital share, but will be severe with a low capital share. In figure 4 the share of capital falls

Figure 4. *Share of Capital in Output and Marginal Product of Capital for the Total Economy of the U.S.S.R., 1950–87*




*Note:* Calculated from Western estimates of capital stocks, employment, and value added for industry and for the whole economy.
*Source:* Authors calculations.

sharply for the economywide production function (a similar graph obtains for the industry estimates). Figure 4 also shows closely related data on the marginal product of capital implied by the CES estimates. The Western GNP data imply high rates of return to capital in the early 1950s, then the rates declined to about 3 percent in 1987.

A CES function with a high capital share acts much like a linear function of capital, so that the marginal product of capital can stay high for as long as the capital share is high. With a very capital-intensive production of goods, including capital goods, the Soviets were close for a while to the model of growth through rapid reproduction of capital described by Feldman in the 1920s as using machines to make more machines (in Domar 1957). Rebelo (1991) shows formally that constant returns to reproducible factors in the capital goods sector are sufficient to generate a constant, sustained rate of growth even without TFP growth.

However, as the capital share begins to fall, the marginal product will begin to decline. The decline can be precipitous when the elasticity of substitution is particularly low. Although we find the extreme values of the marginal product of capital and capital's share in figure 4 surprising, they do not logically rule out the CES form; the capital-labor ratio in a nonmarket economy could be driven to levels that would not be observed in a market economy.

Indeed, the data presented in figure 4 suggest that a market economy could not have gone through the growth process of the Soviet economy between 1950 and 1987. The low wage shares in the early period would probably have prevented any but a subsistence wage equilibrium in those periods. The low marginal product of capital by the mid-1980s would have been inconsistent with equilibrium and would have meant that investment in industry and the capital-labor ratio would have been lower. What would have happened in the early years if there had been a market economy? One possibility is that different technologies would have been adopted. Similarly, in the later period there may well have been other technologies available—but less amenable to central planning—that yielded a positive return to capital. It is also possible that if the extensive growth route had been closed off in a market economy, there would have been more incentive for Soviet entrepreneurs to attempt to improve technology.

### *International Comparisons of Elasticities of Substitution*

As we did with the earlier regression estimates, we put the results on the Soviet elasticity of substitution and TFP growth in an international context. We have fragmentary evidence that the low elasticity for the U.S.S.R. is not unique among planned economies. Rusek (1989) reports the remarkably low elasticity of 0.10 for Czechoslovak industry. Sapir (1980) estimates an elasticity of 0.13 for Yugoslav manufacturing.

The characterization of the Soviet (and Soviet-type) data by a low elasticity of substitution and constant TFP growth provides a natural way to reconcile the lack of success of the Soviet extensive growth strategy with the high payoffs

from capital deepening in Japan, Korea, and other market economies. If East Asian economies have higher elasticities of substitution, then this would explain the much weaker force of diminishing returns to capital deepening. De la Grandville (1989: 479) insightfully speculates "whether part of the explanation of miracle growth in Japan or South Asian countries lies not only in a high saving rate but also in a high elasticity of substitution between factors in their industrial sectors."

Yuhn (1991) tested De la Grandville's conjecture for Korean manufacturing and found an elasticity of substitution near unity. Bairam (1989) presents evidence that Japan's pre–World War II development was characterized by an elasticity of substitution above unity. Wang (1995) argues for an elasticity of substitution above two in manufacturing in Taiwan (China).

High elasticities of substitution would have made it easier for East Asian economies to absorb rapid capital accumulation without severe diminishing returns. Young (1994b) assumes a unit elasticity of substitution in the East Asian tigers he studies and estimates relatively high capital shares (particularly in Singapore). He thus reconciles high and constant output growth in East Asia with rapid increases in capital-output ratios and less-than-spectacular productivity growth (that becomes even negative TFP growth in Singapore). Similarly, Kim and Lau (1994) cannot reject a zero rate of TFP growth in the East Asian NICs, but have an output elasticity of capital beginning at about 0.6 (and then slowly declining to about 0.4).

For market economies in general, a recent study estimating the elasticity parameter from the convergence behavior of the cross-sectional national per capita income data argues that the elasticity of substitution is slightly above one (Chua 1993). The classic studies by Berndt and Wood (1975) and Hudson and Jorgenson (1978) report elasticities of substitution slightly above one for the United States (an estimate that is somewhat higher than in earlier studies).

There has also been a large literature on elasticities of substitution in developing economies. The early focus of this literature was the hypothesis that less-developed economies had fixed-proportions technology (zero elasticity of substitution). This hypothesis has been soundly rejected by the literature. There is little consensus on whether the elasticities of substitution are at or significantly below unity. White's (1978: 33) summary of this literature was that "the estimates tend to clump between 0.5 and 1.2." See also the survey by Morawetz (1976), which reports a similarly wide range of estimates. Elasticities of substitution around or above unity have been found in places as disparate as Brazilian manufacturing (Tyler 1974), Pakistani manufacturing (Mahmood 1992), Venezuelan manufacturing (Sines 1979), Colombian agriculture (Thirsk 1974), Rhodesian GDP (Muzondo 1978), Egyptian manufacturing (Girgis 1974), Puerto Rican industry (Reynolds and Gregory 1965), Philippine manufacturing (Williamson 1971), and Ghanaian industry (Roemer 1975).

Our results in a comparative context suggest the plausible story that diminishing returns to extensive growth were much sharper in the U.S.S.R. than in

market economies because the substitutability of capital for labor was abnormally low. In the concluding section, we will speculate why substitutability may have been low in a planned economy.

*Combining Regression Evidence with Production Function Evidence*

As a final exercise, we inserted the other apparent correlate of declining growth—defense spending—into our production function estimates (we took the midpoint of the Brada and Graves (1988) estimates in table 3 spliced together with the Steinberg (1987, 1990) estimate for 1985–87). Specifically, we allowed the Hicks-neutral rate of technological progress to depend linearly on the share of defense spending in GDP in the production function estimated with Western GDP and capital stock data:

$$\ln(Y/L) = c_0 + c_1 t + c_2 t(\text{defense spending}) + [c_5/(1 + c_5)] \ln [c_6(K/L)^{(1+c_5)/c_5} + (1 - c_6)]. \tag{7}$$

The results for the nondefense parameters are virtually identical to the estimates shown in table 6 and are not reported here. We find that defense spending does indeed have a significant and negative effect on the rate of increase in the total productivity term in the production function. However, the effect is not very quantitatively important: every additional 1 percent of GDP spent on defense lowered productivity growth by 0.07 percent. The increase during 1960–87 of 2.2 percentage points in the defense share thus would have lowered growth by 0.15 percentage points. Moreover, the parameters of the CES function are virtually unchanged from our earlier regression so the low substitutability, diminishing returns story still holds. We also tried equipment investment and R&D spending as independent influences on the technical progress term, but both gave insignificant results.

How do we reconcile our production function estimates with our earlier cross-sectional growth regression evidence using the Levine-Renelt specification? The Soviets' high capital-output ratio and low substitutability of capital for labor implies a lower derivative of growth with respect to the investment rate than in other economies with lower K-Y ratios and more substitutable capital for labor.

To see this, assume zero depreciation and labor growth for simplicity and define $y$ and $k$ as output and capital, respectively, per worker. Assume a CES function $y = A(\gamma k^{\rho} + 1 - \gamma)^{(1/\rho)}$. Growth will be given as a function of the investment ratio ($i/y = \Delta k/y$) as follows:

$$\Delta y/y = \gamma (i/y) [ (k/y) (\gamma + (1-\gamma)k^{-\rho}) ]^{-1}. \tag{8}$$

As is well known, a higher $k/y$ implies a lower marginal effect of the investment ratio on growth simply because a given investment rate translates into lower capital growth. With a unit elasticity of substitution ($\rho = 0$), this is the only way

that the level of capital influences the marginal effect of investment. With a less-than-unit elasticity of substitution ($\rho < 0$), higher capital has an even stronger negative effect on the coefficient on investment in a growth equation. Although obviously not the only explanation, this is consistent with the large negative residual for the U.S.S.R.—and increasingly negative residuals over time—in the cross-sectional regressions.

In the same vein, why were the Soviet "TFP" growth rates, calculated as (growth per worker) − 0.4(capital growth per worker), so poor compared with the TFP growth rates of other economies? If the true production function is CES with a low elasticity of substitution, then a Cobb-Douglas calculation of TFP will yield the following:

$$(9) \qquad \Delta y/y - \gamma\,\Delta k/k = \gamma\,[(\gamma + (1-\gamma)k^{-\rho})^{-1} - 1]\,\Delta k/k + \Delta A/A.$$

If the elasticity of substitution is indeed unity ($\rho = 0$), then the Cobb-Douglas calculation gives the "true" technical change $\Delta A/A$. However, if the elasticity of substitution is less than unity ($\rho < 0$), measured Cobb-Douglas "TFP growth" will be lower, the higher the stock of capital per worker is. The Soviets' high capital stock and low elasticity of substitution would explain the poor Cobb-Douglas productivity growth compared with that of other economies.

We conclude from our reexamination of the aggregate data that—subject to admittedly a priori restrictions on the form of the production function—the original Weitzman (1970) story of Soviet growth deceleration provides a description of the final decades of the U.S.S.R. as well. With neutral technical change and a constant—and low—elasticity of substitution, the decline in Soviet growth can be explained by diminishing returns to capital accumulation alone. This helps explain why the average growth performance was poor when we take into account the rapid capital growth and high education levels. Our comparative approach illuminates further the extensive growth hypothesis of the literature on Soviet growth, and the low substitutability of capital for labor in the U.S.S.R. explains why the outcome of Soviet extensive growth was so different from that of East Asian extensive growth.

## III. Conclusions and Directions for Further Research

One natural question to ask is why Soviet capital-labor substitution would have been more difficult than in market economies such as Japan and Korea and whether this difficulty was related to the Soviets' planned economic system. Recent work on models of endogenous economic growth stresses the notion of a broad concept of capital, including human capital, organizational capital, and the stock of knowledge, which can substitute easily for raw labor and perhaps replace it altogether (Rebelo 1991; Jones and Manuelli 1990; Parente and Prescott 1991). Other recent models of endogenous growth stress the incredible diversity of capital goods and intermediate inputs that are successively explored by mar-

ket economies (see the presentation of such models in Barro and Sala-i-Martin 1995, chapter 6). A broad variety of capital goods would be more substitutable for labor than a narrow range of capital goods. Could one explanation for the Soviets' substitution problems be that, under an autocratically directed economic system, they accumulated a narrow rather than a broad range of capital goods because only certain types of capital were amenable to central allocation? Was substitutability low because they were missing such market-oriented types of physical and human capital as entrepreneurial skills, marketing and distributional skills, and information-intensive physical and human capital (because of the restrictions on information flows)? Do other economies—especially developing ones—with highly state-controlled economic systems also have low substitutability between capital and labor?

Although diminishing returns are weakened if the elasticity of substitution is high, they still apply. Could Soviet-style growth declines await other economies that rely too much on extensive growth? After all, the U.S.S.R. had its period of rapid growth from the 1930s through 1950s, when it appeared to be following a linear output-capital production function, as we have shown. The Soviet experience can be read as a particularly extreme dramatization of the long-run consequences of extensive growth. Weitzman (1990) describes Soviet growth (as analyzed by Ofer 1990) as the best application of the Solow (1956) neoclassical model ever seen. Krugman (1994) has an entertaining description of the unanticipated Soviet growth decline as a metaphor for the alleged future growth slowdown in East Asia. Of course, if the story of low elasticity in the U.S.S.R. is correct, then the metaphor is highly exaggerated.

Or are there other factors that stave off diminishing returns to capital accumulation in East Asia and other market-oriented extensive growers? Rapid growth in human capital would mitigate diminishing returns to physical capital in East Asia. However, we have seen that educational investments were also high in the U.S.S.R. Also, the fact that capital-output ratios rise implies that capital is growing faster than other factors, assuming constant returns to scale. Note that an elasticity of substitution above unity between capital and labor limits diminishing returns asymptotically such that sustained growth through capital accumulation is possible indefinitely. This is an important area for further research.

Our results with the U.S.S.R. in the international cross-sectional growth and productivity regressions suggest that the planned economic system itself was disastrous for long-run economic growth in the U.S.S.R. Although this point may now seem obvious, it was not so apparent in the halcyon days of the 1950s, when the Soviet example was often cited as support for the neoclassical model's prediction that distortions do not have steady-state growth effects. Economic systems with low substitutability may deceptively generate rapid growth with high investment, only to stagnate after some time. Because a heavy degree of planning and government intervention still exist in many developing economies, the eventual fate of the Soviet economic system continues to be of interest.

Table A-1. *Trends in Capital-Output Ratios in Growth-Accounting Studies*

| *Source and economy* | *Period* | *Annual percentage change in capital-output ratio* |
|---|---|---|
| *Western data (this article)* | | |
| U.S.S.R. | 1950–87 | 2.53 |
| *Maddison (1989)* | | |
| Argentina | 1950–84 | 0.61 |
| Brazil | 1950–84 | 0.86 |
| Chile | 1950–84 | –0.22 |
| China | 1950–84 | 2.48 |
| France | 1950–84 | –0.45 |
| Germany, Fed. Rep. of | 1950–84 | 0.16 |
| India | 1950–84 | 1.54 |
| Japan | 1950–84 | –0.91 |
| Korea, Rep. of | 1950–84 | 0.03 |
| Mexico | 1950–84 | 0.50 |
| Taiwan, China | 1950–84 | –0.34 |
| United Kingdom | 1950–84 | 0.62 |
| United States | 1950–84 | –0.07 |
| U.S.S.R. | 1950–84 | 3.75 |
| *Young (1994b)* | | |
| Hong Kong | 1966–91 | 0.84 |
| Singapore | 1970–90 | 2.79 |
| Korea, Rep. of (excluding agriculture) | 1966–90 | 3.62 |
| Taiwan, China (excluding agriculture) | 1966–90 | 2.55 |
| *Kim and Lau (1994)* | | |
| France | 1957–90 | 0.68 |
| Germany, Fed. Rep. of | 1960–90 | 1.16 |
| Hong Kong | 1966–90 | 1.11 |
| Japan | 1957–90 | 3.19 |
| Singapore | 1964–90 | 1.38 |
| Korea, Rep. of | 1960–90 | 3.50 |
| Taiwan, China | 1953–90 | 3.13 |
| United Kingdom | 1957–90 | 0.68 |
| United States | 1948–90 | –0.19 |
| *Elias (1992)* | | |
| Argentina | 1950–80 | 0.39 |
| Brazil | 1950–80 | –0.54 |
| Chile | 1950–80 | –0.39 |
| Colombia | 1950–80 | –0.79 |
| Mexico | 1950–80 | 0.44 |
| Peru | 1950–80 | 1.22 |
| Venezuela | 1950–80 | 0.75 |

| *Source and economy* | *Period* | *Annual percentage change in capital-output ratio* |
|---|---|---|
| *Chenery, Robinson, and Syrquin (1986)* | | |
| Canada | 1947–73 | 0.66 |
| France | 1950–73 | 0.13 |
| Germany, Fed. Rep. of | 1950–73 | 0.33 |
| Italy | 1952–73 | –0.63 |
| Netherlands | 1951–73 | 0.17 |
| United Kingdom | 1949–73 | 0.69 |
| United States | 1949–73 | 0.00 |
| *Benhabib and Spiegel (1994)* | | |
| Hong Kong | 1965–85 | –0.20 |
| Japan | 1965–85 | 2.56 |
| Korea, Rep. of | 1965–85 | 2.78 |
| Singapore | 1965–85 | 2.41 |
| Taiwan, China | 1965–85 | 2.97 |
| United States | 1965–85 | 0.63 |
| Percentile of sample[a] | | |
| 75th | 1965–85 | 1.72 |
| 50th | 1965–85 | 0.80 |
| 25th | 1965–85 | 0.21 |
| *Nehru and Dhareshwar (1993)* | | |
| Japan | 1950–90 | 2.70 |
| Korea, Rep. of | 1950–90 | 3.70 |
| United States | 1950–90 | 0.20 |
| Percentile of sample[b] | | |
| 75th | 1950–90 | 1.84 |
| 50th | 1950–90 | 1.06 |
| 25th | 1950–90 | 0.38 |
| *King and Levine (1994)* | | |
| Hong Kong | 1950–88 | –0.80 |
| Japan | 1950–88 | 2.33 |
| Korea, Rep. of | 1950–88 | 3.05 |
| Singapore | 1950–88 | 2.94 |
| Taiwan, China | 1950–88 | 2.63 |
| United States | 1950–88 | 0.40 |
| Percentile of sample[c] | | |
| 75th | 1950–88 | 1.69 |
| 50th | 1950–88 | 0.95 |
| 25th | 1950–88 | 0.23 |

a. The sample includes seventy-seven countries, excluding Africa.
b. The sample includes seventy-two countries.
c. The sample includes seventy-four countries, excluding Africa.

## REFERENCES

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

Acland-Hood, Mary. 1987. "Estimating Soviet Military R&D Spending." In Carl G. Jacobsen, ed., *The Soviet Defence Enigma.* Oxford: Oxford University Press.

Bairam, Erkin. 1989. "Learning-by-Doing: Variable Elasticity of Substitution and Economic Growth in Japan, 1878–1939." *Journal of Development Studies* 25(3):344–53.

Banerjee, Abhijit, and Michael Spagat. 1991. "Productivity Paralysis and the Complexity Problem: Why Do Centrally Planned Economies Become Prematurely Gray?" *Journal of Comparative Economics* 15(December):646–60.

Barro, Robert J., and Xavier Sala-i-Martin. 1995. *Economic Growth.* New York: McGraw-Hill.

Benhabib, Jess, and Mark M. Spiegel. 1994. "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data." *Journal of Monetary Economics* 34(October):143–73.

Bergson, Abram. 1961. *The Real National Income of Soviet Russia since 1928.* Cambridge, Mass.: Harvard University Press.

———. 1979. "Notes on the Production Function in Soviet Postwar Industrial Growth." *Journal of Comparative Economics* 3(2, June):116–26.

———. 1983. "Technological Progress." In Abram Bergson and Herbert Levine, eds., *The Soviet Economy: Toward the Year 2000.* London: Allen & Unwin.

———. 1987a. "On Soviet Real Investment Growth." *Soviet Studies* 39 (3, July):406–24.

———. 1987b. "Comparative Productivity: The U.S.S.R., Eastern Europe, and the West." *American Economic Review* 77(3):342–57.

———. 1991a. "Real National Income Measurement in Soviet Perspective." Paper prepared for the Bureau of Labor Statistics (BLS) and Eurostat Conference on Economic Statistics for Economies in Transition. Washington D.C. February. Bureau of Labor Statistics, Washington, D.C. Processed.

———. 1991b. "The U.S.S.R. before the Fall: How Poor and Why?" *Journal of Economic Perspectives* 5(4, Fall):29–44.

Berliner, Joseph S. 1976. *The Innovation Decision in Soviet Industry.* Cambridge, Mass.: MIT Press.

Berndt, Ernst, and David O. Wood. 1975. "Technology, Prices, and the Derived Demand for Energy." *The Review of Economics and Statistics* 57(3):259–68.

Brada, Josef C., and Ronald L. Graves. 1988. "Slowdown in Soviet Defense Expenditures." *Southern Economic Journal* 54:969–84.

Chenery, Hollis, Sherman Robinson, and Moises Syrquin. 1986. *Industrialization and Growth.* New York: Oxford University Press.

Chua, Hak. 1993. "Factor Substitution and Slow Convergence." Economic Growth Center Discussion Paper. Yale University, Economic Growth Center, New Haven, Conn. Processed.

CIA (U.S. Central Intelligence Agency). 1982. *U.S.S.R.: Measures of Economic Growth and Development.* Washington, D.C.: U.S. Government Printing Office.

———. Various years. *Handbook of Economic Statistics.* Washington, D.C.: U.S. Government Printing Office.

Coe, David T., and Elhanan Helpman. 1993. "International R&D Spillovers." NBER Working Paper 4444. National Bureau of Economic Research, Cambridge, Mass. Processed.

De Gregorio, Jose. 1992. "Economic Growth in Latin America." *Journal of Development Economics* 39(1):59–85.

De la Grandville, Olivier. 1989. "In Quest of the Slutsky Diamond." *American Economic Review* 79(June):468–81.

de Long, J. Bradford, and Lawrence Summers. 1991. "Equipment, Investment, Relative Prices, and Economic Growth." *Quarterly Journal of Economics* 106(2, May):445–502.

———. 1992. "Equipment Investment and Economic Growth: How Robust is the Nexus?" *Brookings Papers on Economic Activity* no. 2:157–211.

———. 1993. "How Strongly Do Developing Countries Benefit from Equipment Investment?" *Journal of Monetary Economics* 32(3):395–415.

Desai, Padma. 1976. "The Production Function and Technical Change in Postwar Soviet Industry: A Reexamination." *American Economic Review* 66(3, June):372–81.

———. 1987. *The Soviet Economy: Problems and Prospects.* Oxford: Basil Blackwell.

Diamond, Peter, Daniel McFadden, and Miguel Rodriguez. 1978. "Measurement of the Elasticity of Substitution and Bias of Technical Change." In Melvyn A. Fuss and Daniel McFadden, eds., *Production Economics: A Dual Approach to Theory and Applications.* Amsterdam: North-Holland.

Domar, Evsey. 1957. "A Soviet Model of Growth." In *Essays in the Theory of Economic Growth.* New York: Oxford University Press.

Easterly, William, Michael Kremer, Lant Pritchett, and Lawrence H. Summers. 1993. "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks." *Journal of Monetary Economics* 32:459–83.

Elias, Victor J. 1992. *Sources of Growth: A Study of Seven Latin American Economies.* San Francisco: ICS Press.

Ericson, Richard E. 1990. "The Soviet Statistical Debate: Khanin vs. Tssu." In H. S. Rowen and Charles Wolf, Jr., eds., *The Impoverished Superpower.* San Francisco: ICS Press.

Feschbach, Murray. 1983. "Population and Labor Force." In Abram Bergson and H. S. Levine, eds., *The Soviet Economy: Toward the Year 2000.* Boston: Allen & Unwin.

Fischer, Stanley. 1992. "Russia and the Soviet Union: Then and Now." In Olivier J. Blanchard, Kenneth A. Froot, and Jeffrey D. Sachs, eds., *The Transition in Eastern Europe.* Vol. 1. Chicago: The University of Chicago Press.

Girgis, Maurice. 1974. "Aggregation and Misspecification Biases in Estimates of Factor Elasticity of Substitution: The Case of Egypt." *Weltwirtschaftliches Archiv* 110(1):114–47.

Gomulka, Stanislaw, and Mark Schaffer. 1991. "A New Method of Long-Run Growth Accounting, with Applications to the Soviet Economy, 1928–87, and the U.S. Economy, 1949–78. Centre for Economic Performance Discussion Paper 14. London School of Economics and Political Science, London. Processed.

Hanson, Phillip. 1984. "New Evidence about Soviet Investment." *Radio Liberty Research* 13 September.

Hewitt, Daniel. 1993. "Military Expenditures, 1972–1990: The Reasons behind the Post-1985 Fall in World Military Spending." IMF WP/93/18. International Monetary Fund, Fiscal Affairs Department, Washington, D.C. Processed.

Hudson, Edward A., and Dale W. Jorgenson. 1978. "Energy Policy and U.S. Economic Growth." *American Economic Review* 68(2):118–23.

IMF (International Monetary Fund), World Bank, Organisation for Economic Co-operation and Development, and European Bank for Reconstruction and Development. 1991. *A Study of the Soviet Economy.* 3 vols. Washington, D.C.

Joint Economic Committee. 1990. *Measures of Soviet GNP.* Washington, D.C.: U.S. Government Printing Office.

Jones, Larry E., and Rodolfo Manuelli. 1990. "A Convex Model of Equilibrium Growth: Theory and Policy Implications." *Journal of Political Economy* 98(October):1008–38.

Jorgenson, Dale W. 1983. "Econometric Methods for Modeling Producer Behavior." In Zvi Griliches and Michael Intriligator, eds., *Handbook of Econometrics.* Vol. 3. Amsterdam: North-Holland.

Judson, Ruth. 1994. "Do Low Human Capital Coefficients Make Sense? A Puzzle and Some Answers." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass. Processed.

Kellogg, Robert L. 1989. "Inflation in Soviet Investment and Capital Stock Data and Its Impact on Measurement of Total Factor Productivity." Central Intelligence Agency, Office of Soviet Analysis, Washington, D.C. Processed.

Khanin, G. I. 1988. "Ekonomicheskii Rost: Al'ternativnaia Otsenka." *Kommunist* no. 17.

Kim, Jong-Il, and Lawrence J. Lau. 1994. "The Sources of Economic Growth of the East Asian Newly Industrialized Countries." *Journal of the Japanese and International Economies* 8(September):235–71.

King, Robert, and Ross Levine. 1994. "Capital Fundamentalism, Economic Development, and Economic Growth." *Carnegie-Rochester Series on Public Policy* 40(June):259–92.

Krugman, Paul. 1994. "The Myth of Asia's Miracle." *Foreign Affairs* 73(6):62–78.

Landau, Daniel. 1993. "The Economic Impact of Military Expenditures." Policy Research Working Paper 1138. World Bank, Policy Research Department, Washington, D.C. Processed.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4, September): 942–63.

Lichtenberg, Frank R. 1992. "R&D Investment and International Productivity Differences." NBER Working Paper 4161. National Bureau of Economic Research, Cambridge, Mass. Processed.

Maddison, Angus. 1989. *The World Economy in the 20th Century.* Paris: Development Center of the Organisation for Economic Co-operation and Development.

Mahmood, Zafar. 1992. "Factor Price Shocks, Factor Substitution, and Their Implications for Policy." *International Economic Journal* 6(4):63–73.

Marer, Paul, Janos Avray, John O'Connor, Martin Schrenk, and Daniel Swanson. 1992. *Historically Planned Economies: A Guide to the Data.* World Bank: Washington, D.C.

Moorsteen, Richard H., and Raymond P. Powell. 1966. *The Soviet Capital Stock, 1928–1962*. Homewood, Ill.: Irwin.

Morawetz, David. 1976. "Elasticities of Substitution in Industry: What Do We Learn from Econometric Estimates?" *World Development* 4(1):11–15.

Muzondo, Timothy R. 1978. "Sectoral Elasticities of Substitution between Capital and Labour: Rhodesia, 1954–75." *South African Journal of Economics* 46(1):36–47.

Nehru, Vikram, and Ashok Dhareshwar. 1993. "A New Database on Physical Capital Stock: Sources, Methodology, and Results." *Revista de Analisis Económico* 8(1):37–59.

Nishimizu, Mieko, and John M. Page, Jr. 1982. "Total Factor Productivity Growth, Technological Growth, and Technical Efficiency Change: Dimensions of Productivity Change in Yugoslavia, 1965–78." *Economic Journal* 92(368):920–36.

Nove, Alec. 1981. "A Note on Growth, Investment, and Price Indices." *Soviet Studies* 23(1, January):143–45.

Ofer, Gur. 1987. "Soviet Economic Growth: 1928–85." *Journal of Economic Literature* 25 (4, December):1767–1833.

———. 1990. "Macroeconomic Issues of Soviet Reforms." In Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1990*. Cambridge, Mass.: MIT Press.

Pack, Howard, and John M. Page, Jr. 1994. "Accumulation, Exports, and Growth in the High-Performing Asian Economies." *Carnegie-Rochester Series on Public Policy*. 40(June):199–236.

Parente, Stephen L., and Edward C. Prescott. 1991. "Technology Adoption and Growth." NBER Working Paper 3733. National Bureau of Economic Research, Cambridge, Mass. Processed.

Perkins, Dwight W., and Moshe Syrquin. 1989. "Large Countries: The Influence of Size." In Hollis Chenery and T. N. Srinivasan, eds., *Handbook of Development Economics*. Vol. 2. Amsterdam: North Holland.

Powell, Raymond P. 1963. "Industrial Production." In Abram Bergson and Simon Kuznets, eds., *Economic Trends in the Soviet Union*. Cambridge, Mass.: Harvard University Press.

———. 1968. "The Soviet Capital Stock and Related Statistical Series for the War Years." Yale University, Economic Growth Center, New Haven, Conn. Processed.

Rebelo, Sergio. 1991. "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy* 99(3, June):500–21.

Reynolds, L. G., and Peter Gregory. 1965. "Wages, Productivity, and Industrialization in Puerto Rico." Yale University, Economic Growth Center, New Haven, Conn. Processed.

Roemer, Michael. 1975. "The Neoclassical Employment Model applied to Ghanian Manufacturing." *Journal of Development Studies* 11(2):75–92.

Romer, Paul. 1989. "What Determines the Rate of Growth and Technical Change?" PRE Working Paper 279. World Bank, Country Economics Department, Washington, D.C. Processed.

———. 1990. "Capital Accumulation in the Theory of Long-Run Growth." In Robert Barro, ed., *Modern Business Cycle Theory*. Cambridge, Mass.: Harvard University Press.

Rusek, Antonin. 1989. "Industrial Growth in Czechoslovakia, 1971–85: Total Factor Productivity and Capital-Labor Substitution." *Journal of Comparative Economics* 13(2):301–13.

Sapir, Andre. 1980. "Economic Growth and Factor Substitution: What Happened to the Yugoslav Miracle?" *The Economic Journal* 90:294–313.

Sines, Richard H. 1979. "Sectoral Elasticities of Substitution between Labor and Capital in Venezuelan Manufacturing: A Cross-Sectional Micro Analysis." *World Development* 7(1):79–82.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1, February):65–94.

Steinberg, Dmitri. 1987. "Estimating Total Soviet Military Expenditures: An Alternative Approach Based on Reconstructed Soviet National Accounts." In Carl G. Jacobsen, ed., *The Soviet Defence Enigma: Estimating Costs and Burden.* New York: Oxford University Press.

———. 1990. "Trends in Soviet Military Expenditure." *Soviet Studies* 42(4):675–99.

Summers, Robert, and Alan Heston. 1991. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–88." *Quarterly Journal of Economics* 106(May):327–68.

Thirsk, Wayne. 1974. "Factor Substitution in Colombian Agriculture." *American Journal of Agricultural Economics* 56(1):73–84.

Tyler, William G. 1974. "Labour Absorption with Import-Substituting Industrialization: An Examination of Elasticities of Substitution in the Brazilian Manufacturing Sector." *Oxford Economic Papers* 26(1):93–103.

UNESCO (United Nations Educational, Scientific, and Cultural Organization). 1975. *UNESCO Statistical Yearbook.* New York: United Nations.

Wang, Eric C. 1995. "Factor Substitution Approach to Testing the Time Variations in Production: The Case of Taiwan's Manufacturing." *Applied Economics* 27(January):107–16.

Weitzman, Martin L. 1970. "Soviet Postwar Economic Growth and Capital-Labor Substitution." *American Economic Review* 60(5, December):676–92.

———. 1983. "Industrial Production." In Abram Bergson and H. S. Levine, eds., *The Soviet Economy: Toward the Year 2000.* Boston: Allen & Unwin.

———. 1990. "Comment." In Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual: 1990.* Cambridge, Mass.: MIT Press.

White, Lawrence J. 1978. "The Evidence on Appropriate Factor Proportions for Manufacturing in Less Developed Countries: A Survey." *Economic Development and Cultural Change* 27(1):27–59.

Wiles, Peter. 1982. "Soviet Consumption and Investment Prices and the Meaningfulness of Real Investment." *Soviet Studies* 24(2, April):289–95.

Williamson, J. G. 1971. "Relative Price Changes, Adjustment Dynamics and Productivity Growth: The Case of Philippine Manufacturing." *Economic Development and Cultural Change* 19(4):507–26.

World Bank. 1993. *The East Asian Miracle.* New York: Oxford University Press.

Young, Alwyn. 1992. "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore." In Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1992.* Cambridge, Mass.: MIT Press.

———. 1994a. "Lessons from the East Asian NICs: A Contrarian View." *European Economic Review* 38(4)964–73.

———. 1994b. "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." NBER Working Paper 4680. National Bureau of Economic Research, Cambridge, Mass. Processed.

Yuhn, Ky Hyang. 1991. "Economic Growth, Technical Change Biases, and the Elasticity of Substitution: A Test of the De la Grandville Hypothesis." *Review of Economics and Statistics* 73(2):340–46.

THE WORLD BANK ECONOMIC REVIEW VOL. 9, NO. 3 373-395

# Contingent Valuation and Actual Behavior: Predicting Connections to New Water Systems in the State of Kerala, India

Charles C. Griffin, John Briscoe, Bhanwar Singh, Radhika Ramasubban, and Ramesh Bhatia

*In 1988, families in Kerala State in India were surveyed to ascertain their willingness to pay for household connections to a piped water supply system. In 1991 the families in these communities were surveyed again and their actual decisions recorded. This article explores the validity of the findings of the 1988 study on the basis of actual behavior. It looks at the question of benefit revelation: did people behave as they said they would? And it looks at the question of benefit transfer: did people in one site behave as they were predicted to behave, on the basis of the predictions of a behavioral model for a different site? The data were also used to analyze the policy relevance of behavioral modeling.*

The ability to put a value on environmental resources is a core problem in environmentally sustainable development in industrial countries (Carson and Mitchell 1993 and Mitchell and Carson 1989) and developing countries alike (Serageldin and Steer 1994). During the past twenty years there has been a vigorous and contentious debate about the relative merits of various approaches (Brookshire and others 1982; Arrow and others 1993). The "indirect approach" draws conclusions from actual behavior; the "direct approach," or the contingent valuation method, draws conclusions from responses to hypothetical questions. The "benefit-transfer" issue in environmental economics, which is concerned with transferring valuations from one population to estimate how a second population would value the same resource, further complicates the debate (Pearce 1993).

Charles C. Griffin and John Briscoe are with the Eastern Africa Department, Population and Human Resources Division, and with the Transport, Water, and Urban Development Department, Water and Sanitation Division, at the World Bank; Bhanwar Singh and Radhika Ramasubban are with the Centre for Social and Technological Change, Bombay, India; and Ramesh Bhatia is with the Institute of Economic Growth, Delhi University, India, and the International Institute for the Management of Irrigation (IIMI), Colombo, Sri Lanka. Financial support for this work was provided by the Danish International Development Agency, the United Nations Development Programme, the Swiss Development Corporation, the Norwegian Agency for Development, and the World Bank. The study benefited from the assistance and cooperation of the Kerala Water Authority. David Guilkey of the University of North Carolina assisted with the econometric analyses. The data are available from John Briscoe. The authors wish to thank Dale Whittington for helpful advice in revising the manuscript and three reviewers for their excellent comments.

These approaches have been applied to other areas, including water supply policy in developing countries (World Bank 1992). The fundamental issue in water supply policy is predicting the response of consumers to a service to which they have not previously had access or to characteristics that they have not previously experienced. Piped water supply, higher prices, and improved reliability are examples (see Briscoe and others 1990; World Bank Water Demand Research Team 1993).

Mirroring the debate in the environmental economics literature, two fundamental approaches are used to analyze this fundamental issue. The indirect approach involves observing actual behavior, modeling this behavior, and then deriving the willingness to pay for water connections from the value of time spent fetching water or from housing values (see North and Griffin 1993, for example). The direct (contingent valuation) approach, by contrast, involves taking a survey (through a carefully designed and administered questionnaire) of households' willingness to pay specified prices for hypothetical services.

Our first concern is with the validity of the direct approach in predicting actual behavior. In environmental economics, because contingent valuation methods are used to value public goods or environmental amenities, it is not possible to validate the hypothetical responses of the interviewed population through actual market behavior. For a water system, however, it is possible to test the results of a contingent valuation survey by comparing the responses given when the water system was hypothetical to the actual behavior once the water system becomes available.

Our second concern is to test the benefit-transfer hypothesis as it applies to direct valuation approaches. Contingent valuation methods are subject to hypothetical bias, strategic bias, and compliance bias. Hypothetical bias can be reduced if the sample is well aware of the nature of the good; strategic bias can be reduced if the sample has little or nothing to gain by undervaluing the good. Compliance bias can be reduced through careful development of the survey, training, and supervision of the fieldwork. These problems can be reduced, at least hypothetically, by using the benefit-transfer approach, in which the behavior of a group that already has the service is projected onto a second group to predict the second group's willingness to pay for the good or service in question. This approach may be of interest for three reasons: the second group has little knowledge of the good, the second group may behave strategically, and the technique requires little additional data collection. The contingent valuation survey in Kerala was carried out in pairs of communities that differed primarily in their having or not having a water system available. We can therefore test the accuracy of benefit-transfer predictions from the villages where water systems already existed in relation to what we will call "benefit-revelation" predictions based on the contingent valuation responses of the villages where no water system existed.

Our third concern is to test the relative accuracy of simple tabulations of willingness to pay from a contingent valuation survey and predictions from a behavioral model using the same data. Simple tabulations may provide all the

information we need to decide whether people will hook up to a new water system under a planned tariff structure. Behavioral models are also important, however, because they allow us to simulate the impacts of changes in policy variables, estimate price elasticities, quantify changes in welfare, and control for nonpolicy variables that also affect behavior. Yet such models may be misspecified, the functional form may be inappropriate, or the assumed distribution of the errors may be wrong. We are able in this analysis to test actual behavior against both simple tabulated responses and predicted responses based on a behavioral model.

In this article we analyze two surveys of willingness to pay for improved water services in Kerala State in India. The first survey was done in 1988; a follow-up survey of the same households was conducted in 1991. Section I recaps the findings of the 1988 study, and section II describes the 1991 follow-up survey. The article then explores the validity of the findings of the 1988 study on the basis of actual behavior. Section III looks at results for the scarce-water environment and section IV for the saline-water environment. Section V presents conclusions.

## I. Recap of the 1988 Study

The 1988 study of the willingness to pay for water in villages in northern areas of the Indian state of Kerala was part of a multicountry study of willingness to pay for water in rural areas of developing countries (World Bank Water Demand Research Team 1993). Singh and others (1993) reported the results of the 1988 contingent valuation study. The original study conducted contingent valuation surveys among families living in three types of traditional drinking water environments—abundant water, scarce water, and salt-water intrusion (the latter is referred to as saline water in text). Pairs of communities were selected within each of the three water environments. Each pair included a site-A community and a site-B community.

Households in the site-A communities had piped water service already available. Within the site-A communities, two types of households were surveyed: those who had already decided to connect at the existing connection costs and tariffs and those who had decided not to connect. Connectors were asked whether they would continue to connect, for a range of hypothetical tariffs higher than the current tariff. Nonconnectors were asked whether they would connect, for a range of hypothetical tariffs and connection costs.

Households in the site-B communities had no piped system but could expect to have one installed soon. Families were asked whether they would connect for a variety of hypothetical connection costs and monthly charges. All households in both sites in each type of water environment were asked about their willingness to pay if the reliability of the water system were improved.[1] Thus, the im-

1. Reliability was characterized as follows: "Now I would like you to tell me what you would do if the service through the piped water system was greatly improved. Imagine that the water supply was available every day for most of the day, that the flow in the taps was always good, and that the water was clean."

pact on the choice of connecting of three characteristics of water systems—connection charge, tariff, and reliability—was tested through the contingent valuation method.

In the original survey there were 1,150 households distributed approximately evenly across the three types of water environments, including the entire population of connectors in the three A sites, a sample of nonconnectors in the A sites, and a sample of potential connectors in the B sites. Table 1 provides some basic information about the survey sites and shows how willingness to pay varied across connectors and nonconnectors in the A sites and overall for the B sites.

The exact questions posed during the interviews and the econometric issues that had to be resolved are discussed in detail in Singh and others (1993: 1,932–35) and will not be reproduced here. Table 2 summarizes how well the 1988 survey followed best practice in designing contingent valuation surveys, or what has become known as the "Seven Pillars of NOAA" (National Oceanic and Atmospheric Administration) (Portney 1994). These rules are designed to overcome some of the known problems with the technique. Generally, the survey

Table 1. *Location and Types of Survey Sites, with Sample Size and Maximum Willingness to Pay, Kerala, India*

| Site characteristic | *A sites: improved water source available* | | *B sites: no improved water source available* |
|---|---|---|---|
| | Connectors | Nonconnectors | |
| *Water abundant* | | | |
| Location | Ezhuvathuruthy | Ezhuvathuruthy | Nannamukku |
| All households | 66 | 819 | 1,497 |
| Household sample | 66 | 100 | 200 |
| *Water scarce* | | | |
| Location | Elapully | Elapully | Elapully |
| All households | 86 | 723 | 876 |
| Household sample | 86 | 100 | 200 |
| *Water abundant but with saline intrusion* | | | |
| Location | Ezhuvathuruthy | Ezhuvathuruthy | Vallikkunnu |
| All households | 98 | 768 | 1,313 |
| Household sample | 98 | 100 | 200 |
| Total household sample | 250 | 300 | 600 |
| Average maximum monthly tariff bid (rupees) | 19.3 | 8.7 | 5.5 |
| Average maximum connection charge bid (rupees) | n.a. | 355 | 267 |
| Average maximum bid for improved service (rupees) | 25.0 | 9.7 | n.a. |

n.a. Not applicable because the bidding game was not conducted in that site.
*Note:* The exchange rate in 1988 was about 14 rupees per U.S. dollar.
*Source:* Authors' calculations.

Table 2. *Comparison of National Oceanic and Atmospheric Administration (NOAA) Guidelines for Contingent Valuation Surveys with Guidelines Used in the Kerala 1988 Survey*

| *NOAA guidelines* | *Procedures used in the Kerala survey* |
|---|---|
| 1. Interview in person rather than on the telephone. | Interviewed household head personally. |
| 2. Question about a future, hypothetical occurrence rather than a historical event. | Survey asked about willingness to pay for a new or improved water system or a change in tariff policy, not for an existing service. |
| 3. Referendum format in which respondent "votes" on a benefit with a known price (as opposed to open-ended questions). | Bidding game format used: interviewer suggested prices for monthly tariff or connection cost; respondent answered yes or no to each quote. |
| 4. Interviewer begins with a scenario accurately describing the benefit or program. | Interviewer described the exact nature of the good or service to be provided: existing quality of piped water services into the home at various monthly tariffs and connection charges, plus improved quality of service at various monthly tariffs. |
| 5. Survey reminds that payment for the new benefit reduces other consumption. | Survey had no reminders, but questions about other consumption and assets preceded the contingent valuation questions. |
| 6. Survey reminds that substitutes exist for the hypothetical benefit in question. | In-depth questioning about cost, distance, and other characteristics of the household's sources of water (and volume of consumption) preceded the contingent valuation questions. No specific reminder was given during the contingent valuation questions, but it was clear to the respondent that the context of the survey was general use of various sources of water. |
| 7. Follow-up questions to make sure respondent understands the choices made. | No follow-up questions, but interviewer evaluated the quality of response. There was a follow-up survey to ascertain actual behavior after the water system was put in place, and respondents were asked to explain divergence between hypothetical (1988) and actual behavior (1991). |

*Source:* Portney (1994) and authors' review of the Kerala survey and procedures.

meets or exceeds the desired specifications, but the true test is whether the hypothetical responses adequately predict subsequent behavior.

In the earlier analysis (Singh and others 1993), we estimated an econometric model that controlled for individual, household, and community characteristics that could affect the responses to the three policy variables in which we were interested: monthly tariff, connection cost, and reliability. The purpose of estimating the full model was to isolate the impact of the policy variables on the choice to connect and then to perform simulations of how changes in these

policy variables would affect the number of water connections demanded and, consequently, consumer welfare.

Table 3 contains the full model estimated in Singh and others (1993), along with coefficient estimates and *t*-statistics. Monthly tariff was an important determinant of whether respondents were willing to pay for yard connections. For each 1 percent increase in the monthly tariff, the probability of choosing a yard tap fell by almost 1.5 percent. However, connection cost was an even greater impediment limiting connections, probably because of the very high implicit interest rates prevailing. Whereas a ten-rupee increase in the mean monthly tariff quoted in the bidding game would cause connection probabilities to fall by approximately 27 percent, a ten-rupee increase in the implicit monthly connection cost would cause connection probabilities to fall by approximately 82 percent.[2]

Another finding is that household characteristics matter. The decision to connect was positively and strongly affected by nonpolicy variables: higher levels of income, assets, and schooling. Water scarcity also matters. People in scarce-water areas were much more likely to connect, everything else being equal, than those in areas where water is abundant. Improved reliability turns out to be important only for current connectors. Those who were already connected indicated that they would be willing to pay substantially more for more-reliable service. More-reliable service did not, however, affect the probability that nonconnectors would decide to connect to the system.

The findings indicate that large potential welfare gains would be generated by a more liberal connection and pricing policy. Because of the problem of connection costs, we estimated that consumer surplus could be increased by at least 450 percent by amortizing the connection cost and folding it into the monthly tariff. We estimated that connections in the population would rise from the current 250 to about 2,500 under this policy regime.

## II. The 1991 Follow-up Survey

In the period following the 1988 contingent valuation study, improved water services were made available in the scarce- and saline-water areas. Site-B households in these two areas had to decide whether to connect to the improved system. The original site-B families in these two water environments were resurveyed in 1991 to determine whether they had connected.

2. Under reasonable assumptions about the interest rate and amortization schedule, the impact of the tariff and connection charges on the probability of connecting can be equalized, suggesting that respondents treated the tariff as a recurrent cost and the connection charge as a capital cost. The monthly charge for the connection cost in this example is calculated as if the connection cost were financed for six years at an annual real interest rate of 5 percent. A ten-rupee increase in the average monthly cost would thus correspond to an increase in the total connection charge from the bidding game average of Rs219 to Rs842. In fact, rural Indians are likely to face much stiffer credit market conditions, so this example is for most people probably a substantial understatement of the full connection costs when high interest rates are taken into account.

Table 3. *Probit Estimates of the Probability of Choosing a Yard Tap (1988 Survey)*

| *Variable* | *Coefficient* |
|---|---|
| Constant | −0.3009 |
| | (−1.28) |
| Tariff | −0.0605 |
| | (−20.31) |
| Connection charge | −0.0010 |
| | (−13.21) |
| Improved service | −0.0582 |
| | (−1.11) |
| Distance to current source (meters) | 0.00002 |
| | (0.06) |
| Queue at current source (minutes) | 0.0028 |
| | (1.54) |
| Per capita income (rupees) | 0.00002 |
| | (1.93) |
| Household has electricity | 0.3345 |
| | (3.83) |
| Number of rooms in dwelling | 0.0861 |
| | (3.50) |
| Household has females in government service | −0.0997 |
| | (−0.57) |
| Household has males in government service | 0.1664 |
| | (1.78) |
| Hindu household | −0.1908 |
| | (−2.07) |
| Female-headed household | 0.0569 |
| | (0.66) |
| Female respondent | −0.2749 |
| | (−3.70) |
| Some primary school | 0.5092 |
| | (3.39) |
| Primary school complete | 0.6293 |
| | (4.45) |
| Middle school complete | 0.9608 |
| | (6.51) |
| Secondary school complete | 1.1325 |
| | (8.09) |
| More than secondary school | 1.2898 |
| | (7.99) |
| Scarce-water area | 0.3474 |
| | (3.54) |
| Saline-water area | −0.2315 |
| | (−2.19) |
| A-site nonconnecting household | −0.3070 |
| | (−2.45) |
| B-site village household | −0.4921 |
| | (−3.87) |
| Pseudo $R^2$ | 0.28 |
| Sample Size[a] | 9,720 |
| Households | 1,150 |
| $\chi^2$ (22) | 3,272 |

*Note:* Dependent variable: whether respondent would choose to connect at each price quoted. Estimates are weighted by the population of the sampling unit. Standard errors are corrected using a method explained in Singh and others (1993). The estimating equation is significant at better than the 0.00001 level for a likelihood ratio test ($\chi^2$). The omitted site dummy is A-site connecting household. *t*-statistics are in parentheses.

a. The sample size of 9,720 results from multiple observations for each household in the sample. See Singh and others (1993) for a complete explanation.

*Source:* Authors' calculations.

### *Data Issues*

In the 1991 follow-up survey, an attempt was made to contact all of the respondents in the original sample in the two villages. The first issue to address in any follow-up survey is sample attrition. We lost 25 of 200 households (12.5 percent) from the sample in the scarce-water site and 59 of 200 households in the saline-water site (29.5 percent) between 1988 and 1991. The loss of households appears to have been random in the scarce-water site in that it left the original income distribution virtually intact. In the saline-water site, attrition was concentrated among the poorer households.[3]

"Connectors" in this analysis are defined as those who either were already connected to the new system at the time of the follow-up survey, had applied for a connection, or had made a decision to connect once applications were called for by the water authority. Both connectors and nonconnectors were asked a short series of questions concerning their decision and concerning exogenous changes that had come about since the original survey.

We predicted that a household would connect if the maximum bid for connection in 1988 was higher than the actual connection cost in 1991. If the household was connected at the time of the interview in 1991, the respondent was asked how much the connection actually cost. If the household was not connected, the interviewer estimated the cost of connecting for that household on the basis of its distance from the distribution line.[4]

### *Judging the Validity of Predictions*

What constitutes a success or failure in making a prediction based on the 1988 survey? In this section we lay out the criteria to be used in judging the success of the experiment.

The first and most important criterion, as in any sampling procedure, is whether we predict the correct proportion of connectors, notwithstanding whether we predict the exact behavior of each household. We would like to get as close as possible to the correct proportion. If we err, we prefer to err on the side of underpredicting the proportion of connectors because the concern with the direct approach is that its inherent biases tend toward overvaluation. Furthermore, we would not want to recommend installation of a water system on the basis of estimated demand and revenue forecasts that are too optimistic.

The second criterion, aimed directly at the heart of evaluating the technique, is whether specific households behave as they said they would. There are three elements here. The first is gross accuracy, or the proportion of the surviving sample for which actual behavior was correctly predicted. The second is the

3. The exact comparison of the two samples is available from the authors.

4. The connection cost in 1991 was deflated to 1988 rupees so that the numbers could be compared in real terms. We used the average inflation rate from 1980–90 as reported in World Bank (1992), 7.9 percent. This long-term average rate was used to avoid year-to-year fluctuations for this investment and in recognition of the imprecision of annual estimates.

specificity of connector predictions or the proportion of connecting households predicted correctly. The third is the sensitivity of predictions for nonconnectors or the proportion of nonconnecting households predicted correctly.

The statistical epidemiological literature distinguishes between specificity and sensitivity (see Kleinbaum, Kupper, and Morgenstern 1982). Specificity has to do with the problem of false positives, that is, the number of families that were predicted to connect but that actually did not connect. Sensitivity has to do with the problem of false negatives, that is, the number of families that were predicted not to connect but that actually did connect.

### *Comparing Benefit Revelation, Benefit Transfer, and a Behavioral Model*

We apply the above procedures to assess whether the behavior of families at the B sites could best be predicted by benefit-revelation or benefit-transfer methods. There are, accordingly, two research questions. First, did people at the B sites behave as they said they would? We call this the benefit-revelation question. Second, did people at the B sites behave as we had predicted on the basis of the predictions of a behavioral model for the A sites (after substituting the characteristics of households in the B sites)? We call this the benefit-transfer question. Prima facie, it is not obvious which of the two strategies implicit in the above questions would be most promising.

To partially assess the validity of the inferences drawn from the behavioral model (Table 3) estimated using the 1988 data (Singh and others 1993), we also compare predictions from it with actual behavior. Is the actual behavior of the families in the B sites in 1991 predicted accurately by a behavioral model based on the responses of those families in 1988? We call this the behavioral modeling question.

## III. Results for the Scarce-Water Site

This section analyzes the results for the scarce-water site. The results are also robust for the saline-water site, but we do not place equivalent weight on these findings, for reasons to be explained in the next section.

### *Benefit Revelation*

The fundamental question that the follow-up survey was designed to address is the simplest one: did people behave as they said they would? It was necessary to exclude those families for whom predictions could not be made because they bid at the maximum (Rs700) in 1988, and their actual connection costs in 1991 (which varied from house to house, depending on the distance to the public water line) exceeded this maximum bid.[5] This situation applied to 13 of the 161

5. For this group, we do not really observe a maximum bid because we know only that it was at least Rs700. Thus we cannot make a comparison with the connection cost faced in 1991.

Table 4. *Comparison of Actual and Predicted Behavior of B-Site Households in Water-Scarce Areas*
(number of households)

| *Predicted behavior* | *Actual behavior* | | *Total* |
|---|---|---|---|
| | *Did connect* | *Did not connect* | |
| *Benefit revelation*[a] | | | |
| Will connect | 15 | 6 | 21 |
| Will not connect | 7 | 120 | 127 |
| Total | 22 | 126 | 148 |
| *Benefit transfer, all A sites*[b] | | | |
| Will connect | 27 | 76 | 103 |
| Will not connect | 1 | 65 | 66 |
| Total | 28 | 141 | 169 |
| *Benefit transfer, scarce-water A site*[c] | | | |
| Will connect | 28 | 100 | 128 |
| Will not connect | 0 | 41 | 41 |
| Total | 28 | 141 | 169 |
| *Behavioral modeling*[d] | | | |
| Will connect | 10 | 13 | 23 |
| Will not connect | 18 | 128 | 146 |
| Total | 28 | 141 | 169 |

a. Prediction based on 1988 survey of B-site households in the scarce-water area.
b. Prediction based on 1988 survey of A-site households in all three water areas.
c. Prediction based on 1988 survey of A-site households in the scarce-water area.
d. Prediction based on probit model of 1988 B-site households in the scarce-water area.
*Source:* Authors' calculations.

families (7 of which actually connected).[6] The results for the remaining 148 respondents are presented in table 4 (under "benefit revelation").

The results indicate that 14.9 percent of the respondents (22/148) did connect. This is not statistically different (at the 5 percent significance level) from our prediction that 14.2 percent (21/148) of the families would connect. The gross accuracy of the predictions based on benefit revelation was 91 percent, as shown in table 5. That is, the behavior of 91 percent [(15 + 120)/148] of the families was consistent with the intentions they declared in the 1988 contingent valuation survey. The specificity of connector predictions, the percentage of those predicted to connect who actually did connect, was 71 percent (15/21). The sensitivity of nonconnector predictions, the percentage of those predicted not to connect who actually did not, was 94 percent (120/127). Thus, simple tabulations of respondents' answers to the contingent valuation survey in 1988 were remarkably accurate in predicting both the overall proportion of the

6. After losing 25 households because of attrition, we lost another 12 households that had missing values for connection cost in 1991 and 2 with missing values for the 1988 contingent valuation questions, which left us with 161 households for analysis in the 1991 sample.

Table 5. *Summary Statistics on the Accuracy of Predictions of Behavior of B-Site Households in Water-Scarce Areas*
(percentage)

| *Sample statistic* | *Value* |
|---|---|
| *Benefit revelation*[a] | |
| Gross accuracy | 91 |
| Specificity of connector prediction | 71 |
| Sensitivity of prediction for nonconnectors | 94 |
| *Benefit transfer, all A sites*[b] | |
| Gross accuracy | 54 |
| Specificity of connector prediction | 26 |
| Sensitivity of prediction for nonconnectors | 98 |
| *Benefit transfer, scarce-water A site*[c] | |
| Gross accuracy | 41 |
| Specificity of connector prediction | 22 |
| Sensitivity of prediction for nonconnectors | 100 |
| *Behavioral modeling*[d] | |
| Gross accuracy | 82 |
| Specificity of connector prediction | 43 |
| Sensitivity of prediction for nonconnectors | 88 |

*Note:* Gross accuracy is the percentage of the surviving sample for which actual behavior was correctly predicted. Specificity of connector prediction is the percentage of those predicted to connect and who actually did. The sensitivity of prediction for nonconnectors is the percentage of those predicted not to connect who actually did not.

a. Prediction based on 1988 survey of B-site households in the scarce-water area.
b. Prediction based on 1988 survey of A-site households in all three water areas.
c. Prediction based on 1988 survey of A-site households in the scarce-water area.
d. Prediction based on probit model of 1988 B-site households in the scarce-water area.
*Source:* Authors' calculations.

sample connecting and the household-specific choices once the households were given the opportunity of connecting to the piped water supply system in 1991.

In addition to determining whether families connected or not, the follow-up survey asked why. A variety of responses were given during these open-ended discussions. Only one emerged with any consistency—three of the thirteen respondents whose behavior was inconsistent with their 1988 response cited "changed economic circumstances." Two of the seven "unpredicted connectors" fell into this group, as did one of the six "unpredicted nonconnectors." For the nonconnectors, more than 75 percent indicated that "inability to pay the connection cost" was the primary reason for not connecting to the system, just as we had predicted from the analysis of the 1988 data.

Did reliability affect connection decisions in the sample in 1991? We predicted, on the basis of the 1988 results, that households already connected were concerned with reliability, but it was not an important consideration for those who were not already connected to a water system. The results of the ex post

investigations confirm this finding. Only a small proportion of nonconnectors (13 percent) replied that an inadequate quantity of water from the system was the main reason for not connecting, and no respondents mentioned service quality as a decisive reason for not connecting. However, those who had connected in 1991 unanimously expressed dissatisfaction with the reliability of the system. Every connector complained about the quantity of water available from the system during the dry season, and all but one of the connectors found the quantity inadequate even in the monsoon season.

### *Benefit Transfer*

Many reasonable people are justifiably suspicious of answers to hypothetical questions by subjects who have a strong interest in the outcome of the research. Extrapolation based on information gathered from a similar group that is not subject to these problems may reduce these biases and thus be a more reasonable basis for predicting actual behavior. The strategy of benefit transfer depends on the validity of models that can allow us to extrapolate from behavior or valuation of benefits in one area to populations of known characteristics in other areas. This method is still in its infancy (Pearce 1994; Desvousges, Naughton, and Parsons 1992; and Boyle and Bergstrom 1992). In the United States there has been only one legal proceeding in which contingent valuation estimates from one study were used to estimate values in another site, and the court refused to accept the "transferred values" as legitimate evidence (Brookshire and Neill 1992).

Three guidelines have been developed for use in research on benefit transfer. First, the study site should be very similar to the policy site. Second, the policy change or project at the study site should be very similar to that proposed at the policy site. And third, the valuation procedures used at the study site should be analytically sound and carefully conducted (Pearce 1994). The Kerala study follows these guidelines exactly.

To address the benefit-transfer research question, the contingent valuation survey of households at the A sites was used to estimate a model explaining the probability of connection for them. As described in Singh and others (1993), respondents' choices were modeled using a random utility framework, in which the probability that a family chooses to connect is determined by household characteristics, characteristics of the improved supply system, and characteristics of alternative water supplies. Table 3 contains estimates for the full model using all observations for all A and B sites; appendix table A-1 contains analogous estimates for the three subsamples used in this section: all A sites, the scarce-water A site, and the scarce-water B site. Under reasonable assumptions on the distribution of errors, probit estimates are unbiased. They are not efficient, however, because each household recorded repeated bids. The standard errors have been corrected for within-household correlations among groups of observations (described in detail in Singh and others 1993).

Before proceeding, it is important to explain what these estimating equations do and do not do. The coefficients are consistent estimates of the true effect in the population of each variable on the decision to hook up. Given the high *t*-statistics for the important variables, we can be fairly certain that if other samples were drawn from these populations, we would get roughly the same coefficient estimates. However, the pseudo $R^2$ statistics at the bottom of tables 3 and A-1 indicate that only about one-third of the variation in the dependent variable—whether to connect—is explained by the model. Although this statistic is relatively high for cross-sectional models, it suggests that the specified economic and social variables determine only a limited fraction of behavior.

Using this model for one site to predict responses at another site might be problematic for several reasons. For example, some variables might be out of range. The coefficients are valid only for the range of variables occurring in the sample observed in the first site. Another possibility is that the predictable component of behavior may be overwhelmed by unobservable and site-specific effects and random effects. Still another possibility is that comparison across time periods without new information might be problematic, in that some of the observed variables in the equation may have changed, but the changes cannot be observed. One example is income. In predicting behavior in 1991, we bumped up every household's income by the real growth rate of the economy, but doing so fails to capture relative changes in real income across the households.

Using models of behavior at the A sites to predict behavior at the scarce-water B site gives completely inaccurate predictions. This is true whether households in all the A sites or only the households in the scarce-water A site are used (table 4). We assume that the results using the scarce-water A site would be the most accurate, as this site is the closest sample to the scarce-water B site. Using the benefit-transfer approach, we predicted that a very high proportion of the B-site households would connect—76 percent (128/169). From table 4, only 16.6 percent (28/169) of the families connected. The difference is statistically significant. Gross accuracy is 41 percent, specificity of connector predictions is 22 percent, and sensitivity of nonconnector predictions is 100 percent. In short, the performance of the benefit-revelation method (based on the responses to contingent valuation questions at the B site) is vastly superior to that of the benefit-transfer method (based on modeling behavior at the A site and transferring the results to the B site).

We do not have to look far for reasons why the benefit-transfer approach works so poorly in this case. The basic reason is not model specification (in the next subsection we demonstrate that using data from the B site to estimate the coefficients in the model tends to overpredict connectors slightly, but otherwise performs well in explaining behavior), but the much lower bids in the B site in 1988. The descriptive statistics at the bottom of table 1 show that the average connection bid in the B sites in 1988 (Rs267) was only 75 percent of the average for the nonconnecting A sites (Rs355). The average maximum bid for the monthly tariff in the B sites was Rs5.5, only 28 percent of the average for the connecting

A sites (Rs19.3) and 63 percent of the average for the nonconnecting A sites (Rs8.7). The coefficients for the dummy variable for the villages in the B sites in the full model (table 3) are negative and highly significant. In other words, the households in the B sites placed a relatively low value on the new water system in 1988, and they were telling the truth about it: their responses in 1988 predicted their own behavior very well. The respondents in the A sites turned out to be very poor substitutes for the respondents in the B sites, apparently because of unobservable, site-specific factors.

### *Behavioral Modeling*

As described above, the benefit-revelation method gave reliable predictions of actual behavior at the B site. Although this is an important finding, in practice it provides an answer to only part of the question of interest to researchers and policymakers, who are also concerned with likely responses to changes in a cluster of related policy variables (such as tariffs, connection charges, and reliability) while controlling for other confounding variables (such as education and income). The simple tabular analysis of the contingent valuation data from the B site cannot, of course, answer such questions.

The numbers in table 4 were generated by predicting behavior in the scarce-water B site using only the data from that sample and the model appearing in the third column of appendix table A-1.[7] In table 4, the sample size is increased to 169.[8]

We had predicted that 13.6 percent (23/169) of families would connect, whereas in fact 16.6 percent (28/169) of the larger sample of families did connect. Thus, the model meets our criterion of being conservative, that is, it underestimates connections. The difference is not statistically significant at the 5 percent level. The gross accuracy was 82 percent, compared with 91 percent for the simple analysis. The specificity of connector predictions fell from 69 percent in the simple case to 43 percent. The sensitivity of nonconnector predictions was 88 percent, down from 95 percent in the case of the simple statistics.

The behavioral model estimated using the contingent valuation data from the B site does not predict behavior quite as well as the simple descriptive statistics do. Because the behavioral model should give us more information rather than less, this result suggests that functional form and distributional assumptions

7. We also performed the analysis summarized in table 4 (under "behavioral modeling") using the whole-sample estimates presented in table 3. The results are almost identical. The counterpart to table 4 is available from the authors.

8. The change in sample size from 148 to 169 can be explained as follows: we started with 148 in table 4 (under benefit revelation), and then we gained 13 that had been dropped because their bids were at the maximum. For these households, we could predict the probability that they would connect because of the properties of the estimating equation. We gained another 12 households, for which the willingness-to-pay data were missing in 1988, but for which we knew whether they connected in 1991. For these households we could predict whether they would connect using the model. We lost 4 households with a missing value for one or more independent variables in the equations, even though we knew the outcome variables for them.

may be more restrictive than we expected. But we doubt that conclusion and attribute the lower accuracy to the fact that the model allowed us to include those observations at the maximum bid that had to be dropped for table 4 (under "benefit revelation"). These households introduced virtually all the additional error. If we use behavioral modeling to predict for the same sample that appears in table 4 (under "benefit revelation"), we get similar results to those in that table (results not reported here).

Thus, inferences from the econometric model probably provide good, conservative guidance for water policy. Inferences from the model about elasticities (with respect to price, income, and reliability, for instance) and changes in welfare are likely to be reliable. This result suggests that the strong policy conclusions of the initial study (Singh and others 1993) should be taken quite seriously. That article suggested that substantial improvements in welfare would be possible with drastic reform of water sector policies in Kerala. Our results do suggest, however, that considerable care should be taken to make sure that the range of responses to the contingent valuation questions do not artificially censor those who might bid very high or very low. In the simple benefit-revelation case, the censoring requires us to throw out observations; in the behavioral model case, it reduces the predictive accuracy of the model, even though we can predict values for those observations.

## IV. Results for the Saline-Water Site

We will only briefly present the results for the saline-water site because the results buttress the findings from the scarce-water site but are less interesting and less reliable.[9] As already noted, sample attrition was much greater in the saline-water site and was concentrated among the poorer households (we lost 30 percent of the sample, compared with only 13 percent for the scarce-water site). In the saline-water site, we lost twenty-one households for the benefit-revelation exercise because they were at the maximum bid for connection cost. Nineteen of these did not connect, and two did. We lost an additional three households with a missing value for connection cost in 1991.

The results for the saline-water site are shown in table 6. We predicted that 0 percent (0/117) of families would connect, whereas in fact 15.4 percent (18/117) of families did connect. So, although we were infinitely wrong about the proportion connecting, we were at least wrong in the desired direction, underpredicting connectors.

Our failure to predict connectors appears to have been caused by exogenous changes that took place at the saline-water site between surveys. During the

9. We do not include here the additional analyses performed for the scarce-water site because they do not add new information. The results are basically the same as those for the scarce-water site, with the benefit-transfer model performing slightly better than for the scarce-water site (but still badly overpredicting connections) and the behavioral model basically replicating the benefit-revelation results with less accuracy. These results are available from the authors on request.

Table 6. *Comparison of Predicted Behavior of B-Site, Saline Water Area Households in 1988 and Actual Behavior in 1991*
(number of households)

| | *Actual behavior*[b] | | |
|---|---|---|---|
| *Predicted behavior*[a] | *Did connect* | *Did not connect* | *Total* |
| Will connect | 0 | 0 | 0 |
| Will not connect | 18 | 99 | 117 |
| Total | 18 | 99 | 117 |

a. Based on the maximum bid that households were willing to pay for a connection in 1988.
b. Based on the actual number of connections after true cost was known.
*Source:* Authors' calculations.

follow-up survey, 78 percent of the households in this area that did connect, despite our predicting that they would not, cited "unanticipated improved economic circumstances" as the reason for connecting. Another 22 percent said it was because they could borrow money for the connection (in contrast, no one in the scarce-water site borrowed to finance the connection).

The gross accuracy of predictions based on the intentions households declared in the 1988 contingent valuation survey was 85 percent. Thus, even though we predicted no connectors, because the proportion connecting was small, we did very well in predicting behavior for the whole sample. The specificity of connector predictions was 100 percent. We were by definition perfectly correct in getting our predicted connectors right, because we predicted none. The sensitivity of nonconnector predictions was 85 percent.

The results for the saline water site show the fragility of our endeavor to compare bids with behavior three years later, with almost no information about what happened in between. Even so, the results are fairly robust, despite our misgivings even about the 1988 data in the saline-water site. For example, analysis of the 1988 data suggested that a family in a saline-water area was willing to pay less for an improved water supply than a similar family in a water-rich area (see table 3), which was counterintuitive. But more important, between the first and second rounds of the surveys, there were major exogenous changes in the saline-water villages.

## V. Conclusions

Contingent valuation studies suffer from three potential sources of bias (Cummings, Brookshire, and Schultze 1986). First, strategic bias might arise because respondents perceive it to be in their interest to respond inaccurately to the questions. Second, hypothetical bias might arise because respondents are not fully acquainted with the good or service in question. And third, compliance bias might arise because respondents give replies they believe the questioners would find most satisfactory. A priori, in this particular setting, it might be predicted that strategic biases would be relatively high, with respondents underestimating their willingness to pay in the hope that this might lower the tariffs

that would be charged. Hypothetical biases would be relatively low because the good in question—a piped water supply—is familiar to all. Compliance biases might be relatively high, because this has happened with surveys in the Indian subcontinent in the past (for example, Mamdani 1972).

The comparison of actual behavior with that emanating from the contingent valuation survey amounts to a resounding confirmation that, *in this particular study*, the net effect of these biases was small. The caveat regarding "this particular study" cannot be overstressed for several reasons. First is the issue of strategic and compliance bias. Recognizing the potential for these sources of error, the study made strenuous efforts to reduce to a minimum the effects of these biases: by explaining procedures in the survey; by being—and being seen as—independent of the supply agency; and by training interviewers rigorously. There is considerable evidence that "quick and dirty" willingness-to-pay surveys of a similar nature in the past have yielded nonsensical results (Saunders and Warford 1972). Accordingly, the survey instruments were developed with great care in the course of a multicountry study (World Bank Water Demand Research Team 1993) and were already "tried and tested" by the time of the 1988 survey in Kerala. As shown in table 2, the instrument met the NOAA standards five years before the standards were developed. Furthermore, the instrument was carefully pretested in Kerala, and modifications were made so that it was appropriate for the local setting.

Second is the issue of hypothetical bias. Proponents of the contingent valuation methodology have understood for some time that the greatest problem for contingent valuation studies arises not from strategic but from hypothetical bias (see Cummings, Brookshire, and Schultze 1986). Stimulated by the controversy on the damage caused by the Exxon Valdez oil spill in Alaska, a rigorous and heated debate on the contingent valuation methodology has taken place, with much of the attention focused on the issue of hypothetical bias (see Portney 1994; Hanneman 1994; and Diamond and Hausman 1994). Diamond and Hausman (1994: 62) have made a forceful denunciation of the method, arguing that "contingent valuation is a deeply flawed methodology for measuring nonuse values." With respect to this controversy, the Kerala water study simultaneously (a) provides clear evidence that carefully conducted contingent valuation studies can provide reliable information on how people value well-defined goods and services and (apparently paradoxically) (b) does not contradict the concerns that underlie the Diamond-Hausman argument.

The key issue here is that, even though the good at stake in the Kerala study was tangible and simple, there was still a problem with hypothetical bias in parts of the study. This problem emerged in the finding that an important service characteristic—reliability—was of major importance to those who were already connected, but was not perceived as being important by those who had not directly experienced the service. Well-conducted contingent valuation studies can provide reliable and valuable information on behavioral responses to well-defined and well-understood goods such as a household water supply. But

this finding in no way vitiates the very serious problems arising when this method is used to assess such abstract concepts as "existence values." In the words of Diamond and Hausman (1994: 62), "it is precisely the lack of experience both in markets for environmental commodities and in the consequences of such decisions that makes contingent valuation questions so hard to answer and the responses so suspect."

### *Benefit-Transfer Literature*

The results are equally striking for the prospects of using the estimates based on a behavioral model for one population to predict the behavior of another population. Virtually all the characteristics of the study population and the alternative water sources are apparently quite similar at the A and B sites. When behavior at the scarce-water B site was estimated from a well-specified and carefully estimated model of actual behavior at the scarce-water A site, however, predictions were wrong for about half of the sample. The number of connections was overestimated by a factor of four.

This finding is not surprising, for two reasons. First, even when the determinants of behavior are easy to specify (as in this case), detailed models of this behavior are formulated, the full required set of data is collected, and sophisticated statistical tools are applied, less than one-third of variance can be explained (see table 3). Second, the results of a multicountry study of willingness to pay for water (World Bank Water Demand Research Team 1993) shows that both appropriate specifications and parameter estimates vary considerably in different locations.

Attempts to estimate behavior (and thus benefits) in a particular community on the basis of results of studies in other communities in other settings can reach conclusions that are seriously erroneous. This can occur even when the communities, the natural conditions, and the service to be offered are apparently quite similar. Substantial additional information is collected when the expected beneficiaries are interviewed directly.

### *Assessing the Demand for Services through Behavioral Models*

Benefit revelation through direct methods has great potential for assessing the demand for services, especially for capital-intensive and costly services such as water supply in developing countries. Carefully designed and conducted contingent valuation studies can produce reliable estimates of the demand for water and sanitation and are appropriately becoming widely used for this vital function (Whittington and others 1992; Altaf, Jamal, and Whittington 1992).

The sample size requirements are modest (a couple hundred families sufficed in this case). As experience with these studies accumulates, it has been possible to substantially improve quality, increase speed, and reduce cost by the judicious use of off-the-shelf survey components. If the policy interest is limited (in the present case, to the number of families who would connect to a new supply), then simple tabular analyses may suffice. If the policy interest is more complex

(for example, elasticities with respect to price and service reliability; simulations of policy changes; or welfare analysis), then behavioral models using econometric techniques need to be estimated.

### *Caveats*

The contingent valuation method is validated under a very specific set of circumstances. Of particular importance (as stressed earlier) is the fact that hypothetical biases for the service evaluated in this study are much more limited than for many of the environmental resources that the technique is typically used to value. Of equal importance is the careful design and conduct of these studies. In all cases this meant several weeks of pretesting and adaptation to local circumstances, meticulous training and supervision of interviewers, and careful cleaning of data. More specifically, it is noted with concern that the relative success of the set of studies of which the Kerala one is part (World Bank Water Demand Research Team 1993) has inspired some investigators to conduct two-day studies of hundreds of households to determine willingness to pay. In the past such studies gave willingness-to-pay surveys a (well-deserved) bad name. The old adage of "garbage in, garbage out" is certainly applicable to such poorly conducted studies.

In some circumstances, researchers and policy analysts will have access to rich data sets on populations that (a) already have access to the service of interest and (b) are very similar to the population for which the service is to be introduced. In such circumstances, it would appear that carefully specified and estimated models of behavior could be used to predict behavior by the unserved population. The results of this study, however, show that site-specific factors are of major importance and that predictions based on extrapolations may be far off the mark, even when many conditions in the population and environment are apparently similar.

### *What We Learned That Can Improve Survey Design*

Much more care should be exercised in defining the range and the increments for bidding games or referenda. They should be connected as closely as possible to actual costs, or the ranges should be pretested for validity. Often researchers pick very high willingness-to-pay bids for the top of the range, which they think are higher than anyone would pay, yet when the data come back they show that large proportions of the sample have chosen the highest bid. That does not give us good information about willingness to pay. In our case, despite efforts to design the best possible survey in 1988, the top of the ranges for tariffs and connection charges was too low. For example, in the scarce-water site, one connector who bid at the maximum of Rs700 in 1988 actually paid Rs2,547 for the connection. In this case, we could have avoided the problem by reviewing the range of actual connection costs in 1988 before finalizing the survey.

Quick analysis of the survey data immediately after the fieldwork, with a plan to return to the field for additional work, would resolve many questions

that we had in the analysis. In the saline-water site, for instance, we should have caught right away (in 1988) that the bidding behavior was not as we expected, and households should have been reinterviewed. After the 1991 survey, we should have immediately gone back to understand better why six households that had bid Rs100 in 1988 for connections actually hooked up at an average cost of Rs1,380. Doing the analysis quickly in the field, preparing a qualitative questionnaire to complement the quantitative work, and aggressively addressing issues of data quality could raise the validity of benefit-revelation techniques considerably and increase their value for assessing demand.

If other researchers conduct the same type of test reported in this article, we suggest collecting additional information about the households over the elapsed time between the hypothetical bids and the actual choice. Repeated observations on income, assets, family size, community characteristics, precipitation, and traditional water quality would help control for endogenous and exogenous changes over the period.

### *Implications for Kerala*

To the extent that contingent valuation, or benefit revelation, has been validated by this exercise, it has been driven by the nonconnectors. We predicted that very few would connect to the water systems, given the policies in place. We were right. Standard water systems as designed in India and in much of the developing world do not make people better off. People respond by letting the systems fall apart or by telling us, as they did in the 1988 survey, that they would not connect even at the fairly low prices that were quoted. Their behavior in 1991 just turned the hypothetical rejection of the system by the vast majority of the sample into an actual rejection.

We also predicted, on the basis of the 1988 data, that once people connected, they would become concerned about the poor quality and low reliability of the system. They would not care much about these problems unless they connected. We were right again.

Although we would not have been able to make these predictions without a survey and a technique that worked as expected, we still do not know if we were right about the prescriptions we offered for solving these problems. On the basis of simulations, we recommended that the water authority raise tariffs, fold the connection cost into the tariff, and vastly expand access to private connections. We predicted that both connections and revenue would explode. People would consider themselves much better off. They would begin agitating to pay more to create a more reliable system. Making these policy changes, and tracking the results, would be the true test of the work. We know the technique worked. We do not know yet if the economic analysis was accurate, because the changes in policies that we recommended as a result of the analysis have not been tested.

Table A-1. *1988 Survey: Probit Estimates of Probability of Choosing a Yard Tap for Subsamples*

| *Variable* | *All A sites* | *Scarce-water A sites* | *Scarce-water B sites* |
|---|---|---|---|
| Constant | −0.3302 | 0.6083 | −0.1693 |
| | (−0.877) | (1.00) | (−0.47) |
| Tariff | −0.0520 | −0.0461 | −0.0947 |
| | (−15.57) | (−10.54) | (−6.08) |
| Connection charge | −0.0010 | −0.0008 | −0.0017 |
| | (−8.32) | (−3.76) | (−8.49) |
| Improved service | −0.1394 | 0.0678 | |
| | (−3.15) | (0.97) | |
| Distance to current source (meters) | 0.0103 | 0.0120 | 0.0013 |
| | (0.75) | (1.31) | (2.83) |
| Queue at current source (minutes) | 0.0013 | 0.0067 | −0.0076 |
| | (0.64) | (0.82) | (−1.33) |
| Per capita income (rupees) | 0.00004 | 0.00005 | −0.000001 |
| | (1.99) | (1.05) | (−0.04) |
| Household has electricity | 0.0024 | 0.3689 | 0.2937 |
| | (0.02) | (1.29) | (1.35) |
| Number of rooms in dwelling | 0.0441 | 0.0721 | 0.2677 |
| | (1.53) | (1.33) | (2.75) |
| Household has females in government service | −0.1894 | 0.1652 | −0.3037 |
| | (−0.67) | (0.48) | (−0.41) |
| Household has males in government service | −0.1495 | −0.0694 | 0.0741 |
| | (−1.15) | (−0.30) | (0.32) |
| Hindu household | −0.1138 | −0.9599 | −0.3984 |
| | (−0.89) | (−3.24) | (−2.17) |
| Female-headed household | 0.0153 | 0.5330 | −0.1017 |
| | (0.12) | (2.26) | (−0.51) |
| Female respondent | −0.3183 | −0.3582 | 0.1192 |
| | (−2.68) | (−1.66) | (0.71) |
| Some primary school | 0.7960 | 1.2941 | 0.3312 |
| | (2.30) | (2.47) | (1.47) |
| Primary school complete | 1.1315 | 1.3099 | 0.2859 |
| | (3.74) | (3.44) | (1.22) |
| Middle school complete | 1.0451 | 0.3247 | 0.3709 |
| | (3.13) | (0.73) | (1.41) |
| Secondary school complete | 1.4341 | 1.7492 | 0.6238 |
| | (4.77) | (4.60) | (2.63) |
| More than secondary school | 1.7274 | 2.0476 | 0.4956 |
| | (5.30) | (4.53) | (1.36) |
| Scarce-water area | 0.6983 | n.a. | n.a. |
| | (4.41) | | |
| Saline-water area | −0.1446 | n.a. | n.a. |
| | (−0.91) | | |
| A-site nonconnecting household | −0.4557 | −1.1346 | n.a. |
| | (−3.14) | (−5.00) | |
| B-site village household | n.a. | n.a. | n.a. |
| Pseudo $R^2$ | 0.26 | 0.30 | 0.34 |
| Sample size | 5,228 | 1,700 | 1,416 |
| $\chi^2$ (degrees of freedom) | 1,704(21) | 697(19) | 554(17) |

n.a. Variable not appropriate to the sample.

*Note:* Dependent variable: choice to connect in the bidding game. Estimates are weighted by the population of the sampling unit. Standard errors are corrected using a method explained in Singh and others (1993). The estimating equations are significant at better than the 0.00001 level for a likelihood ratio test ($\chi^2$). The omitted site dummy is A-site connecting households. *t*-statistics are in parentheses.

*Source:* Authors' calculations.

## REFERENCES

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

Altaf, Mir Anjum, Haroon Jamal, and Dale Whittington. 1992. "Willingness to Pay for Water in Rural Punjab, Pakistan." Water and Sanitation Report 4. World Bank, Transport, Water, and Urban Development Department, UNDP–World Bank Water and Sanitation Program, Washington, D.C. Processed.

Arrow, Kenneth, Robert Solow, Paul Portney, Edward Leamer, Roy Radner, and Howard Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." Department of the Interior, Washington, D.C. Processed.

Boyle, K. J., and J. C. Bergstrom. 1992. "Benefit Transfer Studies: Myths, Pragmatism, and Idealism." *Water Resources Research* 28(3):657–63.

Briscoe, John, Paulo F. de Castro, Charles Griffin, James North, and Orjan Olsen. 1990. "Towards Equitable and Sustainable Rural Water Supplies: A Contingent Valuation Study in Brazil." *The World Bank Economic Review* 4(2):115–34.

Brookshire, David S., Mark A. Thayer, William D. Schultze, and Ralph D'Arge. 1982. "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches." *American Economic Review* 72(1):165–77.

Brookshire, David S., and H. R. Neill. 1992. "Benefit Transfers: Conceptual Issues and Empirical Issues." *Water Resources Research* 28(3):651–56.

Carson, Richard T., and Robert Cameron Mitchell. 1993. "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water." *Water Resources Research* 29(7):2445–54.

Cummings, Ronald G., David S. Brookshire, and William D. Schultze. 1986. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, N.J.: Rowman and Allanheld.

Desvousges, W. H., M. C. Naughton, and G. R. Parsons. 1992. "Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies." *Water Resources Research* 28(3):675–83.

Diamond, Peter A., and Jerry A. Hausman. 1994. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8(4):45–64.

Hanneman, W. Michael. 1994. "Valuing the Environment through Contingent Valuation." *Journal of Economic Perspectives* 8(4):19–43.

Kleinbaum, D. G., L. L. Kupper, and Hal Morgenstern. 1982. *Epidemiological Methods: Principles and Quantitative Methods.* Belmont, Calif.: Wadsworth.

Mamdani, Mahmood. 1972. *The Myth of Population Control: Family, Caste, and Class in an Indian Village.* New York: Monthly Review Press.

Mitchell, Robert Cameron, and Richard T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington, D.C.: Resources for the Future.

North, James H., and Charles C. Griffin. 1993. "Water Source as a Housing Characteristic: Hedonic Property Valuation and Willingness to Pay for Water." *Water Resources Research* 29(7):1923–29.

Pearce, David W. 1993. *Economic Values and the Natural World.* London: Earthscan Publications.

———. 1994. *Project and Policy Appraisal: Integrating Economics and Environment.* Paris: Organisation for Economic Co-operation and Development (OECD).

Portney, Paul R. 1994. "The Contingent Valuation Debate: Why Economists Should Care." *Journal of Economic Perspectives* 8(4):3–17.

Saunders, Robert J., and Jeremy J. Warford. 1972. *Village Water Supply: Economics and Policy in the Developing World.* Baltimore: Johns Hopkins University Press.

Serageldin, Ismael, and Andrew Steer, eds. 1994. *Valuing the Environment: Proceedings of the first Annual International Conference on Environmentally Sustainable Development.* Washington, D.C.: World Bank.

Singh, Bhanwar, Radhika Ramasubban, Ramesh Bhatia, John Briscoe, Charles C. Griffin, and Chongchun Kim. 1993. "Rural Water Supply in Kerala, India: How to Emerge from a Low-Level Equilibrium Trap." *Water Resources Research* 29(7, July): 1931–42.

Whittington, Dale, Donald T. Lauria, Albert M. Wright, Kyeongae Choe, Jeffrey Hughes, and Venkateswarlu Swarna. 1992. "Household Demand for Improved Sanitation Services: A Case Study of Kumasi, Ghana." Water and Sanitation Report 3. UNDP–World Bank Sanitation Program, Washington, D.C. Processed.

World Bank. 1992. *World Development Report 1992.* New York: Oxford University Press.

World Bank Water Demand Research Team. 1993. "The Demand for Water in Rural Areas: Determinants and Policy Implications." *The World Bank Research Observer* 8(1, January): 47–70.

THE WORLD BANK ECONOMIC REVIEW VOL. 9, NO. 3 397-423

# Political Influence on the Central Bank: International Evidence

Alex Cukierman and Steven B. Webb

*Political influence on the central bank is measured here by looking at the probability that a central bank governor will be replaced shortly after a political change of government. The governor changes about half the time within six months of a nonconstitutional or other radical change of government—a military coup or a restoration of democracy. The governor is much less likely to change within six months following a routine change in the head of government—about one-fourth of the time in developing countries and one-tenth in industrial countries. These indicators vary across countries and correlate statistically with inflation and its variability and with real growth and real interest rates. Differences in the vulnerability of the central bank to political instability, in political instability itself, and in central bank turnover in nonpolitical periods seem to be a major part of the explanation for why developing countries have, on average, higher and more variable inflation than industrial countries do.*

Economists and policymakers generally feel that the degree of autonomy of the central bank from political authorities is an important determinant of policy choices and of economic performance. Empirical verification of these presumptions has been difficult, however, because the autonomy of the central bank is not easily quantified. Most previous studies have used legal indexes from central bank charters to quantify the autonomy of the central bank (Parkin 1986; Grilli, Masciandaro, and Tabellini 1991; Alesina and Summers 1993). These measures help account for cross-country inflation differentials within industrial economies but not within developing economies. Low inflation is not associated with the legal independence of central banks in developing countries because of the small degree of association between actual and legal independence among those countries. In industrial countries, the frequency with which a central bank governor is replaced within three months of the time designated by law is more than ten times higher than in other periods. In developing countries it is only 2.2 times higher.

Alex Cukierman is with the Department of Economics at Tel-Aviv University and with the Center for Economic Research at Tilburg University, the Netherlands, and Steven B. Webb is with the Latin America and the Caribbean Country Department III at the World Bank. The authors acknowledge important research assistance by Charles Guo, Pantelis Kalaitzidakis, Bilin Neyapti, Pedro Rodriguez, and Kenneth Xu and assistance by Stephan Haggard and Yasuhito Asami in reviewing the political data in detail. The authors are thankful for the useful comments of several individuals and groups, including Thomas Havrilesky, Suzanne Lohmann, and Sylvia Maxfield. Research for this article was funded by World Bank research grant RPO 677-77.

Hence, more behavioral indexes of independence are called upon. Cukierman, Webb, and Neyapti (1992) and Cukierman (1992) made an initial step in this direction by using the actual average term in office of the governor as a proxy for central bank independence in developing countries. They found a significant negative association between this proxy and inflation across developing countries. The use of this proxy is based on the presumption that below some threshold a shorter term in office of the chief executive officer (to whom we refer as "governor," although the actual title may be president or chairman) of the bank is associated with lower central bank independence. It should be noted, in this context, that actual terms in office in developing countries are substantially lower than in industrial countries. Cukierman (1994) shows theoretically how the expected length of tenure of the central bank governor relates inversely to inflation because of the governor's effective rate of time preference. This article presses the inquiry further by taking account of how turnover of the governor relates to political events.

Although considerable evidence now establishes the negative association between inflation and central bank autonomy, the reduction of political influence on the central bank is not the only institutional device for assuring price stability. Regardless of whether a country has a dependent central bank, the country may still enjoy price stability if it consistently pegs its currency to that of a country with stable prices. Examples include Argentina in recent years, Belgium, the Netherlands, and some francophone African countries. An even stronger commitment to maintain a fixed exchange rate is to use a foreign currency for legal tender, as in Panama. But it should be emphasized that central bank autonomy and various degrees of commitment to a fixed exchange rate are complementary institutional arrangements and are not mutually exclusive. This article focuses on the documentation and the effects of the varying degrees of political influence on the central bank. We and others have investigated the effects of exchange rate commitments elsewhere. Empirical investigations of the effect of exchange rate pegs on inflation appear in Cukierman, Rodriguez, and Webb (forthcoming) and in Anayadike-Danes (1995). Cukierman, Kiguel, and Liviatan (1992) offers a general discussion of dollarization.

A variety of research—theory, case studies, and statistical analysis—shows that political instability worsens macroeconomic policy and outcomes (Edwards and Tabellini 1991; Haggard, Kaufman, and Webb 1992; Haggard and Webb 1994). The effects are strongest on inflation, but other variables such as growth and the real interest rate are also affected. Some macroeconomic deterioration results from the direct effects of political change on policy, particularly fiscal policy. Some also come from the effect on institutions, such as the central bank, and on their ability and incentive to follow policies for the long-run benefit of the economy. So the research agenda includes the question of the relative importance of different channels through which political instability influences economic outcomes, as well as the underlying questions of whether and how political instability affects institutions such as the central bank.

It seems likely that different kinds of political instability would have quite different effects. If political changes reflected changes in basic attitudes toward economic policy or if they were traumatic and irreversible for the politicians involved, then the instability would motivate politicians to control the central bank tightly and keep it at their disposal to help them stay in power. If, however, the political changes were alternations in power of two or three parties that shared a consensus on many basic tenets of economic policy, then the parties might agree to grant the central bank considerable autonomy to pursue price stability, so that the aspect of economic policy on which they agreed would not suffer from the political contests over other issues. These questions motivated the research for this article, although the article does not answer all of them.

This article presents new behavioral indexes of political influence on the central bank. The indexes focus on the propensity of the governor of the bank to lose office following a political transition. The data base and the indexes derived from it were designed to address four specific questions derived from the broader questions in the previous paragraphs.

*First, is turnover at the central bank significantly different shortly after political transitions than in other periods?* To answer this question, we compare the frequency with which governors are changed in periods shortly after a political transition with that in other periods. We refer to these other periods as "nonpolitical," although recognizing that even then some turnover may result from political influence. Nevertheless, in the presence of heavy political influence we should expect central bank governors to be changed with significantly higher frequency in periods immediately after political transitions, which we call "political periods," than in the nonpolitical periods. The tests presented in the article suggest that this is indeed the case for the entire sample of sixty-seven economies, for the subgroup of developing economies, and, surprisingly, for the industrial economies as well. Given this result, the frequency with which central bank governors are changed in political periods can serve as an index of the extent of political influence over the central bank and thus lead to more refined indicators of central bank autonomy.

*Second, for how long does a political transition, after its occurrence, increase the probability of a change of the central bank governor?* To answer this question we examine a number of two- or three-month successive intervals following political transitions and calculate the frequency of changes at the central bank within each interval. The frequency of changes is high on average immediately after a political transition and usually decreases as more months go by. The cutoff between political and nonpolitical periods is obtained by considering the last interval within which the frequency of changes at the central bank is significantly higher than the average frequency at the ten-month-or-more lag. The demarcation line between political and nonpolitical periods is then defined as the upper bound of this critical interval. With this procedure, periods within six months of the latest political transition are defined as political in the overall sample, and any period beyond that is nonpolitical.

*Third, is the political vulnerability of the central bank systematically related to the level of economic development and the type of political regime?* We calculate a measure of the political vulnerability of the central bank and examine its relation with a country's level of development and the nature of its regime (whether it is always democratic, always authoritarian, or mixed). Vulnerability of the central bank is defined as the fraction of political transitions that are followed, within the subsequent political periods, by a replacement of the governor of the bank.

Frequent removal from office of the bank's governor following political transitions probably reflects gross political influence, because the governor's term in office is not shielded by law or custom from political changes. This type of influence has been common in developing economies, such as Botswana, China, Costa Rica, and Indonesia, and has been particularly high in countries that switch between democratic and authoritarian regimes, such as Brazil, Chile, India, the Republic of Korea, Nicaragua, Peru, and Uruguay. In Argentina before the 1990s, even though the law specified a four-year term, the governor was always replaced when the government changed. Among industrial economies gross political influence on the central bank is rare—Sweden appears to be an exception—but more subtle and mild influences are common. For instance, in the United States, the index of central bank vulnerability is zero, implying the absence of gross political influence. But, as documented in Havrilesky (1992), there are several other, milder, channels of political influence on the Federal Reserve.

*Fourth, is the political vulnerability of the central bank systematically related to measures of economic performance such as inflation, growth, and interest rates?* Cross-sectional and panel-data regressions provide evidence on this question.

The article is organized as follows. Section I presents the data set and discusses the conventions used to organize it. The procedure for distinguishing between political and nonpolitical changes of central bank governors is developed and applied in section II to the entire sample as well as to broad subgroups of countries. Section III discusses indexes of central bank political vulnerability for broad subgroups of countries and investigates the effect of different types of political transitions on central bank vulnerability. The effects of vulnerability and of nonpolitical turnover on inflation, growth, and real rates are briefly examined in section IV. Section V estimates empirically how much of the strong cross-sectional association between inflation and its variance is due to their common association with central bank vulnerability and nonpolitical turnover. Section VI follows with concluding remarks.

## I. The Data

To assess and quantify the degree of political influence on the central bank, we have put together a multicountry data set on political and central bank instability. Political instability is measured by counting political transitions of various types. Instability at the central bank is measured with data on the frequency

and timing of replacement of central bank governors. The research focuses on variables that were available for a large number of countries on a uniform basis.

The sample consists of sixty-seven economies with matched data on political and central bank instability. The economies are listed in table A-1. The sample includes all the major industrial and developing economies, but excludes most Eastern European economies. The data cover 1950–89, where possible, but start later for economies that achieved political independence or established a central bank after 1950. The data are divided into two subperiods: 1950–71, corresponding to the Bretton Woods era, and 1972–89. In each subperiod as well as in the total period, an economy is included only if data are available for at least ten years because data for shorter periods would often be unrepresentative.

### *Political Transitions*

Instability of the executive branch of government seems most relevant for central bank autonomy. Consequently, we selected four types of political change as indicators of political instability: change of the head of government, change of the party in government, change of the fundamental rules of government as embodied in switches from authoritarian to democratic regimes or vice versa, and irregular changes of government from one authoritarian ruler to another. These types of changes form a hierarchy, so that each political event is coded as a level of instability: a change of the head of government without a change of party (low instability); a nonviolent change of party without a change of type of government (medium instability); an irregular change of authoritarian government without a change of form of government—a coup—(type 2 authoritarian); or a change of the form of government between democratic and authoritarian—a coup against a democratic government or a restoration of democracy—(high instability).[1]

Our previous research indicated that the autonomy of the central bank differed markedly between industrial and developing countries (Cukierman, Webb, and Neyapti 1992). Here we examine whether this distinction is also important for the effects of political instability on the central bank. Within each economic group, countries are divided into subgroups with democratic regimes for the whole period, with authoritarian regimes throughout, and with regimes that alternate between democratic and authoritarian. We classified as democratic only the countries and periods in which the head of government was chosen as a result of competitive elections. All others were classified as authoritarian, covering a variety of systems in which the government did not face serious pressure from electoral politics. The classifications thus differ somewhat from what one would

1. Data come from and were checked against several sources. The most comprehensive single source, and usually the initial one, is *The Europa World Yearbook*. Changes of only the economic team (not necessarily including the central bank) are also events that may be relevant for assessing central bank autonomy. To investigate the relevance of these events, we have begun collecting data on changes of the Minister of Finance or Treasury. These data are less readily available than information on changes of the head of government, and we do not yet have information on a broad enough sample to report the results.

code as the degree of political liberty, such as with Gastil's indexes. We chose to exclude from our data political events that challenged and perhaps threatened to change a government but did not do so—strikes, riots, attempted coups, and elections in which the previous government was reelected. To assess the implications for an institution such as the central bank, it seemed better to focus on institutionally well-defined events that actually changed the government to which the bank reported.

### *Changes at the Central Bank*

The raw data on instability at the central bank consist of the actual dates of changes of the governors of the central banks in our sample of countries. The turnover of governors is only an imperfect indicator of actual central bank independence in developing countries (for more discussion of this proxy, see Cukierman, Webb, and Neyapti 1992; Cukierman 1992; and Cukierman and others 1993). Low turnover does not always imply a high level of central bank independence—relatively subservient governors may stay in office longer precisely because they do not stand up to the executive branch. This may be true for countries with exceptionally low turnover rates such as Denmark, Iceland, and the United Kingdom and for countries with stable authoritarian governments. In such countries, low turnover is probably unrelated to independence. Above some threshold turnover rate, however, higher turnover of the governor reflects lower central bank independence.

Sufficiently high turnover rates make the tenure of the central bank governor shorter than that of the executive branch and thus make the governor more susceptible to influence by the executive branch and more discouraged from trying to implement longer-term policies. Because the electoral cycle is at least four years in most countries, the threshold turnover rate is probably between 0.2 and 0.25 (implying a governor's tenure of four to five years). In addition, governors with very short terms of office, such as three years or less, generally have more difficulty in implementing long-term policies (for example, the maintenance of price stability) for any electoral cycle. One limitation of the turnover variable is that all the industrial countries have turnover rates at or below the threshold rate in the years we examine. Turnover rates in developing countries, however, span a range that goes well above the threshold point.

The measures of political influence on the central bank developed here consider the links between political instability and subsequent turnover at the central bank and hence appear to be relevant for both industrial and developing economies. They also make it possible to distinguish between the frequencies of changes at the central bank in political periods, as defined earlier, and in nonpolitical periods. A relatively high frequency of turnover in the political periods indicates substantial influence of political instability on the central bank. A relatively high level of turnover the rest of the time most likely indicates that the central bank is more dependent even in politically tranquil times.

## II. Political Change and the Timing of Central Bank Turnover

To what extent does political instability translate into instability at the central bank? Are there systematic differences in central bank turnover between periods immediately following a political change and nonpolitical periods? If there are differences, how does their magnitude vary with country characteristics? To answer these questions, this section measures political instability by the frequency of political transitions, as defined in section I, and instability at the central bank by the turnover of the governors.

Before attempting to give precise answers to these questions, it is instructive to take a broad look at the data. Table 1 presents central bank turnover figures (measured as the number of changes of central bank governor per month) for various intervals following a political transition. Thus, the average frequency of changes of governor within one month following a political transition is 0.063 per month; within two to three months it is 0.026, and it is only 0.015 at ten months or more after a political transition.[2] These numbers correspond to governors' average terms of office of 1.3, 3.2, and 5.6 years, respectively.[3]

The numbers in table 1 can be interpreted as estimates of the probability per month of a change in central bank governor conditional on being within a time interval, $i$, that follows a political transition by $i$ to $i + 2$ (or by $i$ to $i + 3$) months.[4]

Table 1. *Frequency of Change of Central Bank Governor at Various Intervals by Economy Group, 1950–89*
(number of changes per month)

| *Economy group* | *Number of economies* | *Interval (number of months since a political transition)* | | | | |
|---|---|---|---|---|---|---|
| | | *0–1* | *2–3* | *4–6* | *7–9* | *10 or more* |
| All | 67 | 0.063 | 0.026 | 0.025 | 0.013 | 0.015 |
| *Industrial* | | | | | | |
| All 20 | 24 | 0.008 | 0.013 | 0.013 | 0.003 | 0.009 |
| Democratic only | 18 | 0.020 | 0.009 | 0.015 | 0.002 | 0.009 |
| Mixed | 2 | 0.060 | 0.000 | 0.000 | 0.024 | 0.013 |
| *Developing* | | | | | | |
| All | 47 | 0.096 | 0.041 | 0.034 | 0.022 | 0.018 |
| Democratic only | 9 | 0.046 | 0.016 | 0.065 | 0.011 | 0.025 |
| Authoritarian only | 16 | 0.089 | 0.017 | 0.006 | 0.025 | 0.015 |
| Mixed | 22 | 0.105 | 0.053 | 0.038 | 0.023 | 0.019 |

*Source:* Authors' calculations.

2. When a change in central bank governor occurs within a short time after two or more previous political transitions, the central bank change is attributed only to the most recent political transition.

3. In other words, let $x$ be a turnover number from table 1; then, the corresponding average term in office, in years, is given by $1/12x$.

4. This interpretation requires the probability of two or more changes within a subperiod to be negligible. Because the time intervals considered are only two or three months, this assumption is supported by the data.

Table 1 reveals that, for the entire sample of economies, this probability decreases monotonically with the number of months that have elapsed since the last political transition. It is almost three times larger in the zero-to-one-month interval than in the two-to-six-month intervals. For seven-month intervals and beyond, this probability drops further—to about 60 percent of its value in the two-to-six-month intervals.

Central bank changes within one month of a political transition are highly likely to result directly from the political change, but central bank changes more than nine months after a political transition are unlikely to be the result of a change in government and are more likely to be largely nonpolitical. The challenge is to find the average elapsed time across countries, presumably between one and nine months, at which to set the cutoff between political and nonpolitical periods. We take the probability of turnover long after a political change (ten months or more) as a background rate against which to compare months in the intermediate range. The political period would then be defined to include months at intervals (after a political change) for which the average rate of central bank turnover is significantly higher than the background rate.

The appendix describes the tests for differences between the probability of a change of central bank governor ten or more months after a political change and the probability in each of the intervals with shorter lags, using the normal approximation to the binomial distribution. It shows that for the whole sample the probability of a turnover at the central bank is significantly larger in the zero-to-one-month, two-to-three-month, and four-to-six-month intervals than in the ten-month-and-more interval. The average probability of a change of central bank governor during the seven-to-nine-month interval following a political change does not differ significantly from the probability in the ten-or-more lag period. Hence, the evidence in the appendix supports a choice of six months following a political transition as the cutoff between political and nonpolitical periods in the overall sample.

In summary, for the entire sample of economies, the evidence supports the view that instability at the central bank rises in periods following political transitions and that this increase is significant. The estimated probability of a change of governor at the central bank is more than two times larger in periods within six months after a political transition than in periods that are more removed from political change. Thus, political instability increases instability at the central bank and weakens its independence from political authorities.

The profile of estimated probabilities of a change at the central bank varies between industrial and developing economies and between economies with different political regimes. We distinguish three types of political regimes: stable democracies, stable authoritarian regimes, and mixed regimes (those alternating between democratic and authoritarian regimes). Most industrial economies are stable democracies, none are authoritarian-only for the period under investigation, and only two (Spain and France) are mixed regimes. Developing

economies include all three political regimes—democratic only, authoritarian only, and mixed.

Table 1 also presents the estimated probability of a change of central bank governor at various intervals following a political transition for the different subgroups of economies. The probability of central bank turnover is higher in developing economies than in industrial countries at all intervals . The frequency declines monotonically for developing economies but has two peaks for industrial economies—one in the zero-to-one-month interval and a lower peak in the four-to-six-month interval. The tests reported in the appendix reveal that for developing economies the appropriate cutoff between political and nonpolitical periods is at six months as in the overall sample, and there is clear evidence of a link between political instability and instability at the central bank.

The average frequencies of central bank turnover within the three political subgroups of developing economies are summarized in the lower part of table 1. In the authoritarian and mixed subgroups, the two largest, the frequency in the zero-to-one-month interval is substantially higher than in the subsequent intervals. Statistical tests reveal that in all three cases the difference in frequencies between the zero-to-one-month interval and the ten-month-or-more interval is significant. Comparing the three subgroups, the average frequency of central bank instability in the zero-to-one-month and the two-to-three-month intervals is highest on average in the countries with mixed regimes, second highest in authoritarian-only countries, and lowest in democratic-only countries.

The evidence is less dramatic for industrial economies, but still supports the view that in those countries the average frequency of central bank changes in the zero-to-one-month interval after a political change is significantly higher than the background frequency in the ten-month-or-more interval. The frequencies in the two-to-three-month, four-to-six-month, and seven-to-nine-month intervals are not significantly different from the frequency for the ten-month-or-more lag. Thus, for industrial economies as a subgroup we locate the cutoff between political and nonpolitical periods after the zero-to-one-month interval. We present most of the results in the rest of the article for both the zero-to-one-month and zero-to-six-month cutoffs.

The tendency of mixed regimes to have higher frequencies of central bank changes than democracies in the zero-to-one-month interval also appears within the industrial countries, although this finding is based only on the experience of France and Spain. Also, as was the case for industrial democracies, there is a second peak in the four-to-six-month interval for democratic-only developing economies.

## III. Political Vulnerability of Central Banks

Computing the frequency of central bank turnover in intervals at various lags from political changes was important for assessing the size and duration of the effect of political changes on the propensity for central bank turnover. For a

summary measure of political influence, however, we compute an index of the political vulnerability of the central bank, defined for each country as the fraction of political transitions that are followed promptly by a replacement of the central bank governor:

$$V(i) \equiv \frac{\text{Number of replacements of the central bank governor within } i \text{ months following a political transition}}{\text{Number of political transitions}}, \quad i = 1, 6.$$

Table 2 reports average values of the index of central bank political vulnerability for industrial and developing economies as well as for democratic, authoritarian, and mixed regimes within each group. The overall average value of vulnerability is 0.24. That is, almost a quarter of all political transitions are followed by a replacement of the central bank governor within six months. As with the frequency of turnover, the vulnerability varies widely across country groups and subgroups. Political vulnerability is more than three times larger in developing economies than in industrial ones. A similar picture (not shown) emerges when the sample is broken into two subperiods (1950–71 and 1972–89). Central banks of developing economies with mixed regimes are the most vulnerable on average.

Table A-1 in the appendix presents the vulnerability of central banks to all types of political transitions for individual economies by subgroups. For economies with a small number of political transitions in a subperiod, the vulnerability ratios are highly sensitive to the effect of a truly nonpolitical central bank

Table 2. *Political Vulnerability of Central Banks by Economy Group, 1950–89*

| *Economy group* | *Vulnerability*[a] | | *Frequency of political change (per year)* |
|---|---|---|---|
| | *Within six months* | *Within one month* | |
| *All* | 0.24 | 0.12 | 0.27 |
| Industrial | 0.10 | 0.05 | 0.32 |
| Developing | 0.35 | 0.18 | 0.24 |
| *Industrial* | | | |
| Democratic only | 0.10 | 0.04 | 0.32 |
| Mixed | 0.12 | 0.12 | 0.33 |
| *Developing* | | | |
| Authoritarian only | 0.22 | 0.17 | 0.14 |
| Democratic only | 0.30 | 0.09 | 0.18 |
| Mixed | 0.39 | 0.20 | 0.30 |

a. Average share of political transitions followed by a replacement of the central bank governor within the noted period.

*Source*: Authors' calculations.

change accidentally happening after a political change.[5] Although vulnerability numbers for individual economies in such cases should be viewed cautiously, they are still useful for the overall statistical analysis. Within each subgroup, central banks are arranged in descending order of their political vulnerability in 1972–89. By this measure, Sweden has the most politically vulnerable central bank among industrial democracies in that subperiod. Among democratic-only developing economies, the central bank of Botswana is the most vulnerable. Among developing economies with mixed regimes, the Argentine central bank is the most vulnerable, with sometimes more than one change of governor in the first six months after a political change.[6]

The last column of table 2 reports the average frequency of political transitions per year for each country group. Industrial economies have political transitions more frequently than developing economies, because they have much more frequent democratic changes of government, with and without party changes. Within developing economies, the subgroup with mixed regimes has a frequency of political transitions that is about twice that of either democratic-only or authoritarian-only regimes. Authoritarian-only regimes have the lowest number of political changes of any type.

We now have four indicators of central bank autonomy: the two developed in Cukierman, Webb, and Neyapti (1992)—the index of legal autonomy and the (total) turnover rate of the governor—plus vulnerability and its complement, the frequency of nonpolitical turnover of the governor. Of course the latter two can be reaggregated into the total turnover rate by making appropriate adjustments of units and multiplying vulnerability by the frequency of political transitions in each country. How different are these four measures? Table 3 shows the matrix of correlation coefficients for the four measures and the frequency of political turnover. Most of the various indicators are not closely correlated to one another, although total turnover is closely correlated with its two derivatives. The frequency of political transitions, an indicator of vibrant democracy or of more-fundamental political instability, is not significantly correlated with the various central bank variables.

How is the political vulnerability of the central bank affected by the type of political instability in a country? To answer this question, we regressed the indexes of central bank vulnerability at lags of zero to one month and zero to six months on various types of political transitions, on a dummy for developing countries, and on a dummy for purely authoritarian regimes. Table 4 summarizes the results. With the vulnerability index for the lag of zero to six months,

5. This has been the case at least once in Jamaica and once in the Netherlands. In the Netherlands the decision to replace the governor in 1967 was made prior to the adjoining political transition, but was inplemented after it (de Haan 1995). Thus, vulnerability is a noisy measure of political influence whose quality rises with the number of political transitions in a country.

6. In the 1990s, Argentina has vastly increased its commitment to price stability by upgrading the autonomy of its central monetary institution and by endowing it with the authority to function as a currency board.

Table 3. *Correlation between Various Indicators of Central Bank Autonomy, 1950–89*

| *Indicator* | *Index of legal central bank autonomy* | *Total central bank turnover* | *Vulnerability* | *Frequency of political change* |
|---|---|---|---|---|
| Total central bank turnover | –0.05 | | | |
| Vulnerability (six months) | –0.11 | 0.78 | | |
| Frequency of political transitions | –0.05 | 0.06 | –0.11 | |
| Nonpolitical central bank turnover | –0.02 | 0.88 | 0.60 | –0.21 |

*Source:* Authors' calculations.

the developing-country dummy has a significant positive impact on central bank vulnerability. But none of the indicators of political change had a statistically significant effect on vulnerability. Low- and medium-level changes clearly have no effect, which is consistent with the ambiguous theoretical priors. For high-level political change, the coefficient has a positive sign, as expected, but is not quite significant, although it was with some earlier versions of the sample.

The vulnerability of central banks to political influence depends on the type of political transition. Table 5 reports the indexes of political vulnerability by

Table 4. *The Impact of Political Change on Central Bank Vulnerability, 1950–89*

| *Explanatory variable* | *Estimates* | |
|---|---|---|
| | *For lag of 0–1 month* | *For lag of 0–6 months* |
| Constant | 0.09 | 0.16** |
| | (1.61) | (2.16) |
| High-level political change | 0.73 | 0.96 |
| | (1.46) | (1.44) |
| Type-2 authoritarian transitions | –0.16 | –0.47 |
| | (–0.29) | (–0.65) |
| Medium-level political change | –0.15 | –0.22 |
| | (–0.92) | (–1.01) |
| Low-level political change | –0.09 | –0.19 |
| | (–0.52) | (–0.78) |
| Dummy for purely authoritarian regimes | 0.04 | –0.08 |
| | (0.55) | (–0.79) |
| Dummy for developing countries | 0.09 | 0.24*** |
| | (1.48) | (2.96) |
| Number of observations | 110 | 110 |
| Adjusted $R^2$ | 0.059 | 0.143 |

** Significant at the 5 percent level

*** Significant at the 1 percent level.

*Note:* The estimated equation is a pooled cross–sectional time–series regression in which there are two observations for each country, one for 1950–71 (where available) and one for 1972–89. The dependent variable is central bank vulnerability. *t*-statistics are in parentheses.

*Source:* Authors' calculations.

type of political transition for different economy groupings. The fraction 0.61 at the intersection of the mixed developing-economy row with the "high-level" column means that 61 percent of high-level political transitions in developing economies are followed within six months by a replacement of the central bank governor. The other numbers in the table are defined in a similar manner. There is generally little difference between the vulnerability to medium-level and to low-level political transitions within each country group. The vulnerability to medium- and low-level political transitions is more than twice as high in developing economies as it is in industrial ones. Furthermore, this difference seems to be due to a difference in the level of development, rather than to differences in regimes, because in democratic developing economies vulnerability to medium- and low-level political transitions is more than twice as large as that of industrial democracies. The vulnerability to these types of political transitions in mixed-regime developing economies is of the same order of magnitude as the vulnerability in democratic developing economies.

The highest level of central bank vulnerability occurs in the face of high-level political transitions, which are all in developing economies (with mixed regimes, by definition). For developing economies as a whole, central bank vulnerability to type-2 authoritarian transitions (0.46) is larger than vulnerability to medium- and low-level transitions by a factor of almost two. The 0.46 figure, however, masks quite different tendencies in authoritarian-only regimes, which have an average vulnerability of 0.20, and in mixed regimes, where type-2 authoritarian transitions lead to an average vulnerability of 0.55.

Table 5. *Central Bank Vulnerability (within Six Months) by Type of Political Transition and Economy Group*

| | *Type of political transition* | | | |
|---|---|---|---|---|
| *Economy group* | *High level*[a] | *Type-2 authoritarian*[b] | *Medium level*[c] | *Low level*[d] |
| *All* | 0.58 | 0.46 | 0.15 | 0.19 |
| Industrial | 0.00 | n.a. | 0.11 | 0.08 |
| Developing | 0.61 | 0.46 | 0.24 | 0.26 |
| *Industrial* | | | | |
| Democratic | n.a. | n.a. | 0.11 | 0.08 |
| Mixed | 0.00 | n.a. | 0.12 | 0.17 |
| *Developing* | | | | |
| Authoritarian only | n.a. | 0.20 | n.a. | 0.23 |
| Democratic only | n.a. | n.a. | 0.24 | 0.25 |
| Mixed | 0.61 | 0.55 | 0.24 | 0.28 |

n.a. Not applicable, because there were no political transitions of that type for that category of country.

a. Change of regime, from democratic to authoritarian or vice versa.

b. Replacement of one authoritarian regime by another one.

c. Change of party without a change in regime.

d. Change of head of government without a change in regime or party.

*Source:* Authors' calculations.

Table 6. *Estimates for Central Bank Vulnerability and Inflation, 1950–89*

| *Explanatory variable* | *Estimation with six-month vulnerability index* | | *Estimation with one-month vulnerability index* | |
|---|---|---|---|---|
| | *Dependent variable is transformed inflation, D* | *Dependent variable is standard deviation of D* | *Dependent variable is transformed inflation, D* | *Dependent variable is standard deviation of D* |
| Constant | 0.046* | 0.019 | 0.054** | 0.020 |
| | (1.73) | (1.28) | (2.16) | (1.53) |
| Vulnerability | — | — | 0.164*** | 0.106*** |
| (lag 0–1 months) | | | (3.95) | (4.79) |
| Vulnerability | 0.092*** | 0.070*** | — | — |
| (lag 0–6 months) | (2.89) | (4.02) | | |
| Nonpolitical turnover | — | — | 0.158** | 0.108*** |
| (lag over 1 month) | | | (2.12) | (2.73) |
| Nonpolitical turnover | 0.239*** | 0.108** | — | — |
| (lag over 6 months) | (2.60) | (2.14) | | |
| High-level | 0.312 | 0.237** | 0.208 | 0.193* |
| political change | (1.50) | (2.08) | (1.04) | (1.83) |
| Type-2 authoritarian | 0.308 | 0.190 | 0.268 | 0.167 |
| transitions | (1.38) | (1.56) | (1.24) | (1.45) |
| Medium-level | 0.044 | 0.010 | 0.029 | 0.003 |
| political change | (0.65) | (0.27) | (0.44) | (0.09) |
| Low-level | 0.126* | 0.041 | 0.109 | 0.028 |
| political change | (1.72) | (1.02) | (1.53) | (0.74) |
| Dummy | 0.030 | 0.023 | 0.016 | 0.017 |
| authoritarian only | (0.98) | (1.37) | (0.52) | (1.08) |
| Dummy first period | −0.087*** | −0.021** | −0.083*** | −0.019* |
| (1950–1971) | (−4.53) | (−2.02) | (−4.44) | (−1.95) |
| Dummy | −0.007 | −0.005 | 0.006 | −0.001 |
| developing countries | (−0.26) | (−0.32) | (0.22) | (−0.10) |
| Number of observations | 110 | 110 | 110 | 110 |
| Adjusted $R^2$ | 0.34 | 0.34 | 0.37 | 0.41 |

—Not available.
* Significant at 10 percent level.
** Significant at 5 percent level.
*** Significant at 1 percent level.
*Note:* *t*-statistics are in parentheses.
*Source:* Authors' calculations.

## IV. Effects of Central Bank Vulnerability on the Economy

Besides being of independent interest, measures of central bank autonomy are useful for testing the effect of this autonomy, or its absence, on the economy. Such a project is largely beyond the scope of this article. Nevertheless, this section and the next briefly report evidence on the relation between some of our measures and the performance of the economy as reflected by inflation, real growth, and real interest rates.

Each of these variables is regressed in a pooled cross-sectional time series on central bank vulnerability, on nonpolitical turnover, and on other control variables. Nonpolitical turnover is measured as the average number of nonpolitical changes of central bank governor (more than one month or more than six months after the latest political transition) per year.

### *Inflation*

Previous work has shown a significant positive relationship between inflation or the rate of depreciation in the real value of money ($D$) and the (total) turnover of central bank governors for developing economies (Cukierman, Webb, and Neyapti 1992: tab. 7; Cukierman 1992: tab. 20.2).[7] Table 6 here extends the effort by relating $D$ and its standard deviation to the index of vulnerability, nonpolitical turnover, various types of political instability, and three dummies: one for the Bretton Woods era, one for having only authoritarian regimes in the period, and one for being a developing country. This formulation makes it possible to evaluate the relative importance of lack of central bank autonomy and of political instability for the creation of inflation. Table 6 shows results with the zero-to-one-month and zero-to-six-month measures of vulnerability, and the results are very similar for the two different measures. The regressions in table 6 back the view that the first channel is more important.

The political vulnerability of the central bank as well as nonpolitical turnover at the bank have a positive and significant impact on inflation ($D$) and its variability. Political instability, particularly of the high-level sort, contributes to increased inflation variability. Low-level political instability has a marginally significant effect in raising the level of inflation. Medium-level political instability—alternation between parties (in a democracy)—has essentially zero effect on inflation, probably because frequent transitions of this sort require fundamental consensus on policy issues. Also, worldwide inflation is lower during the Bretton Woods period, even after allowing for the effect of other variables. This is consistent with the view that fixed exchange rates have a stronger disciplinary effect on policy than flexible rates do. But the higher worldwide inflation after 1972 may also be the result of the larger shocks that affected the world economy in the post–Bretton Woods period. The insignificant authoritarian-only dummy means that having an authoritarian rather than a democratic

7. $D = \pi/(1 + \pi)$, where $\pi$ is the annual inflation rate.

regime throughout does not help to explain differences in average inflation, once the central bank and political variables are taken into account. A dummy is not needed for mixed-regime countries, because the variable for high-level political changes is nonzero only in mixed-regime countries.

The most important result in table 6 is the insignificance of the dummy variable for developing economies. Inflation is significantly higher on average in developing economies, and this could not have been accounted for by just looking at the overall frequency of turnover at the central bank (Cukierman, Webb, and Neyapti 1992). Once turnover is disaggregated into its constituent parts, however, and account is taken of the different types of political instability, the distinction between developing and industrial economies no longer contributes to explaining differences in inflation outcomes. In other words, the differences in vulnerability of the central bank to political instability, in central bank turnover in nonpolitical periods, and in political instability can fully account for the developing economies having higher average inflation than industrial economies.

### *Growth*

Recent empirical literature on growth has identified several variables, such as initial gross domestic product (GDP) and education, that are significantly related to real growth (see Barro 1991, for example). Does the degree of anti-inflation commitment by the monetary authorities, for which our indexes of central bank autonomy are a proxy, have any significant impact on growth after controlling for these variables? Table 7 answers this question by presenting growth equations that take account of initial GDP, initial primary and secondary education enrollment ratios, a decade-by-decade change in terms of trade, and our indexes of central bank autonomy. With a full sample of countries, nonpolitical turnover of central bank governors has a marginally significant positive sign, contrary to priors. Brazil, Korea, and Botswana are outliers, however, because they achieved high growth despite high rates of central bank turnover and high vulnerability. With those countries excluded, the six-month vulnerability indicator has a significantly negative sign.[8] This finding supports the view that, other things being equal, higher political dependence of the central bank tends to retard growth in most countries. It is possible that political vulnerability of the central bank is a proxy for general economic and political instability, both of which deter growth, possibly by slowing down investment and innovation. There is some evidence that in developing countries, higher central bank vulnerability is associated with lower levels of investment (Cukierman and others 1993). Sorting out the channels through which political vulnerability of the central bank affects growth clearly deserves further work.

8. Botswana could legitimately be discarded from the sample, because its growth was primarily due to the discovery of diamonds, not to good macroeconomic policy. Brazil and Korea seem to be simply countries where lack of central bank autonomy was not a hindrance to growth, at least not in the period covered here. Further justification for the exclusion of these countries appears in Cukierman and others (1993).

Table 7. *The Impact of Central Bank Vulnerability and Nonpolitical Turnover on Economic Growth, 1960–88*

| *Explanatory variable* | *Estimate* | |
|---|---|---|
| | *Full sample* | *Sample minus three*[a] |
| Constant | −0.15 | 0.73 |
| | (−0.15) | (0.77) |
| Initial GDP | −0.22** | −0.25*** |
| | (−2.46) | (−2.83) |
| Change in terms of trade | 28.9*** | 28.1*** |
| | (4.87) | (5.01) |
| Initial primary education enrollment ratio | 2.03** | 2.53*** |
| | (1.98) | (2.55) |
| Initial secondary education enrollment ratio | 1.59 | 1.34 |
| | (1.22) | (1.04) |
| Nonpolitical turnover of central bank governors | 5.80* | −2.39 |
| | (1.80) | (−0.66) |
| Political vulnerability of central bank | −0.78 | −1.51** |
| | (−1.15) | (−2.30) |
| Dummy for the 1960s | 1.69*** | 1.42*** |
| | (3.01) | (2.61) |
| Dummy for the 1970s | 1.39*** | 1.11** |
| | (2.74) | (2.26) |
| Number of observations | 129 | 120 |
| Adjusted $R^2$ | 0.23 | 0.26 |

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

*Note:* The sample consists of, at most, three observations (one for each decade) for each country. The estimated equation is a pooled cross-sectional time-series regression. The dependent variable is real per capita GDP growth. Central bank vulnerability is characterized in terms of changes in central bank governors that occur within six months of a political transition. *t*-statistics are in parentheses.

a. Botswana, Brazil, and the Republic of Korea are excluded.

*Source:* Authors' calculations.

### *Real Interest Rates on Deposits*

Previous evidence (Alesina and Summers 1993; Cukierman and others 1993) suggests there is a negative relation between the variability of ex post real interest rates and central bank independence. In our work, we have used as proxies for central bank independence the bank's legal independence for industrial economies and the bank's turnover of governors for developing economies. We now reexamine the relation between central bank independence and real interest rates by using central bank vulnerability and nonpolitical turnover as proxies for a lack of central bank independence. The second column of table 8 presents the effect on the variability of real rates. It appears that both central bank vulnerability and nonpolitical turnover significantly increase the variability of ex post real deposit rates. This further supports and amplifies the conclusion that the variability of ex post real deposit rates is lower in countries with more independent central banks. Although it has the expected negative sign, the coefficient on legal independence is not significant.

Table 8. *The Impact of Alternative Measures of Central Bank Autonomy on the Ex Post Real Deposit Rate*

| | Estimate | |
|---|---|---|
| *Explanatory variable* | *Dependent variable is real ex post deposit rate, R* | *Dependent variable is standard deviation of R* |
| Constant | 1.98 | 1.04 |
| | (0.64) | (0.3) |
| Vulnerability | −5.56* | 8.68** |
| (lag 0–6 months) | (−1.89) | (2.1) |
| Nonpolitical turnover | −1.98 | 36.24*** |
| (lag more than 6 months) | (−0.15) | (3.0) |
| Legal independence | −3.26 | −5.60 |
| of central bank | (−0.48) | (−0.7) |
| Number of observations | 34 | 34 |
| Adjusted $R^2$ | 0.05 | 0.52 |

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.
*Note:* *t*-statistics are in parentheses.
*Source:* Authors' calculations based on data from Easterly, Rodriguez, and Schmidt-Hebbel (1992), Cukierman, Webb, and Neyapti (1992), and Cukierman (1992: ch. 19).

The first column in table 8 relates the average level of the ex post real deposit rate to measures of central bank independence. The higher the political vulnerability of the central bank, the lower the average real deposit rate, which corroborates a similar finding in Cukierman and others (1993). The effect of lowering the real deposit rate probably reflects the higher implicit taxation of financial savings in countries with more politically dependent central banks.

## V. Central Bank Independence and the Mean and the Variability of Inflation

The strong cross-country association between the mean and the variability of inflation is a well-known empirical regularity. When legal independence and the total turnover of central bank governors are used as proxies for central bank independence, up to a quarter of this association is accounted for by their common association with central bank independence (Cukierman 1992: ch. 18, 22). For this article, the same experiment was repeated with central bank vulnerability and nonpolitical turnover as indexes of central bank autonomy. Legal independence was added as a regressor in industrial economies. To calculate the fraction of the correlation between the mean and the standard deviation of inflation that results from their common association with central bank independence, we proceeded as follows. First, the cross-sectional covariance between $D$ (the rate of depreciation in the real value of money) and its standard deviation was calculated. Second, both $D$ and its standard deviation were regressed on central bank vulnerability, nonpolitical turnover, and (for industrial economies)

legal independence. The predicted values of $D$ and of its standard deviation were also calculated. Third, the covariance between those predicted values was calculated and compared with the overall covariance between $D$ and its standard deviation.

The experiment was done for the entire 1950–89 period and for the subperiods 1950–71 and 1972–89. The fractions of the covariability between $D$ and its standard deviation that are the result of their mutual link to central bank independence are 0.30 for the whole period and 0.40 and 0.31 for the two subperiods. Thus, with the more refined measures of central bank autonomy presented in this article, about one-third of the association between inflation and its variability is attributable to their common association with central bank independence.

## VI. Concluding Remarks

Stability and other characteristics of government institutions have always been recognized in economic history and in country studies as crucial determinants of macroeconomic stability. Recently this recognition has spread to theoretical work on macroeconomics and has been the focus of some cross-country statistical analysis (Fischer 1991; Edwards and Tabellini 1991; Cukierman, Edwards, and Tabellini 1992; Alesina and others 1995; Haggard, Kaufman, and Webb 1992). A survey appears in Alesina and Perotti (1994). As one would expect, political instability is positively related to inflation and negatively related to growth. There are various theories for how political instability causes macroeconomic instability, most of them not mutually exclusive. One contender is that political instability shortens time horizons of policymakers and that it decreases the ability of the political system to efficiently resolve disputes over real incomes.

This article has demonstrated that political instability causes instability at the central bank as well. But the spillover from political instability to instability at the central bank varies across country groups and types of political transitions. It is particularly large when the political regime changes from democratic to authoritarian or vice versa. This finding supports the view that when political change is deep enough to involve fundamental rules of the game and in other circumstances where political change would probably mean the party in power would not be back soon, if ever, then the expected effect of greater political instability would be shortened time horizons, as discussed above, and therefore less autonomy and stability for the central bank. Each new government would want to use the central bank to try to stay in power as long as possible and would have little concern for the associated detrimental longer-run effects. However, greater frequency of low- and medium-level political change, such as just changing the head of government or the party, does not on average lead to reduced central bank autonomy as proxied by its political vulnerability. Actually, in democratic governments in which the party changes frequently, the ruling party might typically lack the strength to impose its will unilaterally and

might thus agree to a compromise that would endure changes of party. This was explicit in the setup of the reformed Chilean central bank in 1989 (Arriagada and Graham 1994). An analysis of the effect of political instability on central bank independence in countries with democratic and stable rules of the political process appears in Cukierman (1994).

An important issue for future work is the possibility of reverse causality between the performance of the economy and our measures of gross political influence on the central bank. The largely cross-sectional nature of the variables in this article precludes the use of Granger-Sims and other methods to test the exogeneity of our measures of political instability with respect to the performance of the economy. Hence, strictly speaking, it is possible that some of the significant relationships between these two groups of variables are due to causality running from the economy to political influence. This is probably less likely to be the case for real interest rates than for real growth. There is evidence that real growth is affected by political instability (Barro 1991; Alesina and others 1995) which may, in turn, cause instability at the central bank.

The measure of central bank vulnerability, however, is more likely to be exogenous with respect to the economy than other behavioral indexes of political influence on the central bank. Even if political instability responds to the performance of the economy, vulnerability—defined as the ratio between instability at the central bank and political instability—depends on slowly changing institutions and sociopolitical norms, and thus may be largely exogenous to the current economic performance.

Because the main contribution of this article is in the extraction of empirical regularities by matching two new data sets, our approach has been inductive rather than deductive. We deliberately avoided committing to and testing a particular model because we believe that at this early stage the broad regularities in the data can be uncovered more efficiently without positing a particular model. Our hope is that the regularities uncovered here will encourage the construction of more precise models and further empirical testing of hypotheses.

Table A-1. *Political Vulnerability of Central Banks for Individual Economies by Subgroup, 1950–71 and 1972–89*

| Economy | 1950–71 | | | | 1972–89 | | | |
|---|---|---|---|---|---|---|---|---|
| | *Vulnerability (within six months)* | *Number of political transitions* | *Number of central bank turnovers* | | *Vulnerability (within six months)* | *Number of political transitions* | *Number of central bank turnovers* | |
| | | | *Total* | *Nonpolitical*[a] | | | *Total* | *Nonpolitical*[a] |
| *Industrial, democratic* | | | | | | | | |
| Sweden | 0.000 | 1 | 2 | 2 | 0.500 | 4 | 4 | 2 |
| Ireland | 0.000 | 5 | 3 | 3 | 0.286 | 7 | 3 | 1 |
| Japan | 0.200 | 5 | 4 | 3 | 0.222 | 9 | 4 | 2 |
| Belgium | 0.000 | 9 | 2 | 2 | 0.167 | 6 | 3 | 2 |
| Switzerland | 0.059 | 17 | 2 | 1 | 0.167 | 18 | 3 | 0 |
| Finland | 0.105 | 19 | 3 | 1 | 0.125 | 8 | 2 | 1 |
| Italy | 0.071 | 14 | 1 | 0 | 0.077 | 13 | 2 | 1 |
| Australia | 0.333 | 3 | 1 | 0 | 0.000 | 3 | 3 | 3 |
| Austria | 0.000 | 4 | 3 | 3 | 0.000 | 2 | 3 | 3 |
| Canada | 0.000 | 3 | 2 | 2 | 0.000 | 4 | 2 | 2 |
| Denmark | 0.143 | 7 | 2 | 1 | 0.000 | 4 | 0 | 0 |
| Germany, Fed. Rep. | 0.333 | 3 | 2 | 1 | 0.000 | 2 | 2 | 2 |
| United Kingdom | 0.000 | 6 | 2 | 2 | 0.000 | 3 | 2 | 2 |
| Iceland | 0.000 | 3 | 1 | 1 | 0.000 | 7 | 0 | 0 |
| Netherlands | 0.143 | 7 | 1 | 0 | 0.000 | 3 | 1 | 1 |
| New Zealand | 0.250 | 4 | 2 | 1 | 0.000 | 5 | 4 | 4 |
| Norway | 0.000 | 6 | 2 | 2 | 0.000 | 7 | 1 | 1 |
| United States | 0.000 | 4 | 2 | 2 | 0.000 | 4 | 3 | 3 |
| *Industrial, mixed* | | | | | | | | |
| France | 0.056 | 18 | 2 | 1 | 0.500 | 2 | 4 | 3 |
| Spain | n.a. | 0 | 5 | 5 | 0.167 | 6 | 3 | 2 |
| *Developing, authoritarian* | | | | | | | | |
| Egypt | 0.000 | 3 | 8 | 8 | 1.000 | 1 | 4 | 3 |
| Mexico | 0.250 | 4 | 2 | 1 | 0.667 | 3 | 4 | 2 |
| China | — | — | — | — | 0.500 | 2 | 4 | 3 |

*(Table continues on the following page.)*

Table A-1. *(continued)*

| Economy | 1950–71 | | | | 1972–89 | | | |
|---|---|---|---|---|---|---|---|---|
| | *Vulnerability (within six months)* | *Number of political transitions* | *Number of central bank turnovers* | | *Vulnerability (within six months)* | *Number of political transitions* | *Number of central bank turnovers* | |
| | | | *Total* | *Nonpolitical*[a] | | | *Total* | *Nonpolitical*[a] |
| South Africa | 0.000 | 4 | 2 | 2 | 0.500 | 2 | 3 | 2 |
| Taiwan (China) | 0.000 | 3 | 2 | 2 | 0.500 | 4 | 2 | 0 |
| Ethiopia | n.a. | 0 | 0 | 0 | 0.400 | 5 | 5 | 3 |
| Hungary | — | — | — | — | 0.333 | 3 | 2 | 1 |
| Yugoslavia (former) | n.a. | 0 | 5 | 5 | 0.100 | 10 | 4 | 3 |
| Kenya | n.a. | 0 | 2 | 2 | 0.000 | 1 | 2 | 2 |
| Morocco | 0.167 | 6 | 4 | 3 | 0.000 | 4 | 2 | 2 |
| Romania | — | — | 0 | 0 | 0.000 | 1 | 3 | 3 |
| Tanzania | n.a. | 0 | 1 | 1 | 0.000 | 1 | 2 | 2 |
| Uganda | — | — | 0 | 0 | 0.000 | 4 | 2 | 2 |
| Indonesia | 1.000 | 1 | 5 | 4 | — | 0 | 3 | 3 |
| Qatar | — | — | 0 | 0 | — | 0 | 0 | 0 |
| Singapore | n.a. | 0 | 1 | 1 | n.a. | 0 | 6 | 6 |
| Zaire | 0.000 | 1 | 2 | 2 | — | 0 | 4 | 4 |
| *Developing, democratic* | | | | | | | | |
| Botswana | — | — | — | — | 1.000 | 1 | 6 | 5 |
| Costa Rica | 0.400 | 5 | 10 | 8 | 0.750 | 4 | 11 | 8 |
| Malta | 0.000 | 1 | 1 | 1 | 0.500 | 2 | 5 | 4 |
| Jamaica | 1.000 | 1 | 3 | 2 | 0.333 | 3 | 7 | 6 |
| West Samoa | — | — | — | — | 0.333 | 3 | 3 | 2 |
| Barbados | — | — | — | — | 0.000 | 4 | 2 | 2 |
| Israel | 0.000 | 3 | 2 | 2 | 0.000 | 6 | 3 | 3 |
| Bahamas | — | — | — | — | n.a. | 0 | 3 | 3 |
| *Developing, mixed* | | | | | | | | |
| Argentina | 1.714 | 7 | 20 | 8 | 1.111 | 9 | 16 | 6 |
| Brazil | 0.500 | 8 | 9 | 5 | 1.000 | 5 | 11 | 6 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Chile | 0.750 | 4 | 6 | 3 | 1.000 | 1 | 12 | 11 |
| India | 0.000 | 2 | 6 | 6 | 0.667 | 6 | 7 | 3 |
| Korea, Rep. of | n.a. | 0 | 2 | 2 | 0.667 | 3 | 6 | 4 |
| Peru | 0.333 | 6 | 11 | 9 | 0.667 | 3 | 6 | 4 |
| Uruguay | 1.000 | 2 | 6 | 4 | 0.600 | 5 | 6 | 3 |
| Colombia | 0.167 | 6 | 5 | 4 | 0.500 | 4 | 3 | 1 |
| Poland | 0.000 | 1 | 0 | 0 | 0.500 | 6 | 6 | 3 |
| Venezuela | 0.167 | 6 | 5 | 4 | 0.500 | 4 | 7 | 5 |
| Honduras | 0.200 | 5 | 2 | 1 | 0.400 | 5 | 3 | 1 |
| Portugal | 0.000 | 1 | 4 | 4 | 0.385 | 13 | 7 | 2 |
| Turkey | 0.600 | 5 | 7 | 4 | 0.273 | 11 | 7 | 4 |
| Greece | 0.053 | 19 | 2 | 1 | 0.250 | 8 | 5 | 3 |
| Ghana | 0.000 | 4 | 2 | 2 | 0.200 | 5 | 4 | 3 |
| Nigeria | 0.000 | 2 | 2 | 2 | 0.200 | 5 | 3 | 2 |
| Nepal | 0.250 | 4 | 3 | 2 | 0.143 | 7 | 3 | 2 |
| Panama | 1.500 | 2 | 4 | 1 | 0.125 | 8 | 2 | 1 |
| Thailand | 0.250 | 4 | 4 | 3 | 0.111 | 9 | 3 | 2 |
| Malaysia | 0.000 | 3 | 0 | 0 | 0.000 | 2 | 2 | 2 |
| Pakistan | 0.167 | 6 | 4 | 3 | 0.000 | 3 | 5 | 5 |
| Philippines | 0.000 | 4 | 3 | 3 | 0.000 | 2 | 2 | 2 |

—Not available. Vulnerability is not available when data on either the number of political transitions or the number of central bank turnovers are not available.
n.a. Not applicable.

*Note:* Within each subgroup, central banks are arranged in descending order of vulnerability during the 1971–89 subperiod.

a. The nonpolitical turnover is defined as the number of changes in the governor of the central bank that occurred more than six months after a political transition.

*Sources:* Authors' calculations.

## Appendix. Tests for the Determination of the Cutoff between Political and Nonpolitical Periods

To test for possible differences between the probability of a change of central bank governor in the ten-month-or-more intervals and in each of the other intervals in table 1, we use the normal approximation to the binomial distribution. More specifically, let $c_i$ be the number of changes of central bank governor in interval $i$ following a political transition. As in table 1, $i$ may assume the interval values (0–1), (2–3), (4–6), (7–9), and (10 or more). We denote each of these intervals by its lower bound. Thus $i$ assumes the values 0, 2, 4, 7, and 10. Let $n_i$ be the number of intervals of type $i$ in the sample. This number is determined by the total number of political transitions in the sample. The frequency of central bank governor changes within interval $i$ is given by

$$f_i = c_i/n_i \text{ for } i = 0, 2, 4, 7, 10. \tag{A-1}$$

Let $P_i$ be the true conditional probability of a change of central bank governor in interval $i$ following a political transition. The null hypotheses to be tested are

$$H_0: P_i = P_{10} \text{ for } i = 0, 2, 4, 7 \tag{A-2}$$

against each of the alternative hypotheses that $P_i$ is significantly larger than $P_{10}$. The specification of the alternative hypotheses accommodates the possibility that the probability of a turnover at the central bank is larger at shorter lags following a political transition. Under each of the null hypotheses the values

$$z_i \equiv \frac{f_i - f_{10}}{\sqrt{\bar{f}_i(1-\bar{f}_i)\left[\frac{1}{n_i} + \frac{1}{n_{10}}\right]}}, \; i = 0, 2, 4, 7 \tag{A-3a}$$

where

$$\bar{f}_i = \frac{c_i + c_{10}}{n_i + n_{10}}, \; i = 0, 2, 4, 7 \tag{A-3b}$$

have approximately a standard normal distribution, provided both $n_i$ and $n_{10}$ are larger than 30.[9] (See, for example, Huntsberger, Croft, and Billingsley 1980: 302.) The null hypotheses should be accepted for small values of the $z_i$'s and rejected for large positive $z_i$'s. Table A-2 displays these statistics.

9. These conditions are always satisfied for the entire sample of economies and the two main subsamples. For $i = 0, 2$, the terms $n_i$ and $n_{10}$ are the numbers of two-month periods in the appropriate intervals. For $i = 4, 7$, they refer to the number of three-month periods. This variation in the length of the basic time unit is introduced to make its length identical to the number of months within each interval (two months for $i = 0, 2$ and three months for $i = 4, 7$). In either case, the value of $n$ is sufficiently large to make the normal approximation to the binomial valid.

Table A-2. *Values of the Test Statistic $Z_i$ for the Null $H_0$: $P_i = p_{10}$*

| | *Interval (number of months)* | | | |
|---|---|---|---|---|
| *Economy group* | *0–1* | *2–3* | *4–6* | *7–9* |
| All economies | 12.0 | 2.8 | 2.9 | –0.6 |
| Industrial | 3.2 | –0.2 | 1.1 | –1.6 |
| Developing | 13.0 | 4.0 | 3.3 | –0.7 |

*Note:* Sample size is sixty-seven economies.
*Source:* Authors' calculations.

Because a $z_i$ of 1.96 implies that the null is rejected at the 0.05 level, we conclude that the probabilities of a turnover at the central bank are significantly larger for the 0–1, 2–3, and 4–6 intervals than for the 10-or-more intervals. However, there is no significant difference in the probability of a change of central bank governor between the seven-to-nine-month and the ten-or-more-month intervals. The evidence in table A-2 therefore supports the conclusion that for the whole sample the cutoff between political and nonpolitical periods occurs at a lag of six months following a political transition. Accordingly, we define periods of up to six months following a political transition as political and periods seven or more months after the latest political transition as nonpolitical. The respective frequencies are 0.037 and 0.015, respectively. The $z$ statistic testing the significance of the difference between these two estimated probabilities is a huge 9.6.

A replication of these tests separately for each of the two economic categories of countries reveals that for developing countries the appropriate cutoff between political and nonpolitical periods is still at the upper end of the 4–6 interval. For developing economies the $z$ statistic also very significantly rejects the null hypothesis that the frequency within the entire period of zero to six months is no different from the frequency in periods that are seven or more months after a political transition.

In the industrial economies, the probability of a change at the central bank in the 0–1 interval is significantly higher than in the 10-or-more interval. The frequencies in the two-to-three-month, four-to-six-month, and seven-to-nine-month intervals, however, are not significantly different from frequencies in the 10-or-more interval. Thus, for industrial economies we should locate the cutoff between political and nonpolitical periods after the 0–1 interval.

## References

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

Alesina, Alberto, Sule Ozler, Nouriel Roubini, and Phillip Swagel. 1995. "Political Instability and Economic Growth." Draft.

Alesina, Alberto, and Lawrence H. Summers. 1993. "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence." *Journal of Money, Credit and Banking* 25(May):151–62.

Alesina, Alberto, and Roberto Perotti. 1994. "The Political Economy of Growth: A Critical Survey of the Recent Literature." *The World Bank Economic Review* 8(3):351–71.

Anayadike-Danes, M. K. 1995. "Comment on 'Measuring the Independence of Central Banks and Its Effect on Policy Outcomes' by Cukierman, Webb, and Neyapti." *The World Bank Economic Review* 9(2):335–40.

Arriagada, H. G., and Carol Graham. 1994. "Chile: The Maintenance of Adjustment and Macroeconomic Stability during Democratic Transition." In Stephan Haggard and Steven B. Webb, eds., *Voting for Reform: Democracy, Political Liberalization and Economic Adjustment.* New York: Oxford University Press.

Barro, R. J. 1991. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* 106(May):407–43.

Cukierman, Alex. 1992. *Central Bank Strategy, Credibility, and Independence: Theory and Evidence.* Cambridge, Mass.: MIT Press.

———. 1994. "Commitment through Delegation, Political Influence, and Central Bank Independence." In J. O. de Beaufort Wijnholds, S. C. W. Eijffinger, and L. H. Hoogduin, eds., *A Framework for Monetary Stability.* Boston: Kluwer Academic Publishers.

Cukierman, Alex, Sebastian Edwards, and Guido Tabellini. 1992. "Seigniorage and Political Instability." *American Economic Review* 82(June):537–55.

Cukierman Alex, Miguel Kiguel, and Nissan Liviatan. 1992. "How Much to Commit to an Exchange Rate Rule? Balancing Credibility and Flexibility." *Revista de Analisis Económico* 7(1, June):73–89.

Cukierman, Alex, Steven B. Webb, and Bilin Neyapti. 1992. "Measuring the Independence of Central Banks and Its Effect on Policy Outcomes." *The World Bank Economic Review* 6(3):353–98.

Cukierman, Alex, Pantelis Kalaitzidakis, Lawrence H. Summers, and Steven B. Webb. 1993. "Central Bank Independence, Growth, Investment, and Real Rates." *Carnegie-Rochester Conference Series on Public Policy* 29(Autumn):1–46.

Cukierman, Alex, Pedro Rodriguez, and Steven B. Webb. Forthcoming. "Central Bank Autonomy and Exchange Rate Regimes—Their Effects on Monetary Accommodation and Activism." In Harry Huizinga and Sylvester Eijffinger, eds., *Positive Political Economy: Theory and Evidence.*

de Haan, Jakob, 1995. Personal communication. University of Groningen, the Netherlands. August.

Easterly, William, C. A. Rodriguez, and Klaus Schmidt-Hebbel, eds. 1992. "Public Sector Deficits and Macroeconomic Performance." Draft.

Edwards, Sebastian, and Guido Tabellini. 1991. "Explaining Fiscal Policies and Inflation in Developing Countries." *Journal of International Money and Finance* 10 (March, Supp.):16–48.

*The Europa World Yearbook.* Various years. London: Europa Publications.

Fischer, Stanley. 1991. "Growth, Macroeconomics, and Development." In Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1991.* Cambridge, Mass.: National Bureau of Economic Research.

Grilli, Vittorio, Donato Masciandaro, and Guido Tabellini. 1991. "Political and Monetary Institutions and Public Financial Policies in the Industrial Countries." *Economic Policy: A European Forum* 6(October):341–92.

Haggard, Stephan, Robert Kaufman, and Steven B. Webb. 1992. "Democracy, Dictatorship, and Inflation in Middle-Income Countries." World Bank, Country Economics Department, Washington, D.C. Processed.

Haggard, Stephan, and Steven B. Webb. 1994. *Voting for Reform: Democracy, Political Liberalization and Economic Adjustment.* New York: Oxford University Press.

Havrilesky, Thomas. 1992. *The Pressures on American Monetary Policy.* Norwell, Mass.: Kluwer Academic Publishers.

Huntsberger, D. V., D. J. Croft, and Patrick Billingsley. 1980. *Statistical Inference for Management and Economics.* 2d ed. Boston: Allyn and Bacon.

Parkin, Michael. 1986. "Domestic Monetary Institutions and Deficits." In J. M. Buchanan, C. K. Rowley, and Robert Tollison, eds., *Deficits.* New York: Basil Blackwell.

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3: 425-450

# Pioneers for Profit: St. Petersburg Entrepreneurs in Services

Martha de Melo, Gur Ofer, and Olga Sandler

*Russian private entrepreneurs in services are truly pioneering, because many services—especially trade, financial services, and most business and consumer services—are poorly developed. This article uses 1992 and 1993 survey data from St. Petersburg to assess the characteristics of these entrepreneurs, their firms, and the markets in which they deal. Evaluation of the firms' performance establishes how well they are doing and provides insight into the determinants of success. Their performance was surprisingly good and can be attributed to several factors, including the existence of a substantial gap between the desired and the actual levels of many services and the high level of education and motivation of the entrepreneurs themselves. Policy priorities are to achieve macroeconomic stability, a transparent legal framework and simplified regulations, a well-designed tax code, further privatization of real estate, and better access to finance. Direct assistance programs would be useful in providing information, counseling, and financing to small and medium-size firms.*

In McKay (1970), "pioneers for profit" were foreign industrial entrepreneurs in Russia prior to World War I. Today, pioneers for profit are the Russian Federation's new domestic entrepreneurs. Under communism, private ownership of the means of production was not allowed, and markets as we know them did not exist. Prices were fixed and goods were distributed according to a central plan. Thus Russian entrepreneurs, who are creating and managing private businesses in today's unstable macroeconomic and regulatory environment, are entering new territory. And entrepreneurs in services are truly pioneering, because many services—especially trade, financial services, and most business and consumer services—are poorly developed. Such activities were previously suppressed for a number of reasons, including the Marxist doctrine that held them to be speculative and unproductive. As a result, Russians have had relatively little experience in services, as customers or providers.

Our interest in private entrepreneurs in services grew out of a study of the gap between the actual levels of services in the Russian Federation and other former

Martha de Melo is with the Policy Research Department at the World Bank; Gur Ofer is with the Department of Economics at the Hebrew University in Jerusalem; and Olga Sandler is an independent consultant. The authors thank Leila Webster, Alan Gelb, and three anonymous referees for advice and the Leontief Center in St. Petersburg for logistical support in the conduct of the surveys.

Soviet states and those to be expected on the basis of other countries' experiences.[1] Our interest was heightened by early evidence that the source of private sector growth in Central and Eastern European countries, in transition from a centrally planned to a market economy, was largely services. Because service firms are needed in most transactions, in some sense they *are* the market, providing information, communication, and distribution. The market infrastructure that service firms provide facilitates the development of new agricultural and manufacturing activities as well as other services.

Private business activity was officially allowed in the Soviet Union starting in 1987, although only limited private sector development took place before 1991. Since 1991 a new legal framework has been developed. Property rights, contract law, and company law have been woven into the fabric of society, although some laws and decrees are unstable or ambiguous and institutions are weak. Privatization of both small and large state enterprises has proceeded relatively rapidly in the Russian Federation, but its pace has varied considerably by region. Privatization of commercial real estate has proceeded much more slowly for several reasons, including ambiguities in ownership and the absence of private ownership rights for land.

In a typical market economy, the majority of all private firms are small and medium-size, and the majority of these are in services (on services, see, for example, Giarini 1987; on the importance of small firms, see Storey 1983, 1989 and Acs and Audretsch 1993). However, in the late 1980s only 22 percent of Russian firms had under 100 employees, compared with 94 percent in the United States, and a minority of the Russian small and medium-size firms were in services (Brown, Ickes, and Ryterman 1994). The share of services in gross domestic product (GDP) was an estimated 33 percent in Russia in 1990, compared with 50 percent or more in upper-middle-income countries and 67 percent in the United States (Goskomstat, various issues). Thus, the simultaneous growth of services and small firms is an important feature of restructuring for economies in transition to a market economy. Such growth is relatively easy because privatization of small firms is simpler than privatization of large ones and many service firms require only modest investment, simple technology, and limited market infrastructure.

Even so, it is surprising how quickly the emerging market economy in the Russian Federation is being transformed. By 1993, although the number of stores per thousand inhabitants was still only 2 to 2.5, compared with 7 to 10 in most Western European countries, the share of services in GDP had increased to 43 percent (Earle and others 1994; Goskomstat, various issues). Estimates for the first half of 1994 put services at around 50 percent of GDP. Both new firms and privatized firms are contributing to a partly unrecorded but growing private sector, characterized by smaller firms and a strong concentration in services.

1. See Easterly, de Melo, and Ofer (1994) for estimates of the gap in different types of services between the former Soviet Union and other countries. Services are defined there and in this article as all sectors of the economy other than agriculture, manufacturing, construction, and utilities.

This article draws on survey data from 1992 and 1993 to look at the characteristics and performance of entrepreneurs in the services sector in St. Petersburg. Section I describes the surveys and the firms that were included. Section II discusses the characteristics of the owners, their firms, and the markets in which they deal. Section III discusses firm performance and its relation to a variety of internal and external factors. The questions of why performance has been so good and whether it can be expected to continue are addressed here. And section IV draws some policy implications from the survey findings. Judged in the context of the Russian Federation's transition—with its turbulent environment, declining output, and lack of market and private sector tradition—firm performance looks remarkably strong.

## I. The St. Petersburg Surveys

Although the successful development of service firms is vital to the Russian Federation and other economies in transition, in early 1992 little was known about the emergence of service sector entrepreneurs and firms there or about the problems they faced. The lack of information prompted us to undertake our own surveys, which are unique in that they focus on service firms and provide observations of the same firms at two points in time—albeit close together. Our purpose was to understand the entrepreneurs, their firms, and the external business environment; to identify factors that encourage or impede growth; and to recommend policies to accelerate development.

The surveys were undertaken in St. Petersburg—a city with historically close economic and cultural links with Europe and a large number of professionals who have worked in the city's military industrial complex. Although St. Petersburg is not typical of the Russian Federation, the development of a market economy there is likely to precede such a development in many other parts of the Russian Federation and in other former Soviet states. A better understanding of this experience can contribute to some timely policy lessons.

The first survey was undertaken in January 1993 and included eighty-six entrepreneurs (see de Melo and Ofer 1994). Firms were randomly chosen from 10,000 registered firms and spanned the full range of services, including trade, business services, consumer services, transport, finance, and health and education. An estimated one out of three firms responded to an initial inquiry and expressed willingness to be interviewed; a randomly chosen subsample was drawn from this population. Prior to the interview, firms were screened to ensure that each was majority privately owned and majority Russian-owned and that services accounted for 50 percent or more of revenues. No restrictions were placed on a firm's legal form or the number of its employees. An interviewer then personally visited each firm and administered a broadly conceived questionnaire, which took an average of three hours. Most questions were qualitative in nature, although quantitative data were collected for major aggregates such as employment and gross revenues.

Table 1. *Comparison of 1992 Survey and Statistical Office Data on Service Firms, Including Cooperatives, in St. Petersburg by Service*
(percentage of total firms)

| *Service* | *Statistical Office data* | *Survey data* |
|---|---|---|
| Trade | 40 | 37 |
| Business services | 37 | 31 |
| Consumer services | 12 | 15 |
| Health and education | 9 | 8 |
| Transportation | 2 | 5 |
| Finance | 0 | 4 |
| Total | 100 | 100 |

*Note:* The Statistical Office data include some cooperatives and small enterprises with majority public ownership; services account for 52 percent of all activities. The survey data are based on eighty-six firms.

*Source:* Authors' survey data and St. Petersburg Statistical Office (1993).

Ten months later, in November 1993, a shorter, telephone follow-up survey was undertaken. All eighty-six original firms were located, although only eighty-two entrepreneurs could be interviewed. The four firms that could not be interviewed in the follow-up survey were two trading firms and one firm each in business services and house repair. The most surprising finding was that all eighty-six firms interviewed in January 1993 were still functioning at the end of the year. This success is unusual in the face of the high failure rate of small and medium-size enterprises in market economies.

For simplicity, data from the first survey, which reflect end-1992 or the period of operation prior to end-1992, are labeled 1992; data from the second survey, which reflect end-1993 or calendar year 1993, are labeled 1993. Where data from only the first survey are used, the total number of firm observations is eighty-six; where panel data are used, the total number of observations is eighty-two.

The sectoral distribution of survey firms, shown in table 1, is similar to that of all small and medium-size firms and cooperatives in St. Petersburg in 1992.[2] And the survey firms are believed to be typical of the wider population in size, financing, market orientation, and major complaints. Although biases may exist, their experience is interesting and provides insight into the Russian Federation's new private sector.

## II. Survey Findings on Owners, Firms, and Markets

St. Petersburg pioneer entrepreneurs are part of a spontaneous and wide-ranging human conversion to free enterprise. Talented people in state industries and government institutes are opting for a life of risk and opportunity in the newly

2. The data from the St. Petersburg Statistical Office were received only in June 1993; they were not available at the time of the original survey.

emerging private sector. In table 1 their firms are classified by their primary revenue activity into six categories of service activity. According to the survey data, the largest is trade (thirty-two firms or 37 percent), which includes wholesale and mixed-wholesale-and-retail companies, as well as five retail-only companies (including one kiosk owner) and three restaurants. The second-largest category is business services (twenty-seven firms or 31 percent), which includes firms providing computer services, research and development, and legal and management consulting. The third-largest category is consumer services (thirteen firms or 15 percent), which includes travel, repair, custom tailoring, hairdressing, and publishing companies. The other three groups—health and education (public-type services that cater to both businesses and households), transportation, and financial services—include only a few firms each.

### *Owners*

Who are these pioneers for profit and why did they become entrepreneurs and go into services? They are on the whole highly educated, energetic, experienced, and resourceful people. Almost 85 percent have a university education, and 20 percent have graduate-level education. As also found in Ukraine, a strong educational background has provided them with both marketable specialized skills and general analytical abilities useful in setting up and maintaining a business successfully in a complex, difficult environment. (For a comparison with private entrepreneurs in Ukraine, see Stone and Novitskaya 1993.)

Entrepreneurs' explanations of why they started a business were, in order of frequency, to make more money, to become independent, to pursue their profession more freely, and to pursue a hobby (for example, computer programming, commercial design, or the arts). Many entrepreneurs went into services related to their previous public sector activity, such as computer support or educational training. Others established service firms because of the relatively low capital investment required or because they saw the gap in services, such as tourism and restaurants, as a chance to establish a niche. Some 10 percent started out in manufacturing but switched to services, often trade, to replenish their cash flow. One entrepreneur began manufacturing security equipment but, faced with insufficient new orders, switched to repairing and servicing existing security systems.

Virtually all the entrepreneurs interviewed were once employed in the public sector, where they were managers, academics or technical specialists, or workers. Table 2 gives a profile of the owners by background. The former managers represent insiders, the privileged elite of the communist regime. They nevertheless encompass more than the "nomenklatura," a very restricted group of directors, high-level bureaucrats, and party officials. As a group, former managers were slightly older than the survey average, and most had university or technical education. Many of the former academics and technical specialists had completed postuniversity degrees, and they employed a high proportion of professionals in their firms. The former "workers" were either white-collar (clerical

Table 2. *Profile of Owners of Service Firms in St. Petersburg, 1992*

| | *Owner's background* | | | |
|---|---|---|---|---|
| *Indicator* | *Manager* | *Academic or technical specialist* | *Worker* | *Total* |
| Number of owners | 41 | 34 | 11 | 86 |
| Average age | 45 | 41 | 35 | 42 |
| Women (percent) | 17 | 6 | 18 | 13 |
| Higher education (percent) | 80 | 97 | 55 | 84 |
| Previous employment (percent) | | | | |
| State enterprise | 85 | 19 | 55 | 57 |
| Academic institute | 5 | 68 | 9 | 29 |
| Same as present | 47 | 31 | 55 | 42 |

*Note:* The data are based on eighty-one to eighty-six firms, depending on the characteristic. $\chi^2$ values for cross-tabulations: age = 0.06; gender = 0.30; education = 0.00; previous work places = 0.00; and activity = 0.26. (The $\chi^2$ statistic shows the probability that the data are distributed randomly across the cells. A low $\chi^2$ indicates a meaningful difference in owner characteristics.)

*Source:* Authors' calculations.

and administrative) or blue-collar workers; they were typically younger, and had a lower level of education than the other two groups.

The characteristics of firms run by entrepreneurs with different backgrounds are shown in table 3. Firms owned by former managers were more likely to be privatized and to engage in trade. The four former government officials in this group all ran (largely privatized) trading firms whose goods were provided primarily by state enterprises—supporting the popular view that bureaucrats were in a position to arrange good deals for themselves. Their trading firms were large, averaging more than 200 employees, and more likely to have state enterprises as customers. Firms owned by academics or technical specialists were more likely to be new firms, to engage in business services, and to receive foreign aid in some form. Firms owned by white- or blue-collar workers were more likely to be small, and they received no foreign assistance.

In many ways, entrepreneurs of small firms in services behave like entrepreneurs of small firms in other sectors. They rely heavily—and significantly more than in market economies—on personal contacts to accomplish business objectives, substituting such contacts for the information networks and standard government and business procedures that exist elsewhere. For example, in Poland, where private sector development was more advanced, survey data in late 1992 showed that more than 90 percent of entrepreneurs there also found personal contacts to be the most frequent and efficient way of finding customers, even customers abroad (Wyznikiewicz, Pinto, and Grabowski 1993). These personal contacts originate in family relationships, ethnic or regional ties, work relationships, or military service. They are used to hire workers; to sell products, including exports to former Soviet republics; and to deal with the government, banks,

Table 3. *Profile of Service Firms in St. Petersburg by Owners' Backgrounds, 1992*

| | *Owner's background* | | | |
|---|---|---|---|---|
| *Indicator* | *Manager* | *Academic or technical specialist* | *Worker* | *Total* |
| Number of firms | 41 | 34 | 11 | 86 |
| Average number of full-time workers | 59 | 62 | 20 | 55 |
| Privatized (percent) | 51 | 12 | 27 | 33 |
| Received foreign aid (percent) | 14 | 29 | 0 | 18 |
| Service (percent) | | | | |
| Trade | 51 | 21 | 36 | 37 |
| Financial and business services | 22 | 49 | 27 | 35 |
| Consumer services | 20 | 12 | 9 | 15 |
| Health and education | 2 | 15 | 9 | 8 |
| Transportation | 5 | 0 | 18 | 5 |

*Note:* The data are based on seventy-seven to eighty-six firms, depending on the characteristic. $\chi^2$ values: privatized = 0.01; activity = 0.02; and foreign aid = 0.52.

*Source:* Authors' calculations.

and state enterprises. Relying on personal contacts, however, is less efficient than using established, impersonal market institutions, and it can limit the ability to expand business activity.

In other ways, entrepreneurs in services differ from entrepreneurs in other sectors. A comparison with St. Petersburg entrepreneurs in private manufacturing (see Webster and Charap 1993) reveals that service sector entrepreneurs are more diverse in age, gender, and work experience; they operate newer, more independent firms; they face more competition in their output markets; and they are more active in seeking new markets through diversification.

### *Firms*

Survey firms were quite young: as of January 1993 their average age since start-up or their becoming majority private was less than two years, compared with eight years for Polish firms at that time (Wyznikiewicz, Pinto, and Grabowski 1993). Most survey firms were new firms and less than one-third had been privatized. Trading firms were particularly likely to have been privatized and accounted for two-thirds of all privatizations in the survey. Several consumer services had also been privatized. Most firms providing business services were new, suggesting that conditions of entry are particularly important for such firms.

The distinction between new and privatized firms, however, is not as sharp as might be expected. Of the new firms, most originated within a public organization, often with no formal arrangements but having important economic links. For example, the owner of a new firm might moonlight on the job in the host organization, rent space from it (in theory or in fact), use its equipment, or

provide it with services. In many cases such symbiotic relations gradually erode, leading to full separation, but twenty-two firms classified as new in the survey still maintained one or more of these relations with their host organization. (This is also true in other transition countries. In the former Czechoslovakia, for example, new entrepreneurs' personal contacts with former state enterprise managers were important in maintaining access to materials, second-hand equipment, and market information, according to McDermott and Mejstrik 1992.) For privatized firms in the St. Petersburg sample, the minority public ownership was relatively high in a few firms, but it accounted for only 3 percent of enterprise equity on average. Some 80 percent of equity was held by individuals; the remaining shares were held by other firms.

Most survey firms became private—either as new or privatized companies—in 1990 and 1991. As of January 1993 their average age was twenty-one months. Only two were founded in 1987, the year Michail Gorbachev officially allowed private businesses to open. About half the firms were registered as joint-stock closed or its equivalent—limited liability. The rest were registered as private individual, small enterprise, leasehold, or cooperative. Introduced in 1990, the joint-stock designation is popular because owners are not personally responsible for their debts. An additional seven firms had converted to this form by end-1993.

*Firm size.* Firms in the survey varied widely in size, from a health care company with one person and monthly revenues of $100 (R10,000) to two firms with 1,000 employees or more and a trading company with monthly revenues of $20 million (R2 billion).[3] Survey data indicate that 87 percent of private firms have fewer than 100 employees, compared with 22 percent under the old regime. However, this figure understates the extent of change in the industrial structure, because only registered firms are included in the survey. Registration entails responsibilities, such as bank accounts and tax payments, that discourage some businesses. Thus, many new businesses are undeclared and part of Russia's "shadow economy," recently officially estimated at more than 20 percent of GDP. Unregistered businesses include self-employed crafts workers and service providers who operate with or without a license.

Monthly revenues for different activities are shown in table 4, although Russia's high inflation makes such estimates difficult. The median monthly revenue for survey firms during the last quarter of 1992 was $20,000 (R2 million), suggesting an annual turnover of about a quarter of a million dollars. As expected, gross revenues of trading firms were high, reflecting the high value of goods purchased for a given value added. But revenues of the few survey firms in finance and transport were higher, reflecting the former's large profits and the latter's need to recover capital costs. In general, premises reflected revenues,

3. Rubles are converted to U.S. dollars using an estimated purchasing power parity exchange rate of R100 = $1. This compares to the nominal exchange rate of R400 = $1 in the fourth quarter of 1992 on the Moscow International Currency Exchange Market.

Table 4. *Size of Service Firms in St. Petersburg by Revenue and Employees, 1992*

| | *Monthly revenue (R1,000)* | | *Number of full-time employees* | |
|---|---|---|---|---|
| *Main activity* | *Average* | *Median* | *Average* | *Median* |
| Trade | 98,711 | 4,000 | 95 | 25 |
| Business services | 9,200 | 1,500 | 22 | 11 |
| Consumer services | 6,522 | 600 | 40 | 17 |
| Health and education | 985 | 500 | 10 | 7 |
| Transportation | 9,925 | 9,750 | 5 | 6 |
| Finance | 38,750 | 38,750 | 146 | 41 |
| Total | 40,263 | 2,000 | 55 | 13 |

*Note:* Data on monthly revenue, or gross sales, are based on sixty-four firms; data on full-time employees are based on eighty-five firms. The estimated purchasing power exchange rate for rubles is R100 = $1.

*Source:* Authors' calculations.

with high-revenue firms occupying large, well-equipped, and well-guarded offices.

The median number of full-time employees in 1992, also shown in table 4, was thirteen, and the average, fifty-five. The median is quite typical of small and medium-size enterprises in Western Europe; the average is strongly affected by a few large firms (a transportation company and two trade companies) that had more than 800 employees. Trading firms tended to be larger than average; business services, smaller than average; and consumer services, about average.

*Diversification.* Diversification has both static and dynamic dimensions. In 1992 many firms—especially those in trade or business services—were engaged in several activities, basically hedging their bets. These multiple activities were generally small-scale and sometimes were carried out as subsidiaries, under separate legal structures. This early emphasis on diversification is consistent with the contemporary experience of Ukrainian private enterprises that were found to "grow by diversification, not expansion" (Stone and Novitskaya 1993). During 1993 St. Petersburg entrepreneurs in services focused more on lowering their costs or upgrading the quality of their product than on undertaking additional activities. This strategy is consistent with the Johnson and Loveman (1995) view that firms are more likely to survive when the attention of management is concentrated on a set of core activities.

In addition to engaging in multiple activities at a given point in time, firms also diversified over time. This was particularly common before 1992, when many firms were using trial and error to explore a range of activities. During 1993 the rate of diversification moderated, with most firms maintaining the same primary activity but some still looking for the right niche. Eighteen firms—most of them quite small—changed their main activity between 1992 and 1993.

Patterns of diversification over time can be interesting to policymakers; the popularity of a given activity may indicate which firms should be followed up by tax authorities, which activities should receive priority attention for simplifica-

tion and clarification of regulations, and which entrepreneurs might require assistance during the transition. For example, the fact that no firms diversified into consumer services but several firms diversified out suggests that consumer services should not be high on the tax collector's list. Also, the fact that research and development firms continued to diversify out of business services and into trading or manufacturing and construction suggests that they might require assistance during the transition.

The majority of entrepreneurs engaged at some point in trading, suggesting that substantial gains in efficiency could be achieved by simplifying and clarifying trade regulations. In 1992 about 33 percent of survey firms moved into trading, or new areas of trading, and only 5 percent curtailed overall trade operations. Within trading, entrepreneurs moved away from wholesale trading in foodstuffs, for which margins had narrowed and supplies—especially from Ukraine—had been disrupted. In 1993, 25 percent of firms initiated or expanded trading activities, but 10 percent curtailed or stopped such activities. Trading was still profitable but was no longer the only lucrative pursuit. The diminishing attraction of trading is consistent with developments elsewhere. In Poland, for example, the speculative gains in trading in 1990 and 1991 were disappearing by late 1992.

Financial services continued to be profitable over the two-year period, and interest in real estate and related construction and renovation activities was growing strongly in 1993, again suggesting areas of focus for authorities. About 10 percent of survey firms started or expanded some real estate activity: real estate services, building repair, building construction, or architectural design. Only one firm cut back on its real estate activities. About 10 percent of firms also diversified out of services into manufacturing or construction, but this trend is difficult to interpret without indications of diversification rates in the opposite direction.

*Financing.* All over the world, small and medium-size enterprises obtain their financing primarily from the personal resources of their founders and reinvested profits, and only secondarily from bank credit. This is true for St. Petersburg entrepreneurs, who as a group rely very little on bank credit. On average, firms were established mainly with the owner's personal financial resources, supplemented in part by early profits and loans from family and friends. Working capital and investments have been financed largely from internal profits. Bank credits contributed only 7 to 11 percent of financing—compared with an estimated one-third to a half in the United States (Storey 1982). Government programs contributed only 1 to 2 percent.

The median nominal interest rate paid by survey entrepreneurs in 1992 was 120 percent, with inflation running at an annual rate of over 2,000 percent. In reality, real interest rates were much higher than nominal rates would imply because of discounting up front, extra fees, and bribes, but they were still probably strongly negative in real terms. Not one entrepreneur had a loan with a

maturity beyond twelve months, although several said their short-term credit lines could be rolled over. A number of entrepreneurs said that available loans were too short-term to interest them.

In 1992 only 22 percent of St. Petersburg service firms had bank credits, and this share dropped to 18 percent in 1993. These figures can be compared with data from other surveys in late 1992 and early 1993 showing 40 percent of private firms receiving bank credit in Poland (Wyznikiewicz, Pinto, and Grabowski 1993) and 50 percent in Ukraine (Stone and Novitskaya 1993). In both cases, bank credit covered only a small share of total financing needs, but credit access was clearly higher than for the St. Petersburg entrepreneurs in services. Access to bank credit may have been lower in the Russian Federation than elsewhere for the new private sector at this time, but it appears also that banks prefer to lend to manufacturing firms—because collateral is available, because the owners have better personal connections with banks, and because manufacturing firms are still viewed as socially more useful.

The reason for the drop in the share of firms with bank credit from 1992 to 1993 was not stricter lending criteria. Rather, it seems to have been the entrepreneurs' reluctance to borrow, given higher—albeit still negative—interest rates and uncertainty about future input and output prices. According to the 1993 survey data, nominal interest rates had risen to 250 percent or more on an annual basis, with inflation running at about 1,000 percent, and entrepreneurs said the maximum annual interest rate they would pay for a ruble loan was from 100 to 300 percent. In other surveys undertaken around this time, Stone and Novitskaya (1993) and Johnson and Loveman (1995) found that private entrepreneurs in Ukraine and Poland were also reluctant to borrow at high nominal interest rates, even when real interest rates were low or negative. In Ukraine, the main explanation was uncertainty, as here; in Poland, it was cash flow problems arising from early repayment of principle under high nominal rates.

*Main problems.* Entrepreneurs were asked about the main problems faced by their firms. Table 5 shows that in 1992 the most frequently mentioned problem—especially by firms in trade and financial services—was the instability and uncertainty in the macroeconomic and legal environment, including crime and corruption. (This was also true for private entrepreneurs in Ukraine, according to Stone and Novitskaya 1993. In Poland, where inflation had moderated by late 1992, entrepreneurs' biggest complaint was lack of financing.) Almost half the trading firms in the 1992 St. Petersburg survey were subject to pressure from the mafia in deciding from whom to buy, where to sell, and how to price their products. Problems of financing and taxes ranked next in importance.

In 1993 the same three problems were dominant. Most firms said the power of the mafia was as strong or stronger than before, and taxes were an increasing problem for firms remaining in trade and consumer services. Trading firms were subject to higher tax rates than were firms engaged in other activities, and consumer services were still suffering from low demand. Financing was the main

Table 5. *Main Problems of Service Firms in St. Petersburg, 1992 and 1993*
(percentage of firms)

| *Main activity* | *Instability* | *Finance* | *Taxes* | *Demand* | *Workspace and utilities* | *Other* |
|---|---|---|---|---|---|---|
| *1992* | | | | | | |
| Trade | 38 | 28 | 14 | 3 | 0 | 17 |
| Business services | 24 | 28 | 28 | 12 | 4 | 4 |
| Consumer services | 25 | 8 | 25 | 8 | 8 | 24 |
| Health and education | 17 | 50 | 0 | 17 | 0 | 17 |
| Transportation | 0 | 50 | 25 | 25 | 0 | 0 |
| Finance | 67 | 0 | 0 | 0 | 33 | 0 |
| Total | 29 | 27 | 19 | 9 | 4 | 13 |
| *1993* | | | | | | |
| Trade | 19 | 19 | 33 | 0 | 11 | 18 |
| Business services | 29 | 19 | 19 | 5 | 5 | 25 |
| Consumer services | 20 | 20 | 40 | 0 | 20 | 0 |
| Health and education | 33 | 33 | 0 | 0 | 17 | 17 |
| Transportation | 0 | 75 | 0 | 25 | 0 | 0 |
| Finance | 25 | 0 | 25 | 0 | 0 | 50 |
| Manufacturing and construction | 29 | 43 | 0 | 0 | 14 | 14 |
| Total | 23 | 24 | 23 | 3 | 10 | 19 |

*Note:* Data are based on seventy-nine firms in each year. Percentages may not add to 100 because of rounding.
*Source:* Authors' calculations.

problem for half the firms changing activity between 1992 and 1993, and was a prime reason for the change. For most firms, however, financing was needed for expansion, not survival.

Other problems mentioned were low demand (caused in part by a shortage of cash in state organizations, the closing of state enterprises, increased competition, the firm's own low-quality products, and market domination by more established firms); lack of affordable workspace with adequate utilities; a scarcity of good workers; delays in banking transactions, especially for transfers outside the city limits; and unsatisfactory business relations—especially difficulties in finding honest partners. These last two problems particularly affected entrepreneurs in trade and business services.

### *Markets*

In fully functioning market economies, there are numerous examples of market failure. Until recently, however, markets did not exist at all in the Russian Federation. For new private firms there, the problem is not market failure but market creation and development.

As shown in table 6, the firms' market orientation varied according to the background of owners. Former managers were more likely to deal with the state sector and households. Former academics and technical specialists were more likely to deal with other private firms and to import current inputs, and they were the only own-

Table 6. *Market Orientation of Service Firms in St. Petersburg by Type of Owner, 1992*
(average percentage share)

| | *Owner's background* | | | |
|---|---|---|---|---|
| *Market orientation* | *Manager* | *Academic or technical specialist* | *Worker* | *Total* |
| *Customers* | | | | |
| Public administration | 1 | 7 | 8 | 4 |
| State enterprises | 27 | 27 | 8 | 24 |
| Private companies | 20 | 43 | 31 | 31 |
| Households and individuals | 51 | 19 | 52 | 39 |
| Foreigners | 0 | 5 | 0 | 2 |
| *Suppliers* | | | | |
| State enterprises | 71 | 49 | 36 | 58 |
| Private companies | 13 | 28 | 49 | 23 |
| Individuals | 9 | 14 | 14 | 12 |
| Foreigners | 6 | 9 | 1 | 7 |

*Note:* Data for customers are based on seventy-eight firms, and data for suppliers are based on fifty-six firms.
*Source:* Authors' calculations.

ers to export their services. Former workers were the least likely to deal with state enterprises, presumably because they had few useful contacts; they relied on private companies as suppliers and households as customers.

For all the survey firms, state enterprises were clearly the most important input suppliers, although they were less important than for private manufacturing firms in Russia and Czechoslovakia in the early 1990s (see Webster and Charap 1993 for Russia and McDermott and Mejstrik 1992 for Czechoslovakia). Customers were more diverse than suppliers; and sales concentration was moderate, with less than 15 percent of firms receiving more than half their revenues from a single customer, typically a state enterprise. Most entrepreneurs sold their services mainly locally and in surrounding towns and looked to their competitors for pricing guidelines (for further findings on wages and prices in private service firms, see de Melo and Ofer 1994). Business services were the most dependent on a single large customer, with 70 percent receiving a quarter or more of their revenues from one source. In 1993, survey firms—especially those whose main customer in 1992 was the state sector—made a clear shift away from state sector customers and toward private firms and households. One explanation was the failure of public enterprises to pay their bills; another was the drop in state sector output.

The low reported average export share in sales reflects the reluctance of several trading firms engaged in the export of raw materials (oil, timber, and metals) to talk about these operations, which may not have been legal. In response to more general questions, 25 percent of the firms said they were engaged in some form of export, and 5 percent identified exports as their primary sales

market. Of exporters who spoke freely, half were traders, acting as intermediaries; the others provided exports of tourism, transportation, design, and information services. Most firms exported to the former Soviet states; only a quarter exported elsewhere. Constraints to exporting were said to include payment delays, government bureaucracy, and marketing, but exporting was also discouraged by foreign exchange surrender requirements, unfavorable foreign exchange conversion rates by the central bank, capital gains tax on increases in the domestic currency value of foreign currency deposits, high bank fees on foreign exchange transactions, and visibility with mafia groups. Despite these impediments, 80 percent of exporting firms increased exports in 1992; in 1993, 40 percent increased exports.

Competition was moderate and increasing. In 1992 about 50 percent of the entrepreneurs said there were ten or more companies—typically small, private firms—competing in their output market. Less than 10 percent said they had no competition. Most entrepreneurs felt the pressure of competition and reacted by reducing prices in real terms, improving quality, and diversifying. Other strategies included cutting costs, increasing market share by expansion, and agreeing on market shares with competitors. In 1993, some 70 percent of firms dealing in all markets—real estate, capital equipment, material inputs, service inputs, and service sales—found them more competitive than a year earlier. Others found conditions the same. Only a few found markets less competitive.

## III. Firm Performance

The firms' performance was evaluated to establish how well the new pioneers for profit are doing and to gain some insight into the determinants of success.

### *Performance Indicators*

The interviewer asked the entrepreneur for a qualitative assessment of five performance indicators: current profits, growth of profits, growth of real sales, expansion plans, and capacity utilization.[4] A sixth performance indicator was developed from quantitative data on employment, on the assumption that growing employment signals success.

In the original survey, an overall grade of 1 to 4 was assigned to each firm by the interviewer. It was based partly on the performance indicators and partly on a qualitative impression of the firm, taking into account the qualities of the entrepreneur and a personal inspection of the facilities. Three-quarters of firms were classified as moderately successful (grade 3) or highly successful (grade 4). A comparable interviewer grade was not possible in 1993 because the follow-up survey was conducted by telephone. Instead, the owner was asked to compare the overall well-being of the firm in 1993 with that for the period ending in 1992.

4. The original survey also asked entrepreneurs about their share of new equipment and found that one-third had 100 percent new equipment and over half had more than 50 percent new equipment.

By this evaluation, only 15 percent noted a deterioration in their firm's general well-being during 1993 (grade 1), and more than 50 percent cited an improvement (grades 3 and 4). A composite grade was calculated each year from the available overall grade and five of the performance indicators (see table 7). Capacity utilization, because it did not show positive and significant intercorrelation with the other indicators, was excluded from the composite grade.

Table 7. *Performance of Service Firms in St. Petersburg, 1992 and 1993*
(percentage of firms)

| *Indicator* | *1992* | *1993* |
|---|---|---|
| *Current profits* | | |
| Negative | 4 | 5 |
| Break even | 12 | 10 |
| Small profits | 49 | 60 |
| Large profits | 35 | 25 |
| *Growth of profits* | | |
| Less profitable | 26 | 22 |
| About the same | 21 | 29 |
| More profitable | 53 | 49 |
| *Growth of real sales* | | |
| Decreased | 19 | 12 |
| No significant change | 18 | 13 |
| Increased | 63 | 75 |
| *Expansion plans* | | |
| Will contract | 0 | 0 |
| No change | 14 | 13 |
| Will expand | 86 | 87 |
| *Capacity utilization* | | |
| Capacity underutilized | 19 | 20 |
| Near full capacity | 64 | 72 |
| Capacity constrained | 16 | 8 |
| *Growth in employment* | | |
| Decreased | 8 | 35 |
| Stayed the same | 27 | 4 |
| Increased up to and including the median | 15 | 10 |
| Increased above the median | 50 | 51 |
| *Composite grade* | | |
| 1 = lowest | 0 | 0 |
| 2 | 15 | 10 |
| 3 | 44 | 54 |
| 4 = highest | 41 | 36 |

*Note:* Data are based on seventy-one to eighty-one firms, except for those on employment growth in 1992, which are based on fifty-two firms. Performance indicators are qualitative except for employment growth, which is based on number of employees.

*Source:* Authors' calculations.

The most significant correlations among the five intercorrelated performance indicators are between expansion plans on the one hand and current profits and growth of sales on the other, and between growth of sales and growth of profits. In 1992 the five intercorrelated performance indicators are also strongly correlated with the overall grade; in 1993, equivalent correlations are positive but somewhat weaker. All performance indicators are significantly correlated with the composite grade, and most of the correlations are quite strong, suggesting that the composite grade is a good indicator of performance. Although the calculated composite grade remained at the same level in 1993 as in 1992, the inclusion of dynamic performance indicators such as growth in profits, sales, and employment means that the same rating is a sign of continued improvement. Most of the firms that responded that there was no change in 1993 were already successful in 1992.

Together, the performance indicators and overall grades give the impression of good and improving performance. This impression is reinforced by the oral testimonies of many owners, who spoke of improvement and expansion in a variety of directions: internal organizational changes, stabilization of working patterns, greater specialization, improvements in quality, diversification into promising new areas, investment in equipment and real estate, improvements in the skill levels and work ethic of employees, development of good client relations, and expansion of markets and business connections—including those outside the Russian Federation. Some even cited improvement in relations with various government officials, including tax authorities.

An interesting question is whether particular indicators or overall grades can predict subsequent performance. If so, surveys could be used to anticipate problems. With this in mind, we examined cross-period correlations. Most cross-period correlations were found to be either significantly lower or weaker than correlations within periods, but a few of the indicators were significantly correlated across periods. Expansion plans in 1992 appeared to be a particularly good predictor of 1993 performance, especially as measured by growth of sales. In fact, 80 percent of firms with concrete expansion plans in 1992 experienced an increase in real sales in 1993, and 90 percent of firms experiencing an increase in real sales had concrete expansion plans in the previous period. Several factors might explain the lack of strong intertemporal correlation among other variables. The unstable environment might result in radical changes in the firms' financial positions in the short run, and short-term cycles in firm performance might lead to inevitable corrections of short-term deviations from longer-term trends.

Employment growth—the only performance indicator that was quantified—exhibited the lowest and least significant levels of correlation with other indicators. This is not surprising. Although expanding employment can be a clear sign of success, some firms may have raised profits and productivity by reducing employment. Also, employment growth may cut into current profits, even if it leads to higher profits over time.

Table 8. *Employment Growth in Service Firms in St. Petersburg by Various Indicators, 1992 and 1993*
(annualized percent change)

| *Indicator* | *Full-time workers 1992* | *All workers[a] 1993* |
|---|---|---|
| *All firms* | 28 | 65 |
| *Main activity in 1992* | | |
| Trade | 140 | 36 |
| Business services | 156 | 26 |
| Consumer services | 0 | 37 |
| Health and education | 116 | 104 |
| Transportation | -3 | 107 |
| Finance | 507 | 35 |
| Firms that changed activity in 1993 | 18 | -2 |
| Firms that did not change activity in 1993 | 28 | 71 |
| *Main activity in 1993* | | |
| Trade | 139 | 14 |
| Business services | 209 | 96 |
| Consumer services | 0 | 0 |
| Health and education | 124 | 93 |
| Transportation | -3 | 107 |
| Finance | 232 | 117 |
| Manufacturing and construction | 11 | -41 |
| *Number of workers in the firm in 1992* | | |
| Less than 8 | 42 | 162 |
| 8 to 39 | 34 | -20 |
| 40 to 99 | 79 | 3 |
| 100 or more | 22 | 91 |
| *Firm history* | | |
| New | 103 | 40 |
| Completely privatized | 110 | 60 |
| Privatized with residual state ownership | 3 | 81 |

*Note:* Data are based on forty-nine firms.
a. Includes the full-time equivalent of part-time workers.
*Source:* Authors' calculations.

Table 8 shows that employment of full-time workers in survey firms increased at an annual rate of 28 percent in 1992, and employment of all workers increased at an annual rate of 65 percent in 1993.[5] This rapid growth indicates that private services are making an important social contribution by providing employment for state sector employees who lose their jobs. Analysis of the components of employment growth in 1993 shows that reliance on part-time employees declined. Firms hired a higher share of full-time workers, reflecting perhaps

5. Discussion is based on forty-nine firms that provided answers for both 1992 and 1993. The definition of employment changes because growth for all employees is not available for the earlier period, yet this measure is preferred. The full-time equivalent of part-time workers is calculated as half-time where information on hours per week is not available.

a growing appreciation of their own investment in the human capital of their employees but also the convenience of hiring on long-term contract rather than permanent hire. Contract hiring, which provides substantial flexibility to the employer but reduces workers' job security, increased significantly between 1992 and 1993.

Based on the categorization of firms by main activity in 1992, activity-specific growth rates varied widely in 1992 and in 1993. In 1992, employment growth was concentrated in business services, trade, and finance—that is, those activities that were the most repressed under communism. A period of consolidation with lower growth followed in 1993. Firms providing public-type services (health and education) increased employment in both periods. Based on the categorization of firms by main activity in 1993, employment in trade slowed noticeably, but in business services and finance it remained high, reflecting the diversification trends discussed earlier. Firms engaged in consumer services in both 1992 and 1993 experienced zero growth, reflecting low demand; however, firms that diversified out of consumer services increased their employment.

Employment growth in 1992 was concentrated in medium-size firms (with forty to ninety-nine workers). In 1993 these firms stabilized, and growth was generated by either very small, new firms or large firms with 100 or more employees. From the perspective of ownership, employment growth was more rapid over the two periods in new firms; this is understandable because such firms often started small. Employment growth in completely privatized firms was much higher in 1992 than in privatized firms with residual state ownership, probably because these latter firms were ten times larger and hence more likely to have surplus labor.

### *Why Such Good Performance?*

There are several possible explanations for the good overall performance of the firms surveyed. Two possible explanations are particularly difficult to gauge but could be important. First, there may be biases in the sample selection. Successful firms may have been more likely to respond to the interview request letter. Also, a positive performance bias may prevail among registered firms, from which the sample was drawn, with much of the failure and instability occurring among the undocumented population of unregistered firms. Both sources of selection bias are possible but difficult to verify. Second, despite the general negative effects of inflation, some firms, especially in trade and financial services, may have profited from it.

Other possible explanations that may be important are listed below:

- There is a substantial gap between desired and actual levels of many services—such as trade, financial services, and business services—and new private service firms have an opportunity to fill this gap. Despite declines in real incomes, this gap exists because such services were repressed under the previous regime and also because the centralized trading system collapsed.

Service firms in St. Petersburg are performing somewhat better than comparable manufacturing firms, giving support to this explanation (see de Melo and Ofer 1994).

- The private sector is able to attract high-quality, quick-learning entrepeneurs, given the disarray in public administration, budget-supported government agencies, and academic institutions. These entrepreneurs have formed a critical mass that permits some firms to operate successfully by dealing primarily with other private businesses. Their capabilities do not imply that the initial quality of services was necessarily high but rather that their learning curve is steep. In fact, several entrepreneurs claimed improvements in the quality of their product and the work habits of employees in the short period between surveys.
- New, smaller, private firms have an advantage over large, established firms, which are less flexible, may have managers that are less motivated, and may be burdened with social expenditures, including expenditures on housing. A recent survey of large, privatized firms in Moscow and Vladimir oblasts supports this explanation (Webster and others 1994).
- The environment is one in which competition, although increasing, has not yet reduced margins to the minimum, and so companies are able to change activities and to absorb bad debts, loss of customers, and other hazards without incurring bankruptcy. Wyznikiewicz, Pinto, and Grabowski (1993) provides survey evidence on the decline in trading margins in Poland, which in 1992 and 1993 was further along in the transition process than the Russian Federation was.

An interesting comparison can be made with registered private firms in Ukraine, where 20 percent of firms surveyed in April 1992 were found to have stopped operation (no reason given) eight months later. The Ukraine survey covered manufacturing firms and included a higher percentage of small firms (defined as having less than twenty-five employees) and a much higher percentage of firms with bank credit. It also had less scope for selection bias. Otherwise, the characteristics of firms, the entrepreneurs, and their environment appear to be similar (see Stone and Novitskaya 1993).

### *Internal and External Factors Affecting Performance*

Internal and external factors associated with firm performance include activity, owner's background, firm history, firm size, output market, bank financing, and main problems imposed by the business environment. These are discussed below. Other factors—such as age of firms and entrepreneurs, gender of owner, owner's previous workplace, size of entrepreneur's ownership share, and share of professionals in the labor force—were uncorrelated with performance indicators in both periods and are therefore not discussed. It is also of interest to note that service entrepreneurs in St. Petersburg have had very little contact, not to mention assistance, from government agencies or Western firms.

An analysis of the relation between performance and main activity shows that by far the best performance in both periods was demonstrated by financial firms, followed by trade and business service firms. The weakest showing was by consumer service firms, which suffered from declining demand for such services as hairdressing, tailoring, and watch repair.

An analysis of performance by type of entrepreneur suggests that general characteristics of entrepreneurship and capabilities, rather than specific managerial experience or skills, are important. Thus, two characteristics showing significant correlation with firm performance during 1992 were the number of other firms owned by the entrepreneur and the entrepreneur's level of education. Interestingly, the entrepreneur's level of satisfaction with the firm's employees was negatively correlated with performance. One interpretation is that an entrepreneur's ability to recognize labor problems was a sign of professionalism. In several successful firms, including a private hospital and an investment firm, the owners were concerned with worker performance and instituted very strict hiring policies.

In addition, there was an association between firm performance levels and the entrepreneur's background, as shown in table 9. Firms run by former academics or technical specialists performed better than the other two groups by most criteria during both periods. This was true even though firms run by former state managers were more concentrated in trade, a more profitable activity, and even though many state managers had previous experience in

Table 9. *Performance of Service Firms in St. Petersburg by Owner's Background, 1992 and 1993*
(average composite grade)

| | *Owner's background* | | | |
|---|---|---|---|---|
| *Indicator* | *Manager* | *Academic or technical specialist* | *Worker* | *Total* |
| *1992* | | | | |
| Current profits | 3.1 | 3.3 | 2.9 | 3.2 |
| Growth of profits | 2.8 | 3.1 | 2.8 | 2.9 |
| Growth of sales | 3.0 | 3.4 | 2.8 | 3.2 |
| Growth in employment | 2.7 | 3.5 | 3.3 | 3.1 |
| Expansion plans | 3.8 | 3.8 | 3.7 | 3.8 |
| *1993* | | | | |
| Current profits | 3.1 | 3.0 | 3.0 | 3.1 |
| Growth of profits | 2.7 | 3.1 | 3.2 | 2.9 |
| Growth of sales | 3.4 | 3.7 | 2.9 | 3.5 |
| Growth in employment | 2.3 | 3.2 | 3.3 | 2.8 |
| Expansion plans | 3.8 | 3.9 | 3.7 | 2.8 |

*Note:* The performance indicator values are scaled from a low of 1 to a high of 4. Data are based on seventy-three to eighty-one firms in 1992 and seventy-one to eighty-one firms in 1993, except those for employment growth, which are based on fifty-two firms.

*Source:* Authors' calculations.

service activities. This finding suggests that in services the insider status may not have the same importance it has for private firms in manufacturing (Webster and Charap 1993). Also, continuity may reflect the conservative nature of traditional managers, but services, which are strongly customer-oriented, require a radical departure from previous practices. Another possible explanation is that the analytical skills acquired by former specialists give them a comparative advantage in confronting the uncharted waters of the private service sector.

There was also some variation in performance by firm ownership. In 1992 privatized firms ranked lower than new firms. These differences narrowed during 1993 as some privatized firms cut their large, inherited work forces. Also, within the category of privatized firms, those with residual state ownership improved their rankings in 1993, which suggests that they may have benefited from special privileges given to state enterprises.

According to almost all indicators, large firms, as defined by revenue, performed consistently better than small firms did, and this was true for both survey periods (table 10). The better performance of large firms may be because higher revenues are associated with financial and other resources that help firms to deal with the bureaucracy, including tax authorities, and to obtain credit. Large firms, as defined by number of employees, also performed better than small firms, although the relationship is somewhat weaker.

Table 10. *Performance of Service Firms in St. Petersburg by Revenue, Customers, and Credit History, 1992 and 1993*
(composite grade)

| *Indicator* | *1992* | *1993* |
|---|---|---|
| *Monthly revenue* (1,000 rubles) | | |
| Less than 100 | 2.5 | 2.5 |
| 100 to 999 | 2.9 | 3.1 |
| 1,000 to 9,999 | 3.3 | 3.2 |
| 10,000 or more | 3.5 | 3.3 |
| Total | 3.2 | 3.2 |
| *Main customer* | | |
| State sector | 3.1 | 3.3 |
| Private firms | 3.5 | 3.3 |
| Households | 2.8 | 3.2 |
| Total | 3.1 | 3.2 |
| *Credit history* | | |
| No credit | 3.1 | 3.2 |
| Received credit | 3.5 | 3.3 |

*Note:* Data for monthly revenue are based on eighty-one firms in both years; data for main customer on fifty-six firms; and data for credit history on eighty firms. Monthly revenue categories are adjusted for inflation in 1993; see table 4 for ruble and dollar values. The composite grade runs from 1 (the lowest) to 4 (the highest); grades shown are averages for the group.

*Source:* Authors' calculations.

Data from the original survey indicated that firms selling to other private firms performed better than those selling to the state (table 10). Firms selling to households performed less well than the other two groups and were typically privatized, had a low share of professional workers, had low employment growth, and were run by former state enterprise managers who were continuing in the same activity, had less education, and owned no other companies. The conventional wisdom is that private companies working with the state are in a position to take advantage of rents created by distortions and to make lucrative profits. But the evidence on this is mixed.

In 1992, firms selling mainly to the state sector had lower overall performance than those selling mainly to other sectors. But in 1993, firms reducing their share of sales to the state showed weaker performance, and firms maintaining or increasing their sales to the state performed equally to or better than firms selling mainly to other sectors. The association between exports and overall performance is clearer. The seven firms that increased their export sales share over time performed better than average in both 1992 and 1993, and the four firms that reduced export shares were consistently weak overall performers (not shown).

In 1992, firms that received bank credit were more successful, but not in 1993 (table 10). One reason may be that performance moderated in trading firms, which are often able to obtain credit by using traded goods as collateral. Another may be that real interest rates were significantly higher in 1993 than in 1992, with a resulting reduction in implicit subsidies to borrowers. To the extent that a positive correlation between performance and bank credit does exist, we would like to think that the causality runs from the quality of the firm and the soundness of its expansion plans to its ability to receive credit. However, reverse causality, where the independent ability to receive cheap bank credits helps to improve performance, cannot be ruled out—particularly if real interest rates were negative, as they appear to have been. In some cases, obtaining credit may simply reflect the cleverness of the owner who perceives a substantial interest rate subsidy. Alternatively, some entrepreneurs may have received bank credit through corruption or personal connections. Receiving credit was correlated in the survey with the use of bribes and with owners who previously worked in the state sector.

Table 11 shows how the main problems varied with the firms' composite grades. In both 1992 and 1993, macroeconomic and legal instability was the most frequently mentioned problem for strong firms. Financing was an important constraint for both moderate and strong firms, suggesting that improved access to credit could foster more rapid private sector growth. Taxes were mentioned with higher frequency in 1993 than in 1992 by all categories of firms, and the share of weak firms that mentioned taxes as their main constraint increased from one-third to a half. The importance of taxes to weak firms may be explained by their severe shortage of funds, a shortage that could be eased by lower taxes but not by credit, because the firms were ineligible for credit. Because strong firms were

Table 11. *Performance of Service Firms in St. Petersburg by Main Problems, 1992 and 1993*
(percentage of firms)

| *Performance level* | *Instability* | *Finance* | *Taxes* | *Demand* | *Workspace and utilities* | *Other* |
|---|---|---|---|---|---|---|
| *1992* | | | | | | |
| Weak | 8.3 | 25.0 | 33.3 | 25.0 | 0.0 | 8.3 |
| Moderate | 33.3 | 19.4 | 19.4 | 8.3 | 5.6 | 14.0 |
| Strong | 32.3 | 35.5 | 12.9 | 3.2 | 3.2 | 12.8 |
| Total | 29.1 | 26.6 | 19.0 | 8.9 | 3.8 | 16.5 |
| *1993* | | | | | | |
| Weak | 12.5 | 12.5 | 50.0 | 0.0 | 0.0 | 25.0 |
| Moderate | 19.1 | 26.2 | 21.4 | 4.8 | 9.5 | 19.1 |
| Strong | 31.0 | 24.1 | 17.2 | 0.0 | 13.8 | 13.9 |
| Total | 22.8 | 24.1 | 22.8 | 2.5 | 10.1 | 17.7 |

*Note:* Data are based on seventy-nine firms for both years. $\chi^2$ values: 1993 = 0.39; 1992 = 0.54. Weak firms have composite grades less than 2.5; moderate firms, between 2.5 and 3.5; and strong firms, over 3.5.

*Source:* Authors' calculations.

relatively liquid—and possibly because they were able to avoid or evade taxes—they complained less about high taxes than weak firms did.

## IV. Conclusions and Policy Implications

The newly emerging private sector in St. Petersburg is developing despite an uncertain and unstable political and economic environment. It is coping with both the short-term shock effect of disrupted trade, declining output, and social disorder and the longer-term transition effect of changing ownership, laws, and institutions associated with moving from a centrally planned to a market economy. In 1992, entrepreneurs were exploring their options and looking for a niche. In 1993 the pace of diversification moderated, and output markets became increasingly dominated by the private sector. Firms received little financial support in either year and complained about taxes and lack of financing.

Firm performance was on average quite good from the time firms were established until the end of 1992, and this performance improved in 1993. Explanations for the good performance cannot be found in any dramatic improvement in the Russian Federation's economic, legal, administrative, or political environment during 1993. Inflation continued at an annual rate of close to 1,000 percent, production continued to deteriorate, and political instability persisted. Rather, the explanation for good performance seems to lie partly in the gap between demand and the previous dearth of services—especially trade, business services, and financial services—and partly in the personal qualities of the St. Petersburg pioneers for profit—their high level of education, technical experience, and entrepreneurial attributes.

A reasonable prognosis for services would be continuing growth with lower profit margins. On the one hand, the lack of a well-developed service sector has an advantage; the development of small and medium-size firms has not been blocked by large, preexisting firms whose managers collude with politicians and the bureaucracy to preserve their economic monopolies. On the other hand, profit margins will be squeezed as competition continues to increase; private service entrepreneurs may find it more difficult to survive. (See McDermott and Mejstrik (1992) for an account of difficulties faced by small and medium-size manufacturing enterprises in Czechoslovakia during the early stages of transition.)

The survey findings have several clear policy implications. First is the need to establish a stable macroeconomic environment—an objective frequently mentioned by the entrepreneurs surveyed. Students of private sector development in Eastern Europe single out relative stability as the key element in private sector development in Poland, Hungary, and the Czech Republic and instability as the key element in delaying such growth in Romania, Bulgaria, and the Russian Federation. See Earle and others (1994) and Johnson and Loveman (1995). Entrepreneurs' aversion to high interest rates is one reason why rapid and broadly based stabilization is necessary.

Second, as part of the general business environment, a stable legal framework is also important, together with transparent and evenhanded implementation of government regulations. Complex or changing laws and procedures, confusing or conflicting instructions, favoritism, and bribe taking all work against small and medium-size firms that have less financial and human resources to deal with these obstacles than large firms do. Aside from regulations covering the entry and operation of businesses, those pertaining to the rent and renovation of real estate are important.

Third, given the relatively low access of small and medium-size firms to bank credit, and the possible conflict between credit expansion and macroeconomic stabilization, the next priority is arguably the creation of a well-designed tax code that permits and encourages self-financing. Private sector activities are expanding and clearly must form an important part of the tax base. But taxes need to be rationalized (for example, discriminatory treatment of trade and other services should be eliminated) and set at levels that are moderate enough to permit substantial reinvestment of profits. Specific tax relief for small and medium-size enterprises—perhaps on reinvestment of a limited amount of nominal profits—would allow these firms to grow even more quickly, generating more jobs and providing the distribution and business service infrastructure required to stimulate and facilitate private sector development in agriculture and manufacturing.

Fourth, continued privatization of real estate is an extremely important factor in private sector development. The experience in Eastern Europe demonstrates that the transfer of ownership in real estate assets to small private businesses may be the most important aspect of the privatization process—more

important than the transfer of the business itself (Earle and others 1994). The slowness of real estate privatization in the Russian Federation is no doubt a major barrier to faster growth of private firms.

Fifth, private firms need better access to financing in order to grow rapidly. This requires actions on both the demand and supply sides. Banks need to develop procedures and personnel to evaluate credit applications, and firms need to develop good business proposals. In regard to the latter, direct assistance would be helpful, although important issues arise on how such programs are designed and delivered. See de Melo and Ofer (1994), Kondratowicz and Maciejewski (1994), Piasecki (1992), and de Koning and Snijders (1992) for a variety of specific policy recommendations for direct assistance programs.

At the time of the St. Petersburg surveys, the private sector in the Russian Federation was growing without the support of the government or of Western business. Direct assistance may not have priority over macroeconomic stabilization, but if such assistance is warranted in established market economies, it is clearly justified in the Russian Federation, where the market infrastructure is still weak and private growth so important. Given the high education level of the St. Petersburg entrepreneurs, information, counseling, and training should have a good payoff there.

## References

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

Acs, Zoltan J., and David B. Audretsch. 1993. *Small Firms and Entrepreneurship: An East-West Perspective.* New York: Cambridge University Press.

Brown, Annette, Barry Ickes, and Randi Ryterman. 1994. "The Myth of Monopoly: A New View of Industrial Structure." World Bank Policy Research Working Paper 1331. World Bank, Policy Research Department, Washington, D.C. Processed.

de Koning, Aad, and Jacqueline Snijders. 1992. "Policy on Small- and Medium-sized Enterprises in Countries of the European Community." *International Small Business Journal* 10(3):25–39.

de Melo, Martha, and Gur Ofer. 1994. *Private Service Firms in a Transitional Economy: Findings of a Survey in St. Petersburg.* Studies of Economies in Transformation 11. Washington, D.C.: World Bank.

Earle, John S., Roman Frydman, Andrzej Rapaczynski, and Joel Turkewitz. 1994. *Small Privatization: The Transformation of Retail Trade and Consumer Services in the Czech Republic, Hungary, and Poland.* Prague: Central European University Press.

Easterly, William, Martha de Melo, and Gur Ofer. 1994. "Services as a Major Source of Growth in Russia and Other Former Soviet States." Policy Research Working Paper 1292. World Bank, Policy Research Department, Washington, D.C. Processed.

Giarini, Orio, ed. 1987. *The Emerging Service Economy.* Oxford: Pergamon Press.

Goskomstat. Various issues. *Social/Economic Situation of Russia.* Moscow: Goskomstat.

Johnson, Simon, and Gary Loveman. 1995. *Starting Over in Eastern Europe: Entrepreneurship and Economic Renewal.* Boston: Harvard Business School Press.

Kondratowicz, Andrzej, and Wojciech Maciejewski. 1994. *Small and Medium Private Enterprises in Poland.* Warsaw: Adam Smith Research Center.

McDermott, Gerald A., and Michal Mejstrik. 1992. "The Role of Small Firms in the Industrial Development and Transformation of Czechoslovakia." *Small Business Economics* (4):179–200.

McKay, John. 1970. *Pioneers for Profit: Foreign Entrepreneurship and Russian Industrialization, 1885–1913.* Chicago: University of Chicago Press.

Piasecki, Bogdan, ed. 1992. *Policy on Small and Medium-sized Enterprises in Central and Eastern European Countries: Survey and Comparative Study.* Organizing Committee, 19th International Small Business Congress, Warsaw.

Stone, Andrew, and Irina Novitskaya. 1993. "Ukrainian Private Enterprise: Profiting against the Odds." World Bank, Private Sector Development Department, Washington, D.C. Processed.

Storey, David. 1982. *Entrepeneurship and the New Firm.* New York: Praeger Publishers.

———. 1983. *The Small Firm: An International Survey.* New York: St. Martin's Press.

———. 1989. *The Performance of Small Firms: Profits, Jobs, and Failures.* London and New York: Routledge.

St. Petersburg Statistical Office. 1993. *Main Activity Indicators for Small Enterprises and Cooperatives in St. Petersburg and Leningrad Region in 1992.* St. Petersburg.

Webster, Leila, and Joshua Charap. 1993. *A Survey of Private Manufacturers in St. Petersburg.* World Bank Technical Paper 228. Washington, D.C.

Webster, Leila, with Juergen Franz, Igor Artimiev, and Harold Wackman. 1994. *Newly Privatized Russian Enterprises.* World Bank Technical Paper 241. Washington, D.C.

Wyznikiewicz, Bogdan, Brian Pinto, and Maciej Grabowski. 1993. "Coping with Capitalism: The New Polish Entrepreneurs." World Bank, Private Sector Development Department, Washington, D.C. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3: 451-475

# Apprenticeship Contracts, Small Enterprises, and Credit Markets in Ghana

Ann D. Velenchik

*This article uses data from a 1992 survey of manufacturing enterprises in Ghana to describe the importance of apprenticeship in the manufacturing sector and analyze the structure of the contracts in which apprenticeship takes place. The article presents three major findings. First, the training of apprentices is both a widespread activity and a part of the training of a large fraction of entrepreneurs and manufacturing workers. Second, two primary types of contracts are apparent in the data: those in which apprentices pay fees for their training and those in which they do not. Third, for those firms training apprentices, the choice of contract type is strongly correlated with other characteristics of the firm, particularly its use of credit. Apprenticeship fees are one among many informal sources of firm finance.*

Many studies of manufacturing enterprises in Africa, particularly those focused on small firms, have noted the importance of apprenticeship in the backgrounds of individual entrepreneurs (House, Ikiara, and McCormick 1993; Harris 1971), as a component of current labor forces (Liedholm and Mead 1986), and as a mechanism for providing training (Squire 1981; Berry 1985). In most of these studies, however, discussions of apprenticeship are incidental to the main subject, and attention has rarely been focused on the institution of apprenticeship itself. Although there is a small literature in education and anthropology focusing on the structure and educational content of individual firms' apprenticeship programs, the economics literature includes no systematic portrait of either the prevalence and scope of apprenticeship in manufacturing or the microeconomic details of the contracts involved. This article begins the process of filling in both those gaps by providing a quantitative assessment of the importance of apprenticeship in Ghanaian manufacturing and presenting an analysis of the contractual relationships between apprentices and employers that characterize this institution.

My analysis of the determinants of the structure of apprenticeship contracts places the institution of apprenticeship at the intersection of labor and financial

Ann D. Velenchik is with the Department of Economics at Wellesley College. She would like to thank Tyler Biggs, Paul Collier, Tom Downes, Chris Udry, participants in the North East Universities Development Consortium Conference and the African Studies Association Meetings, and three anonymous referees for their comments and their insights on this issue. The research was conducted, in part, while the author was Visiting Scholar in the Department of Economics at Northwestern University.

markets and highlights the relationship between firms' behavior in credit markets and contract choice in the apprenticeship arena. The discussion extends the examination of contracting behavior in manufacturing and of African economic institutions more generally. The growing literature on economic institutions in developing countries provides a strong source of motivation for this work.

Another motivation comes from the role of apprenticeship as a training institution in Africa, where skilled labor is generally considered to be quite scarce. The past two decades have seen a dramatic expansion of educational systems throughout the continent (see Psacharopoulos and Woodhall 1985 and Knight and Sabot 1990), many of which have included the establishment of government vocational and technical schools geared toward providing skilled workers to the manufacturing sector. In many countries these schools exist alongside the traditional apprenticeship system. A great deal has been written about the expansion of formal education in Africa, with particular attention on evaluating the impact of education on productivity, earnings, inequality, and growth (see Collier and Lal 1986; Hazlewood and others 1989; and Knight and Sabot 1990). Little is known, however, about the role of apprenticeship in generating human capital for the manufacturing sector. Understanding the structure of apprenticeship contracts is a first step toward analyzing how useful apprenticeship is in skill formation, productivity increase, and growth in the manufacturing sector.

This article presents three major findings. First, the training of apprentices is both a widespread activity and a part of the training of a large fraction of entrepreneurs and manufacturing workers. Second, two primary types of contracts are apparent in the data: those in which apprentices pay fees for their training and those in which they do not. The third finding is that, for those firms training apprentices, the choice of contract type appears to be strongly correlated with other characteristics of the firm, particularly its use of credit. Because firms charging fees are more likely to use informal rather than formal sources of working capital, apprenticeship fees are one among many informal sources of firm finance.

As background for the discussions of the prevalence and structure of apprenticeship contracts, section I presents a description of the data and the study from which it is drawn. A broad quantitative discussion of the incidence of apprenticeship is presented in section II. A description of contracting forms, their importance, and some characteristics of the individuals and firms that participate in them is presented in section III. Section IV presents an analytical framework for the examination of contract choice, and section V includes an application of this framework using the Ghanaian data. Finally, section VI presents a summary and, most important, raises some questions for future research in this area.

## I. The Data

The data used in this study are drawn from a survey of 185 manufacturing firms in Ghana conducted in the summer of 1992. The survey was the pilot

study for the Regional Program on Enterprise Development (RPED), a multiyear study of the manufacturing sector in several African countries (Burundi, Cameroon, Côte d'Ivoire, Ghana, Kenya, Rwanda, Tanzania, Zambia, and Zimbabwe). The RPED is organized by the World Bank and funded by the Canadian and several European governments. The fieldwork for the Ghana study was conducted by a team of researchers from the Centre for the Study of African Economies at Oxford University and the Department of Economics at the University of Ghana.

The RPED provides an overview of how manufacturing firms in these economies perform during the post–structural adjustment period and focuses on a wide variety of areas of firm behavior. The survey was administered to the firm owners or managers. It includes modules covering the history of the firm; current production and sales; the personal history of the entrepreneur; labor force structure, wages, and apprenticeship contracts; use and adaptation of technology; firm response to the regulatory environment; and experience with conflict and conflict resolution. The survey also includes a sample of each firm's workers and apprentices, providing an additional sample of 545 workers and 212 apprentices. This article uses data drawn from almost all of the modules; hence it allows for an examination of apprenticeship contracts within the context of a broad range of firm characteristics and activities.

This analysis focuses on Ghana only. Although comparative cross-country analysis is a long-term goal of this line of research, it seems best to begin with detailed single-country analyses. Ghana is a good starting point both because, as the pilot, the data were available for analysis and because preliminary analyses indicate that apprenticeship is more prevalent in Ghana than in the other RPED countries. This prevalence is consistent with other findings, which indicate that apprenticeship institutions are more common in West than in East Africa (Liedholm and Mead 1986).

Details about the sampling frame and sample construction are available in Baah-Nuakoh and Teal (1993). The sample was drawn from four manufacturing sectors: food and beverage processing; metalworking; textiles and garments; and woodworking and furniture. The background paper by University of Ghana (1992) indicates that these four sectors account for 77 percent of manufacturing enterprises and 65 percent of manufacturing employment, so this study gives quite broad coverage of the manufacturing sector as a whole. The firms included in the sample range widely in size and ownership, from individually owned small enterprises to subsidiaries of multinational corporations and parastatal operations.

## II. Apprenticeship: Prevalence and Concentration

The quantitative importance of apprenticeship in the manufacturing sector can be measured in three dimensions: the fraction of current entrepreneurs who were trained as apprentices, the fraction of workers in the subsample who were

trained as apprentices, and the fraction of firms that are currently training apprentices.

Apprenticeship is fairly widespread in Ghanaian manufacturing. Of the 151 individual entrepreneurs interviewed in the survey (managers of multinational enterprises, parastatals, and other corporate entities were omitted from this module), 82, or 56 percent, had been apprentices. Most (74 percent) had done their training in the field in which they were currently in business. Of the 545 workers in the sample, 45 percent had been trained as apprentices. Finally, 97 of the firms, or 52 percent, were training apprentices at the time of the survey. These findings are consistent with other small-enterprise surveys in the region. Liedholm and Chuta (1976) find that apprentices comprised 42 percent of the labor force of small firms in Sierra Leone. Liedholm and Chuta also note that most entrepreneurs in manufacturing had learned their trades as apprentices. Steel and Webster (1992) find that 44 percent of Ghanaian microentrepreneurs who entered business after 1984 had been apprentices, and these results are quite similar to those obtained here.

Apprenticeship training is most important in the small firms, as seen in table 1. Firms owned by former apprentices are smaller, and were smaller at the time of their founding, than firms owned by individuals who were not apprentices. Former apprentices account for 62 percent of owners of firms with fewer than thirty workers, but only for 23 percent of owners of firms with more than thirty workers. The mean current size of firms owned by former apprentices is fourteen workers, and the mean size at the time of founding of these firms was eight workers. The mean size of non-apprentice-owned firms is fifty-eight workers, having been an average of thirty-eight workers at founding. Similarly, the mean size of firms in which apprentices are trained is thirty-three workers, and the mean size of firms not training apprentices is seventy-seven workers. Apprenticeship is also found in the larger firms: table 1 indicates that 36 percent of the large firms in the sample were training apprentices. The nature of the apprenticeship programs in large firms differs substantially from that found in the smaller firms, as shown in section III.

Table 1. *Apprenticeship and Firm Size, Ghana, 1992*

| *Firm size (by number of workers)* | *Number of firms* | *Firms owned by former apprentices (percent)* | *Workers who were apprentices (percent)* | *Firms that train apprentices (percent)* |
|---|---|---|---|---|
| Micro (less than 5) | 40 | 54 | 37 | 37 |
| Small (5–29) | 81 | 66 | 57 | 72 |
| Medium (30–99) | 39 | 30 | 43 | 38 |
| Large (100 or more) | 25 | 9 | 30 | 36 |

*Source:* Author's calculations.

Table 2. *Apprenticeship and Sector, Ghana, 1992*

| *Sector* | *Number of firms* | *Firms owned by former apprentices (percent)* | *Workers who were apprentices (percent)* | *Firms that train apprentices (percent)* |
|---|---|---|---|---|
| Food | 42 | 20 | 21 | 17 |
| Textiles and garments | 41 | 69 | 46 | 56 |
| Woodworking and furniture | 54 | 59 | 53 | 68 |
| Metalworking | 48 | 55 | 54 | 62 |

*Source*: Author's calculations.

Apprenticeship training is least important in the food sector, but of similar importance in each of the other three sectors, as seen in table 2. The relative unimportance of apprenticeship in the food sector is driven in part by the size distribution of firms in this sector. Forty-eight percent of food firms are either large enterprises or microenterprises, the sizes in which apprenticeship is least common. In addition, however, only 17 percent of the small firms in the food sector employ apprentices. This is probably because of the low level of skills required by food processing activities (most of the firms are either microbakeries or large, automated beverage processors). The low level of skills is also evident in the fact that in food firms that train apprentices, the mean duration of the training is less than half the mean duration of training in other sectors.

Entrepreneurs who were themselves trained as apprentices are nearly twice as likely to train apprentices as are those who were not. Thirty-eight percent of the owners who were not apprentices were training apprentices at the time of the survey, compared with 74 percent of owners who were apprentices. Unsurprisingly, apprenticeship training is passed from generation to generation. Apprenticeship is widely considered to be training for self-employment, and that impression is borne out in these data as well.

Most of the apprentices in the sample are training in skilled production activities: as bakers in the food processing sector, as tailors and dressmakers in the garment sector, as welders in the metalworking sector, and as furniture makers and joiners in the wood sector. Some of the larger firms' apprentices are training in the maintenance and repair of equipment, but the dominant training received by apprentices is in productive crafts.

## III. Apprenticeship Contracts

Apprenticeship is a form of skill acquisition in the context of production, or on-the-job training. The apprentice, in addition to learning a trade, provides labor services to the firm. The relationship between the apprentice and the firm, therefore, includes two types of transactions, one being the provision of labor services by the apprentice to the firm, the other being the provision of training services by the firm to the apprentice. The two facets of the apprentice-firm

relationship give rise to two components of the financial relationship between apprentices and firms.

It might be expected that the labor services and training services are exchanged directly and payments made only when the values of the two services are unequal, so that monetary flows between the two parties are the net value of the compensation to the apprentice for labor minus that to the firm for training. In this case, there are three kinds of arrangements: payments by firms to apprentices when the value of labor services exceeds the value of training; payments by apprentices to firms when the value of training exceeds the value of labor services; and no payments of either type when the values of the labor and training services are equal. This approach to understanding the financial arrangements surrounding apprenticeship is clearly described in Coy (1989).

In reality, however, the monetary arrangements surrounding apprenticeship training in Ghanaian manufacturing are more complex. In addition to wage payments, many apprentices receive allowances for food, housing, clothing, or transportation. These payments are made both in cash and in kind but are identified by firms as distinct from wage payments. In larger firms, these allowances are paid to all workers and have been used as a way to generate nontaxable compensation.[1] In smaller firms, allowances are often paid to apprentices even when wages are not. Historically, apprentices often lived with their masters during training. Although this practice has largely died out, provision of basic living expenses by the master continues to be a common feature of these contracts. In the following analysis, these allowances will be included in apprentice earnings as payments by firms to apprentices, although we will measure them separately when possible.

More important, there are contracts in which apprentices both pay fees to the firm and receive wages, living allowances, or both, in return. In total, then, the ninety-seven firms in the sample that trained apprentices in 1992 offered contracts of four types: contracts with neither fees nor wage-allowance payments; contracts without fees but with wage-allowance payments; contracts with fees but without wage-allowance payments; and contracts with both fees and wage-allowance payments. The characteristics of these contracts, their relative importance in the sample, and some characteristics of the individuals and firms engaged in them are presented in table 3.

As shown in table 3, the two most common contract forms (second and fourth columns) involve payments of wages or allowances. The clear distinction between contract types is in the payment of fees rather than the payment of wages and allowances. The significance of this distinction will be explored in the next two sections.

Most firms in which apprentices are not paid (contract types in the first and third columns) are in textiles and garments. In these two contract types there is also a

1. The government recently passed a compensation consolidation act, so allowances will not be paid separately from wages in the future.

Table 3. *Characteristics of Apprenticeship Contracts, Ghana, 1992*

| *Sector and characteristic* | *No-fee-charging firms* | | *Fee-charging firms* | |
|---|---|---|---|---|
| | *With unpaid apprentices* | *With paid apprentices* | *With unpaid apprentices* | *With paid apprentices* |
| Number of firms | 6 | 35 | 16 | 43 |
| Firms by sector (percent) | | | | |
| Food | 1 | 4 | 2 | 0 |
| Textiles and garments | 3 | 1 | 9 | 10 |
| Woodworking and furniture | 0 | 15 | 3 | 19 |
| Metalworking | 2 | 14 | 2 | 12 |
| Mean firm size (number of workers and apprentices) | 42 | 49 | 12 | 13 |
| Mean monthly earnings of apprentices (cedis) | 0 | 13,302 | 0 | 5,442 |
| Mean monthly wages of apprentices (cedis) | 0 | 9,650 | 0 | 2,579 |
| Mean monthly allowances of apprentices (cedis) | 0 | 3,695 | 0 | 2,645 |
| Firms paying allowances but not wages (percent) | n.a. | 6 | n.a. | 58 |
| Mean total apprenticeship fee (cash and in-kind, cedis) | 0 | 0 | 35,100 | 33,985 |
| Mean cash fee paid at start of apprenticeship (cedis) | 0 | 0 | 28,125 | 17,037 |
| Mean cash fee paid at end of apprenticeship (cedis) | 0 | 0 | 4,062 | 8,536 |
| Mean in-kind fee paid at start of apprenticeship (cedis) | 0 | 0 | 1,275 | 3,087 |
| Mean in-kind fee paid at end of apprenticeship (cedis) | 0 | 0 | 1,637 | 4,090 |
| Mean duration of apprenticeship (years) | 1.7 | 2.2 | 2.6 | 2.9 |
| Mean age of apprentices (years) | 17 | 22 | 23 | 21 |
| Mean education of apprentices (years) | 11 | 11 | 11 | 10 |
| Mean share of workers who are female (percent) | 50 | 5 | 73 | 12 |

n.a. Not applicable.
*Source:* Author's calculations.

high concentration of female apprentices. In many of these seamstress operations, the apprentices not only pay initial fees and provide their own sewing machines, but they also pay monthly tuition. In essence, these firms operate largely as schools where young women spend the period between formal education and marriage. In fact, several of these young women told me that their dressmaking apprenticeship was a form of dowry, selected and paid for by their fathers, because these marketable skills would make them more attractive potential wives. The training of apprentices is a significant source of revenue for these entrepreneurs. Steel and Webster (1992) describe a seamstress who used expanded training of apprentices as a means of survival during periods of low demand for output.

Although garment firms are more likely to charge apprentices fees than are firms in other sectors, the practice is not limited to these firms. A significant fraction of firms in both woodworking and furniture and in metalworking also charge apprenticeship fees. This implies that a discussion of contracting behavior is more than an examination of sector-specific institutions and must be driven by factors that are common across sectors.

Firms that charge fees are substantially smaller than firms that do not. Apprenticeship fees are part of a training institution that seems to be unique to the informal sector. The basic characteristics of the institution of apprenticeship as it exists in both West and East Africa include the payment of fees, the provision of living allowances, and the marking of the end of training with a party hosted by the apprentice. These characteristics have been described elsewhere in the economics, education, and anthropology literatures (Berry 1985; Bas 1989; King 1977; Coy 1989). The basic characteristics are very similar across countries, although the institution appears to be most common in West Africa.

None of the authors, however, have attempted to explain why these traditional apprenticeship programs are structured as they are, although each has identified one characteristic as particularly salient. King focuses on the fact that under such arrangements in both Kenya and Nigeria "apprentices pay for their training" (King 1977: 27). Berry, in a discussion of apprenticeship among mechanics in Nigeria, where wages are not provided but living allowances are, states that "although apprentices were not very skilled or efficient, the fact that they did not have to be paid was crucial in a trade where receipts are likely to be irregular and working capital is therefore implicitly expensive" (Berry 1985: 142). Although neither statement is fully accurate for my sample, both reflect a significant factor driving the structure of apprenticeship contracts in small-scale manufacturing.

## IV. Human Capital Theory

Because apprenticeship is a particular form of on-the-job training, apprenticeship contracts can be analyzed using basic human capital theory, following the discussion of training in Becker (1975). Assuming that labor and training markets are competitive and capital markets are perfect, the cost of training, which is also the returns received by the firm for providing it, is exactly equal to the present value of the stream of returns generated by the training. Given that framework, some simple theory provides a mechanism for understanding the apprenticeship contracts described above.

Assume that training takes place in period 0, after which time the apprentice is qualified in a trade and works for one period. The training entails a cost, $C$, which includes the time spent by the master and other employees in working with the apprentice. During the training period, the apprentice's productivity is $MP_0$, which is the marginal product of an unskilled worker in this firm or outside. Once the training is completed, the apprentice's productivity rises to $MP_1$

if the apprentice stays in the firm, and $MP_1^*$ if the apprentice goes to work elsewhere. The assumption of competitive markets implies that $C = (MP_1 - MP_0)/(1 + r)$, where $r$ is the one-period interest rate.

The apprentice's productivity in the firm that provided the training in relation to that in a different firm provides a measure of the degree of firm-specificity of the training. Define an index of firm-specificity, $a$, as $a = (MP_1^* - MP_0)/(MP_1 - MP_0)$. Fully specific training, which raises the apprentice's productivity only if the apprentice stays in the firm, is equivalent to $a = 0$. Fully general training, which is as productive in other firms as it is in the firm providing it, corresponds to $a = 1$. Intermediate cases can be indexed by values between 0 and 1.

Investments in fully general human capital will be made entirely by workers, because the returns to such training cannot be captured by firms. As Becker (1975) and others point out (Hashimoto 1981; Hashimoto and Yu 1980), workers may also make part of the investment in specific human capital. Because such human capital has value only as long as the employment relationship is maintained, both parties have an interest in avoiding inefficient terminations of this relationship. Sharing this investment, so that both workers and firms bear some costs and reap some benefits, avoids inefficient separation decisions by both parties. By sharing the investment, a worker would accept a wage that is lower than marginal productivity during training and receive a share of the returns to this investment thereafter, generating a positively sloped wage profile. For simplicity, in the analysis that follows I will assume that firms bear all the costs of specific training, so that the apprentice's share of the total training cost is exactly proportional to the degree of generality of the human capital, generating an apprentice investment of $aC$. The thrust of the results, however, is the same if the worker also shares in the cost of specific human capital investments.

Although the apprentice will ultimately bear training costs of $aC$, a number of strategies can be used for financing this investment, each giving rise to particular contract forms. Payments can be made either as deductions from wages or as direct payment to entrepreneurs and can take place before, during, or after the training. Payments made before the training may be thought of as apprentice financing of training, payments made during training as pay-as-you-go financing, and payments made after training as employer financing. These payment strategies can occur in combination, but it is useful to begin by describing them independently.

Under a fully pay-as-you-go strategy, apprentices pay the full amount of their training cost through deductions from wages during the training period, receiving $W_0 = MP_0 - aC$. Once training is completed, workers receive $W_1 = MP_0 + a(MP_1 - MP_0)$ if they stay with the firm and $W_1^* = MP_1^*$ if they leave the firm and work elsewhere. Given the definition of $a$, $W_1 = W_1^*$. If $MP_0 - aC \geq 0$, this type of financing arrangement is consistent with contract types in which no fees are paid and zero or positive wages are received during training. If $MP_0 - aC < 0$, the contract is one in which wages are zero and fees are paid on a continuous basis during training. No contracts of this type were observed in these data.

With apprentice financing, the apprentice pays for the training before it begins. To finance the full amount, the apprentice pays a fee of $aC$ at the beginning of period 0, and receives a wage of $W_0 = MP_0$ during the training period. The second-period wages are the same as under the pay-as-you-go strategy. The apprentice-financing strategy would be manifested as the contract type in which both fees and wages during training are positive. Apprentices may also choose to finance only part of their training in advance, paying for the rest through deductions from wages during the training period. In that case the apprentice pays a fee of $F < aC$ before training begins and receives wages of $W_0 = MP_0 - aC + F$ during training. The post-training wages will be the same as in the strategies described above. This strategy is consistent with both contract types in which fees are paid, where the case with no wage payments is a special case in which $MP_0 - aC + F = 0$.

Employer-financing strategies allow apprentices to pay their share of training costs by deductions from wages in the second period. Under full employer financing, the apprentice bears none of the costs during training and receives a wage of $W_0 = MP_0$. At the end of training the apprentice owes the firm $aC(1 + r)$, which is paid as a deduction from the wage in the post-training period, which is then $W_1 = MP_0$. If the apprentice bears part of the cost during training, the apprentice receives $W_0 > MP_0 - aC$ but less than $MP_0$ during training, at the end of which time the apprentice owes the firm $[aC - (MP_0 - W_0)](1 + r)$. Repayment through wage deductions in the post-training period yields $W_1 = MP_0 + (1 + r)(MP_0 - W_0)$. The second term reflects the fact that the worker receives the return on that fraction of the cost of training actually paid through wage deduction during the training period. This wage is less than the wage the worker would receive in alternative employment.

Both apprentice- and employer-financing arrangements contain credit components and therefore give rise to default risk and moral hazard issues. These issues are not the same as those surrounding the sharing of investments in specific human capital. Those arrangements are made to avoid inefficient separations, but not failure to repay debts or to provide services that have already been paid for, as is the case regarding the financing of the apprentice's portion of the investment, which is the issue under consideration here. Mechanisms for dealing with default risk and moral hazard issues can be incorporated into contracts by reducing the incentives to default and providing contract enforcement mechanisms to deter individually desirable breaches of contract.

When apprentices finance their training by paying fees in advance, firms then have an incentive to fail to provide that training or to dismiss the apprentice before training is completed. One mechanism for reducing default incentives is for the apprentice to defer part of the payment of the fee to the end of the training period. In 40 percent of the firms that charge fees an average of 20 percent of the fees are deferred to the end of training. Older firms are less likely to have deferred fees, indicating that reputation factors may provide another disincentive for breach of contract. Although firms do not engage in repeated

contracts with any individual apprentice, they do need to continue to attract apprentices and hence must worry about reputation effects. Firms for which apprenticeship is more important also may find reputation more important, which is verified by the finding that firms in which apprentices form a larger fraction of the labor force are less likely to receive deferred fees.

Under employer financing, the apprentice has an incentive to borrow the training costs from the firm and then leave for a higher-paying job without paying those costs back. It should be noted that this situation is different from that of the firm's having financed investment in specific human capital and wishing to retain the worker in order to appropriate the returns. In that case, the worker has not borrowed from the firm, and the worker's earnings would not be enhanced by leaving, so there is no real problem of moral hazard. Under employer financing of general human capital, the worker can actually benefit from leaving before the debt is repaid. To prevent this occurrence, firms might require apprentices to post a performance bond at the beginning of the training period, to be refunded once the worker's debt has been paid off. Although such a scheme is possible, it seems unlikely, because an apprentice who could post a bond large enough to cover the debt would also be able to pay that amount as a fee, thereby financing that part of the training directly. An alternative type of bonding scheme is through deferred compensation, in which workers are paid lower wages during the training period and higher wages during the post-training period. By deferring some of the compensation to the later period, the firm induces the apprentice to stay. This mechanism would be indistinguishable in practice from shared investments in specific human capital, as both would generate upward-sloping wage profiles.

Moral hazard problems may be diminished by external contract enforcement mechanisms. Agreements between apprentices and masters that generate a bonding of the apprentice to the master while also defining the master's obligations may be made verbally or in writing. External enforcement of these agreements may be provided by the legal system, trade unions, crafts associations, and other social networks. Nearly half of the firms in the sample have unionized labor forces or belong to trade associations that set explicit rules governing apprenticeship. Among the firms in which fees are charged, it is common for either the beginning or the end of the apprenticeship, or both, to be marked by a ceremony hosted by the apprentice's family. This ceremony is generally the source of the in-kind fees recorded in the data and may be the reason for some of the deferred fees as well. The ceremony usually includes the provision of drinks by the apprentice and the pouring of libation to the ancestors. These ceremonies are attended by relatives, friends, and business associates of both the master and the apprentice. Masters and apprentices often come from the same social network, with 85 percent of the apprentices in the sample having found their masters through the recommendations of friends and relatives. The guests at the ceremony, including the ancestors, are witnesses to the contract itself, and the social network and kinship ties provide a strong external enforcement mechanism.

Apprenticeship contracts can be written to combine several of these financing strategies: the apprentice may finance part of the training through initial fees, part through deductions from wages, and part through lower wages in the post-training period. Although the characteristics of contracts can be observed, the combination of strategies underlying each contract cannot be. However, the circumstances under which each strategy might be expected to be chosen can be described and related to observed contract types.

## V. The Determinants of Contract Choice

The choice of financing strategies, and hence the choice of contract type, is influenced by the characteristics of the training itself as well as by attributes of the firms and individuals participating in the apprenticeship. The foregoing discussion indicated that a contract type that does not involve payments of fees by apprentices can be generated by both pay-as-you-go and employer-financing strategies. Therefore, the only clear distinction between strategies that can be made through an examination of contract types is between those that involve some apprentice financing through fee payments and those that do not. The following analysis consequently focuses on what determines the use of apprentice financing as evidenced by the charging of fees.

Before turning to a discussion of contract choice it is useful to present some evidence about the validity of interpreting apprenticeship fees as apprentice financing. The cost of training includes both time and equipment, and these data indicate that firms charging fees are more than three times as likely to require their apprentices to provide their own tools than are firms that do not charge fees. In providing their own tools, these apprentices are financing the capital costs of their training, so it seems reasonable to interpret the fees as payment for the labor and materials costs of the training.

Contract choice should be influenced by the degree of specificity of the training. The more specific the training, the lower the cost of training to be borne by the apprentice. This would make it easier for the apprentice to finance the training through deductions from wages, as in the contract types without fees. More specific training also gives rise to steep wage profiles designed to induce the worker to stay with the firm. Such profiles would also be used when employers finance the acquisition of general training, which is the second strategy giving rise to contract types without fees. Thus, contracts without fees should, on average, involve a greater fraction of specific training.

The data set does not include information about the content of training; I use information about the post-training behavior of apprentices as a proxy. The more specific the training, the greater the probability that the apprentice will continue working for the firm after training is completed. If the training offered by firms that do not charge fees is indeed more specific, a larger fraction of those apprentices would continue in the firm. The data set provides three means

Table 4. *Contract Type and Workers Remaining in Their Apprenticeship Firms, Ghana, 1992*
(percent)

| *Item* | *Fee-charging firms* | *No-fee-charging firms* |
|---|---|---|
| Mean share of surveyed apprentices who finished training the previous year and stayed on at the firm | 20 | 76 |
| Mean share of current apprentices surveyed who plan to stay on after completing their training | 12 | 50 |
| Share of surveyed workers who are working in the firm in which they did an apprenticeship and paid a fee for it | 28 | n.a. |
| Share of surveyed workers who are working in firms other than that in which they did an apprenticeship and paid a fee for it | 55 | n.a. |

n.a. Not applicable.
*Source:* Author's calculations.

for examining this relationship, and the relevant data are presented in table 4. These data support the hypothesis that firms not charging fees offer more specific training than do firms charging fees.

It would also be expected that firms offering specific training or employer financing of general training would offer post-training earnings profiles designed to discourage quits. That is, more specific training should be associated with steeper earnings profiles providing higher returns to seniority. This is indeed the case for these data, in which the returns to an additional year of seniority among workers in firms in which apprentices are not charged fees is more than twice as high (4.6 percent compared with 2.0 percent) as in firms that charge fees.[2] This behavior is consistent with Hashimoto's (1981) argument about the sharing of investments in specific human capital and is consistent as well with the foregoing discussion of employer financing.

It should be noted that decisions about the type of training to offer and the financing strategy involved may be made simultaneously. A firm offering more specific training will bear a larger fraction of the costs, and the ability to do so may be driven by the same factors determining the firm's use of employer financing of the worker's share of costs. Total training costs might also be relevant to contract choice. In particular, lower training costs make it more likely that the apprentice can pay through deductions from wages during the training period. There is no clear association, however, between the level of training costs and the choice between apprentice and employer financing. Once costs are too high to be borne easily as wage deductions, there is no obvious reason to expect

2. These estimates were obtained by estimating an earnings function for the 191 individuals in the sample of workers who are employed by firms currently training apprentices. The returns to experience were allowed to differ according to the fee-charging practice of the individuals' employers. *t*-tests indicate that the coefficients are different at a 5 percent significance level.

apprentice or employer financing to be more likely. Similarly, the duration of training, which might affect costs, is also influenced by financing decisions. Employer financing requires that the apprentice stay on after training is completed and would give rise to longer measured apprenticeships for any given training costs. Although table 3 indicates that apprenticeships in which fees are charged are of longer average duration, this could be due to a number of factors, including the more general nature of the training, and cannot be clearly related to costs.

In addition to the characteristics of training, contract choice is also influenced by the characteristics of firms and apprentices. The evidence in table 3 indicates that firms charging fees are significantly smaller on average than those not charging fees. Small size in itself, however, is not a characteristic that should exert independent influence on contract choice. Firm size is endogenous and determined by a number of other factors that may also influence the choice of contract type. A principal task of the rest of this analysis will be to identify factors that, although they may be correlated with size, are more reasonable sources of differences in firm behavior.

One set of possibilities is related to the general labor market behavior of larger firms. There is a large literature (see Brown and Medoff 1989) documenting the widespread existence of large earnings differentials associated with employment in larger firms. The factors that cause large firms to "overpay" their workers may also cause them to "undercharge" their apprentices, in which case contract choice is simply an extension of a broader class of uncompetitive labor market behavior. If larger firms do not charge fees as part of a general pattern of paying rents to their employees, then explanations of contract choice would rely on explanations of the employer-size wage premium. Although this article does not discuss all of the possible sources of the employer-size wage premium, several of these explanations are also relevant here.

Do firms that do not charge fees provide rents to their workers and apprentices? The evidence in table 3 indicates that apprentices who are not charged fees do indeed earn more during their apprenticeships than their fee-paying counterparts do. But these differences could be generated by differences in the degree of specificity of the training and overall training costs, and thus cannot be easily interpreted as rents. Apprentices who are not charged fees may well be advantaged, but we cannot measure this advantage using current earnings without having data on training costs.

There is strong evidence, however, that nonapprenticed employees are advantaged by employment in larger firms. Jones and Teal (1994) find evidence of large-employer size effects in estimates for the whole sample of workers. They estimated that the premiums for small, medium, and large firms in relation to microenterprises are 21, 38, and 55 percent, respectively. Estimates for the subsample of workers employed in firms offering apprenticeship also show a substantial advantage to employment in larger firms. Medium-size firms pay 40 percent more than firms with fewer than thirty workers do, and large firms pay 73 percent more.

The fact that employees of larger firms earn rents does not necessarily imply that employees of all firms in which no fees are charged also earn rents, because the correlation between contract choice and size is not perfect. Among the twenty-four firms with more than thirty workers, five charge fees, and there is no statistically significant difference in wages between large firms that charge fees and those that do not. Among the seventy-three firms with fewer than thirty workers, twenty-one do not charge their apprentices. Interestingly, among the smaller firms, those that do charge their apprentices actually pay their workers more. Overall, there is no clear evidence that all firms that do not charge fees also pay rents to their employees. To the extent that the decision to charge fees is strongly correlated with firm size, however, broader explanations of large-firm behavior, particularly in regard to wage premiums, may provide insight into contract choice.

A source of these wage premiums that may also explain contract choice is unionization. Larger firms are more likely to be unionized, and wages in union firms are 9 percent higher than wages in non-union firms (Jones and Teal 1994). If unions also cause firms to undercharge for training, or if unions forbid the charging of fees, unionization would also drive contract choice. Seventeen of the firms in this sample are unionized, and none of them charge fees. Although it is possible that unionization and contract choice are both caused by a third factor (perhaps the profitability of the firm), it seems reasonable to attribute at least part of the decision not to charge fees to union influence. But there are also twenty-three non-union firms that do not charge fees, so the union factor is not a complete explanation of contract choice.

Another common explanation of employer-size wage premiums is rent sharing. If larger employers are able to earn economic profits or rents, they may share these rents with their workers and apprentices. There are several potential sources of rents. Market power in output markets would allow these firms to charge higher prices. These data include no obvious measure of market power, so we cannot directly explore the relation between market power, profits, and contract choice. Quantity discounts on raw materials are another source of rents. Differences in the nature of raw materials used across firm sizes make direct comparisons of raw materials costs impossible, although the subsequent discussion will explore differences in value added that may capture variation in input costs in relation to output values.

If firms face differential costs of financing their working capital needs, this would also give rise to rents for those firms having cheaper or easier access to funds. Differential costs of working capital would also give rise to variation in contract choice if, as described above, choice of contract type is based on a decision about financing of training. Employer financing of training is a form of informal credit from firms to apprentices, and it is likely that firms finance training when they have access to funds at a lower interest rate than that available to apprentices or when firms are able to access more funds than apprentices can. This is the idea that Berry (1985) reflects in her discussion of the expense of working capital to firms using apprentices.

Table 5. *Contract Type and Firm Credit Market Behavior, Ghana, 1992*
(percent unless noted otherwise)

| Variable | *No-fee-charging firms* | | *Fee-charging firms* | | *t-test for difference in means*[a] |
|---|---|---|---|---|---|
| | *With unpaid apprentices* | *With paid apprentices* | *With unpaid apprentices* | *With paid apprentices* | |
| Value added per worker per year[b] (cedis) | 414,852 | 830,324 | 322,103 | 173,183 | |
| *Relations with banking sector* | | | | | |
| Share of firms receiving bank loans | 16 | 6 | 6 | 0 | |
| Share of firms receiving overdraft facilities | 33 | 17 | 6 | 7 | ** |
| Share of firms holding bank accounts | 100 | 79 | 75 | 63 | * |
| *Relations with suppliers* | | | | | |
| Share of firms receiving credit from suppliers | 33 | 35 | 19 | 32 | |
| Mean length of supplier credit for those receiving it (days) | 45 | 51 | 9 | 29 | ** |
| Mean implied interest on supplier credit for those receiving it | 7 | 4 | 10 | 13 | ** |
| *Informal credit markets* | | | | | |
| Share of firms receiving informal credit | 33 | 26 | 44 | 51 | ** |
| Share of firms using informal safekeeping services or mobile bankers | 0 | 4 | 19 | 18 | ** |
| Share of firms participating in informal savings groups | 0 | 3 | 0 | 7 | |
| *Relations with customers* | | | | | |
| Share of firms making sales on cash basis | 50 | 38 | 75 | 66 | ** |
| Share of firms making sales on credit basis | 67 | 62 | 44 | 27 | ** |
| Share of firms receiving advance payment from customers | 50 | 44 | 56 | 68 | ** |

*Note:* The sample size is ninety-seven firms.

a. Difference in means between fee-charging and no-fee-charging firms (for firms paying their apprentices a wage and those that do not). * and ** indicate the test is significant at 10 and 5 percent, respectively.

b. Data were available for only ninety-one firms for this variable.

*Source:* Author's calculations.

Although a firm's access to or cost of financing cannot be measured, its behavior in other arenas involving financing can be used as an indicator. If the hypothesis above is correct, firms that charge apprenticeship fees will differ from firms that do not in terms of their use of alternative sources of finance, their relationships with other business partners, and their behavior and experience in credit markets generally. In particular, firms that charge fees to their apprentices would be expected to differ from other firms in their mechanisms for obtaining working capital.

Normally, businesses meet their working capital needs through some combination of their own liquidity and formal sources of credit such as banks and input suppliers.[3] Firms with weak cash flow and lacking access to bank and supplier credit will use informal sources of finance, such as borrowing from money lenders, or holding funds with informal savings collectors or mobile bankers in order to have access to funds at month's end, or requiring advance payments from customers. This use of informal finance can be seen as a spillover from formal financing markets, in which agents who are rationed out of formal sources turn to informal markets.

From the perspective of contract choice, firms with good cash flows or access to formal finance at reasonable costs would be expected to be able to finance their apprentices' training and would be less likely to charge fees. These firms would also face lower costs for working capital and could share this advantage with workers by not charging fees. The charging of fees to apprentices, which is similar to the practice of requiring advance payments from customers, is itself a form of informal finance, because the fees are generally paid before the training is received. Therefore, firms charging apprentices fees would also probably use other informal sources of credit. To the extent that these sources are more expensive than formal sources, these firms would not be earning rents that could be shared with workers.

Data about several features of firms' credit and financial market behavior are presented in table 5. Although it would be desirable to include a measure of the strength of the firms' cash flows, the data set does not include information that would allow such a measure to be constructed. The average value added (value of sales minus cost of raw materials) per worker per year is a blunt measure of the availability of rents to the firm and is relevant for examination of the rent-sharing hypothesis. These data were available for only ninety-one of the ninety-seven firms. Firms that charge for apprenticeship have significantly lower value added per worker, indicating less access to rents than firms that do not charge.

A firm is considered to have received bank credit if it had either a loan or an overdraft facility at a bank or other formal financial institution in the past three years. Bank accounts include both demand and time deposits at all formal financial institutions. From table 5, firms that charge for apprenticeship are less in-

3. Supplier credit lies at the border between formal and informal credit. For this analysis it is treated as formal because supplier credit is a common feature of business practices in developed manufacturing sectors throughout the world and is part of their formal credit practices.

volved with the banking sector, both as borrowers and as depositors, than are firms offering the other form of contract, and this indicates limited use of the banking system as a source of working capital.

Firms were asked to describe the form of payment (cash, credit, or advance payment) they used with the primary suppliers of their three major raw materials and then to give the details of the last transaction of each type they conducted. Each firm can engage in all three types of transactions. A firm is said to receive supplier credit if any of its principal raw materials suppliers extend credit. Table 5 presents two characteristics of supplier credit for those firms receiving it. The first is the mean value of the elapsed time between delivery of the goods and final payment, and the second is the implicit interest rate paid for the credit, measured in terms of the cash discount forgone. Although firms that charged their apprentices were no less likely to have received credit from suppliers than were firms that did not charge, they did get shorter repayment periods and pay higher interest rates, in terms of forgone cash discounts, than the other firms did. This evidence supports the hypothesis that firms without strong cash flows or access to credit from the formal sector are more likely to charge fees.

Firms were asked if they had borrowed money on the informal credit market (this includes loans from friends, relatives, suppliers, clients, other enterprises, and informal moneylenders) in the past three years. They were also asked if they had participated in any informal savings groups or had used the services of a mobile banker or savings collector—an individual who holds cash for safekeeping in exchange for a fee. As shown in table 5, firms that charge fees are more likely to be involved in the informal credit market, both as borrowers and as users of savings collectors than are firms that do not.

Firms were asked to describe the form of payment they used with their most important customers or the form they used most frequently when they had a large number of infrequent customers. They were also asked to describe the details of their most recent transaction of each type. From table 5, firms charging apprenticeship fees were more likely to deal in cash or advance payment with their clients and less likely to extend them credit. Because the elapsed time between order and delivery of goods is far less than the mean duration of apprenticeship, it seems reasonable to conclude that firms that do not finance their production costs out of working capital would be unlikely to finance training. Worker financing of training is itself a source of informal credit. The evidence indicates that firms that are more likely to use informal credit use several forms of informal credit, including borrowing from their customers in the form of advance payment and borrowing from their apprentices in the form of advance fees. Contract choice in apprenticeship markets, then, seems to mimic financing choices in other arenas and to be part of an overall financing strategy emphasizing informal credit arrangements.

The relation between these credit market behaviors and the choice of apprenticeship contract is summarized in table 6, which presents the results of a probit estimation of the influence of such capital market features on the probability

Table 6. *Probit Estimates of the Influence of Capital Market Features on the Probability That a Firm Charges a Fee to Apprentices, Ghana, 1992*

| | *Estimation excluding value added per worker* | | *Estimation including value added per worker* | |
|---|---|---|---|---|
| *Variable* | *Coefficient* | *Mean effect on probability*[a] | *Coefficient* | *Mean effect on probability*[a] |
| *Sector dummy* | | | | |
| Food | 0.156 | 0.058 | 0.557 | 0.115 |
| | (0.214) | | (0.697) | |
| Woodworking and furniture | −0.587 | −0.181 | −0.596 | −0.071 |
| | (1.188) | | (1.107) | |
| Metalworking | −1.170** | −0.285 | −1.058** | −0.100 |
| | (2.309) | | (1.892) | |
| Firm size | −0.0005 | −0.0001 | −0.0017 | −0.0002 |
| | (0.262) | | (0.744) | |
| Value added per worker | | | −1.44e-6** | −3.20e-7 |
| | | | (2.865) | |
| *Capital market behavior dummy*[b] | | | | |
| Firm received bank credit | −0.475 | −0.148 | −0.193 | −0.034 |
| | (0.931) | | (0.307) | |
| Firm holds bank account | −0.641* | −0.194 | −0.374 | −0.049 |
| | (1.633) | | (0.906) | |
| Firm received informal credit | 0.462 | 0.180 | 0.270 | 0.049 |
| | (1.326) | | (0.697) | |
| Firm used mobile banker | 1.275* | 0.465 | 1.149 | 0.289 |
| | (1.805) | | (1.509) | |
| Firm made sales on cash basis | 1.011** | 0.384 | 1.002** | 0.242 |
| | (2.771) | | (2.595) | |
| Firm made sales on credit basis | −1.084** | −0.273 | −1.002** | −0.297 |
| | (2.863) | | (2.391) | |
| Firm made sales with advance payment | 0.927** | 0.356 | 0.732* | 0.161 |
| | (2.305) | | (1.639) | |
| Constant | 0.542 | n.a. | 1.022 | n.a. |
| | (0.829) | | (1.435) | |
| Number of observations | 97 | | 91 | |
| $\chi^2$ | 45.13 | | 48.36 | |
| Pseudo $R^2$ | 0.3432 | | 0.4020 | |
| Log likelihood | -43.175 | | −35.959 | |

* Significant at 10 percent.

** Significant at 5 percent.

n.a. Not applicable.

*Note:* The dependent variable takes the value 1 if the firm charges a fee. *t*-statistics are in parentheses.

a. The mean effect on probability is the effect of each variable on the probability a firm charges fees, calculated as the mean of the effect across all observations.

b. The value is 1 if the variable holds, 0 otherwise.

*Source:* Author's calculations.

that a firm charges a fee to apprentices. The dependent variable takes the value 1 if the firm charges a fee. The independent variables include sectoral dummies, firm size, and value added per worker, followed by a series of dummy variables that take the value 1 if the firm made sales on a cash basis, made sales on a credit basis, made sales with advance payment, received bank credit, held bank accounts, received informal credit, and used informal mobile banking services.[4] The effect of each variable on the probability that a firm charges fees, calculated as the mean of the effect across all observations, is reported in the second column for each set of estimates. For the continuous variables, this represents the change in probability in response to a one-unit change in the value of the variable. For the dummy variables, it represents the difference in the probability of charging a fee generated by having rather than not having that characteristic. The table reports estimates of two versions of the probit, one that excludes and one that includes the value-added measure.

There is substantial correlation among the independent variables, particularly those related to banking and informal credit, and this correlation results in imprecise estimates. One mechanism for dealing with the imprecision is to compress some of the measures. The results in table 7 were constructed by replacing the two banking measures with a single dummy that takes the value 1 if the firm had any relationship (borrowing or deposits) with formal banking institutions and replacing the two informal credit markets with a single dummy that takes the value 1 if the firm had any relation (borrowing or using mobile bankers) with informal credit institutions.

These results indicate that a relation between credit market behavior and the use of apprenticeship fees exists and that firms with more access to rents, as measured by value added, are indeed less likely to charge fees. The magnitudes of the credit market effects are smaller in the estimates where value added is included. This is not unreasonable, given that profitability, as proxied by value added, would be a determinant of firms' access to formal credit, generating some substantial correlation among the variables in these regressions.

The credit market results show that firms that do not provide financing to their customers are less likely to provide financing to apprentices. Use of the formal banking sector is negatively correlated, but use of informal credit markets is positively correlated, with the charging of apprenticeship fees. It must be emphasized, however, that this shows only correlation, not causation. Both credit market behavior and contract choice would be influenced by factors that control the firm's access to and cost of funds. The probit provides results consistent with the notion that worker financing of apprenticeship results from poor firm access to formal sources of working capital. The probit analysis is not a test of that hypothesis, however, nor is it a direct test of the hypothesis that firms which do not charge apprenticeship fees are engaging in rent sharing with their workers.

4. Because there is no variation across contract types in the use of supplier credit, supplier credit indicators were not included in the probit.

Table 7. *Probit Estimates with Formal and Informal Banking Dummy Variables, Ghana, 1992*

| Variable | Estimation excluding value added per worker | | Estimation including value added per worker | |
|---|---|---|---|---|
| | Coefficient | Mean effect on probability[a] | Coefficient | Mean effect on probability[a] |
| *Sector dummy* | | | | |
| Food | 0.125 | 0.048 | 0.399 | 0.136 |
| | (0.175) | | (0.510) | |
| Woodworking and furniture | −0.531 | −0.208 | −0.479 | −0.147 |
| | (1.109) | | (0.927) | |
| Metalworking | −1.120** | −0.397 | −0.948* | −0.256 |
| | (2.280) | | (1.785) | |
| Firm size | −0.0004 | −0.0001 | −0.0010 | −0.0002 |
| | (0.211) | | (0.440) | |
| Value added per worker | | | −1.32e-6** | −2.95e-7 |
| | | | (2.663) | |
| *Capital market behavior dummy*[b] | | | | |
| Firm had relationship with formal bank | −1.024** | −0.372 | −0.646 | −0.190 |
| | (2.366) | | (1.450) | |
| Firm had relationship with informal financial sector | 0.701** | 0.323 | 0.488 | 0.167 |
| | (2.011) | | (1.278) | |
| Firm made sales on cash basis | 1.068** | 0.323 | 0.968** | 0.318 |
| | (3.094) | | (2.626) | |
| Firm made sales on credit basis | −0.962** | −0.354 | −0.900** | −0.247 |
| | (2.675) | | (2.252) | |
| Firm made sales with advance payment | 0.893** | 0.287 | 0.701* | 0.237 |
| | (2.309) | | (1.625) | |
| Constant | 0.689 | n.a. | 1.121 | n.a. |
| | (1.030) | | (1.589) | |
| Number of observations | 97 | | 91 | |
| $\chi^2$ | 44.74 | | 47.56 | |
| Pseudo $R^2$ | 0.3403 | | 0.3954 | |
| Log likelihood | −43.368 | | −36.359 | |

* Significant at 10 percent.
** Significant at 5 percent.
n.a. Not applicable.
*Note:* *t*-statistics are in parentheses.
a. The effect of each variable on the probability that a firm charges fees, calculated as the mean of the effect across all observations.
b. The value is 1 if the variable holds, 0 otherwise.
*Source:* Author's calculations.

It would also be expected that characteristics of the apprentices influence the type of contracts in which they are found. The evidence in table 3 shows little difference in the educational characteristics or ages of apprentices, although it does indicate that women tend to be concentrated in contract types that do not involve wage payments but do involve fees. In addition to these personal charac-

Table 8. *Probit Estimates of the Probability an Apprentice Paid a Fee, Ghana, 1992*

| *Variable* | *Coefficient* | *Mean effect on probability* |
|---|---|---|
| Age (years) | 0.0389 | .0079 |
| | (0.197) | |
| Age squared | −0.0006 | −.0001 |
| | (0.137) | |
| Highest schooling-level dummy | | |
| Primary | −0.1014 | −.0403 |
| | (0.210) | |
| Middle school | 0.4422 | .1703 |
| | (1.151) | |
| Secondary school | −0.6054 | −.2278 |
| | (1.441) | |
| Apprentice is male | −0.6909** | −.2556 |
| | (2.196) | |
| Apprentice sends remittances to parents | −0.7003** | −.2586 |
| | (2.281) | |
| Apprentice receives cash from parents | 0.8768** | .3083 |
| | (2.755) | |
| Constant | 0.2086 | n.a. |
| | (0.088) | |

** Significant at 5 percent.

n.a. Not applicable.

*Note:* For the dummy variables, the value is 1 if the variable holds, 0 otherwise. The sample size is 209, the log-likelihood is −78.502864, the $\chi^2$ (8) is 49.93, and the pseudo $R^2$ is 0.2413. The absolute values of *t*-statistics are in parentheses.

*Source:* Author's calculations.

teristics, it would also be expected that the apprentices' ability to finance their apprenticeship influences contract choice. Most apprenticeship fees are paid by parents, and many apprentices have little knowledge of the specific arrangements that were made. If parents are the main source of funds for apprentices, then it would be expected that the apprentices for whom fees are paid are more able to draw on parental resources. The survey included questions to the apprentices about their receipt of funds from and remittance of funds to their parents. These variables, along with more standard measures of human capital, are included in probit estimates of the probability an individual apprentice paid a fee. These estimates are presented in table 8.

As expected, male apprentices are less likely to have paid fees.[5] In the total sample, access to parental financial resources increases the likelihood of being in a paying apprenticeship by 31 percent; in contrast, apprentices who send cash home are

5. Given the differences between men and women in this sample, an attempt was made to analyze their behavior separately. Small sample size for the women (sixty-four) and lack of variation across them in personal characteristics such as education and age made estimation of the same probit impossible, but the available results for women do indicate the same positive relationship between receiving cash from parents and paying fees as is seen for the sample as a whole.

26 percent less likely to have paid fees. Other human capital variables, including education and age (a proxy for experience) do not exert statistically significant influence on contract choice, although it seems that the most educated individuals—those who have completed secondary school—are less likely to be in apprenticeships in which fees are charged. If the qualifications and abilities of the two groups—those paying and those not paying fees—of apprentices do differ, the differences are not captured in standard measures of human capital.

Overall, the evidence indicates that contract choice in apprenticeship, insofar as it is based on decisions about the financing of training, is influenced by the relative access to funds of firms and apprentices. In contexts such as manufacturing in Ghana, where firms are small and informal and not necessarily clearly distinguishable from households, it seems reasonable that some firms would be unable to finance training as easily as their apprentices can.

## VI. Summary and Directions for Future Research

A great deal of training takes place on the job in the manufacturing sector in Ghana, particularly in small firms. Although this article does not explore the content of this training and therefore cannot assess its quality, it seems clear that the training of a new generation of craft workers is an important part of economic activity in the small-scale manufacturing sector. Equally important, the evidence indicates that these firms provide relatively general training, generating skills that are useful in other firms and that seem to provide apprentices with preparation for entrepreneurship. Both King (1977) and Berry (1985) note that apprenticeship in a trade is a route to self-employment, and the importance of apprenticeship training in the backgrounds of the entrepreneurs in this sample supports this idea. The important implication of this fact is that apprenticeship should not be excluded from general analyses of vocational training in Ghana and that the human capital being generated in small-scale manufacturing should be considered when policy in this area is made.

The human capital approach taken here provides insights into the characteristics of traditional apprenticeship programs. This article has expanded the ideas of King and Berry by looking more closely at the financing arrangements involved in apprenticeship contracts. King's contention that apprentices pay for their training is correct, because all apprentices ultimately bear the cost of the general portion of their training. But what is truly interesting about fee-paid apprenticeships is that they involve apprentice financing as well. Berry's contention that the usefulness of apprentices to firms is that they do not have to be paid is insightful because it focuses on the importance of constraints on working capital in motivating the structure of apprenticeship contracts. Although these contract forms do alleviate constraints on working capital, the fact that apprentices do not receive wages is unlikely to actually increase firm profits. The small payments received by apprentices would be exactly offset by their low productivity. Only if apprentices received less than their productivity minus training

costs would this practice actually enhance firm profits. This scenario would be possible only if the firm charged the apprentices more than their training cost, which the firm could do only if it had some market power. Given the large number of small firms engaged in the training of apprentices in any trade, this kind of market power seems unlikely.

The correlation between apprenticeship contract choice and other credit market behavior neither shows causation nor identifies the root cause of differences in these behaviors across firms. It does indicate, however, that the practice of firm financing of training is related to the way in which firms finance other activities. This implies that changes in the functioning of urban capital markets could influence patterns of human capital formation as well as cause the more well-known effects on technological change, employment, and growth.

This analysis has provided a first step to understanding the role of apprenticeship in manufacturing in Ghana. It does, however, neglect some important questions, including the determinants of the firm's decision to train apprentices and to offer general or specific training and the influence of apprenticeship training on the employment opportunities, career paths, and wages of apprentices. Future work must address these issues. In an era of budgetary crisis and structural adjustment that are forcing many African countries to reduce their education budgets, the role of alternative training institutions such as apprenticeship must be explored.

## References

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

Baah-Nuakoh, Amoah, and Francis Teal. 1993. "Economic Reform and the Manufacturing Sector in Ghana: Final Report." World Bank, Regional Program on Enterprise Development (RPED), Washington, D.C. Processed.

Bas, Daniel. 1989. "On-the-Job Training in Africa." *International Labour Review* 128(4):485–96.

Becker, Gary. 1975. *Human Capital.* 2d ed. Chicago: University of Chicago Press.

Berry, Sara. 1985. *Fathers Work for Their Sons.* Berkeley: University of California Press.

Brown, Charles, and James Medoff. 1989. "The Employer Size–Wage Effect." *Journal of Political Economy* 97(October):1027–59.

Collier, Paul, and Deepak Lal. 1986. *Labour and Poverty in Kenya.* New York: Oxford University Press.

Coy, Michael, ed. 1989. *Apprenticeship: From Theory to Method and Back Again.* Albany: State University of New York Press.

Harris, John. 1971. "Nigerian Entrepreneurship in Industry." In Peter Kilby, ed., *Entrepreneurship and Economic Development.* New York: The Free Press.

Hashimoto, Masanori. 1981. "Firm-Specific Human Capital as a Shared Investment." *American Economic Review* 71(3):475–82.

Hashimoto, Masanori, and Ben Yu. 1980. "Specific Capital, Employment Contracts, and Wage Rigidity." *Bell Journal of Economics* 11(2):536–49.

Hazlewood, Arthur, Jane Armitage, Albert Berry, John Knight, and Richard Sabot. 1989. *Education, Work, and Pay in East Africa.* Oxford, U.K.: Clarendon Press.

House, William, Gerrishon Ikiara, and Dorothy McCormick. 1993. "Urban Self-Employment in Kenya: Panacea or Viable Strategy?" *World Development* 21(July):1205–23.

Jones, Tracy, and Francis Teal. 1994. "Earnings and Wages in the Ghanian Manufacturing Sector: A Preliminary Report on a Panel Data Survey." University of Oxford, Centre for the Study of African Economies, Oxford, U.K. Processed.

King, Kenneth. 1977. *The African Artisan.* London: Heinemann.

Knight, John, and Richard Sabot. 1990. *Education, Productivity, and Inequality: The East African Natural Experiment.* New York: Oxford University Press.

Liedholm, Carl, and Enyinna Chuta. 1976. "The Economics of Rural and Urban Small Scale Industries in Sierra Leone." African Rural Economy Paper 14. Michigan State University, East Lansing. Processed.

Liedholm, Carl, and Donald Mead. 1986. "Small Scale Industry." In Robert Berg and Jennifer Whitaker, eds., *Strategies for African Development.* Berkeley: University of California Press.

Psacharopoulos, George, and Maureen Woodhall. 1985. *Education for Development: An Analysis of Investment Choices.* New York: Oxford University Press.

Squire, Lyn. 1981. *Employment Policy in Developing Countries: A Survey of Issues and Evidence.* New York: Oxford University Press.

Steel, William, and Leila Webster. 1992. "How Small Enterprises in Ghana Have Responded to Adjustment." *The World Bank Economic Review* 6(3):423–38.

University of Ghana. 1992. "Country Background Paper: Economic Reform and the Manufacturing Sector in Ghana." Department of Economics, Legon, and University of Oxford, Centre for the Study of African Economies, Oxford, U.K. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3: 477–508

# Inequality and Growth Reconsidered: Lessons from East Asia

Nancy Birdsall, David Ross, and Richard Sabot

*East Asian economies have experienced rapid growth over three decades, with relatively low levels of income inequality, and appear to have also achieved reductions in income inequality. We argue that policies that reduced poverty and income inequality, such as emphasizing high-quality basic education and augmenting labor demand, also stimulated growth. Closing two virtuous circles, rapid growth and reduced inequality led to higher demand for, and supply of, education. Moreover, low levels of income inequality may have directly stimulated growth. We present cross-economy regression results that are consistent with a positive causal effect of low inequality on economic growth and with low inequality of income as an independent contributing factor to East Asia's rapid growth. We conclude that policies for sharing growth can also stimulate growth. In particular, investment in education is a key to sustained growth, both because it contributes directly through productivity effects and because it reduces income inequality.*

It has long been the conventional wisdom that there is a tradeoff between augmenting growth and reducing inequality. Two explanations are given for the view that an unequal distribution of income is necessary for, or the likely consequence of, rapid economic growth. The first, following Kaldor (1978), is that because a high level of savings is a prerequisite of rapid growth, income must be concentrated in the hands of the rich, whose marginal propensity to save is relatively high. The second, following Kuznets (1955), is that as labor shifts from sectors with low productivity to sectors with high productivity, aggregate inequality must initially increase substantially and only later decrease; Robinson (1976: 437) observed that this pattern had "acquired the force of economic law."

The conventional wisdom has in fact been repeatedly questioned. Twenty years ago, prior to the fulfillment of East Asia's remarkable potential, Ahluwalia

Nancy Birdsall is Executive Vice-President at the Inter-American Development Bank, David Ross is with the Department of Economics at Bryn Mawr College, and Richard Sabot is with the Center for Development Economics at Williams College and the Policy Research Department at the World Bank. The authors are grateful to Kristin Forbes, Rebecca Foster, and Jennifer Keller for their able assistance. The authors also wish to thank for helpful comments the participants in the authors' session of the 1994 meeting of the American Economics Association and in seminars at St. Antony's College, Oxford; Williams College; Kennedy School of Government, Harvard University; and the Inter-American Development Bank, where this article was presented.

(1974: 13) concluded, on the basis of the cross-country data then available, that "[T]here is no strong pattern relating changes in the distribution of income to the rate of growth of GNP. . . . This suggests that there is little firm empirical basis for the view that higher rates of growth inevitably generate greater inequality." Anand and Kanbur (1993) demonstrate the sensitivity of the U-shaped relation between inequality and per capita income to changes in functional form and choice of observations. Alesina and Perotti (1994), Alesina and Rodrik (1994), Fields (1989), and Persson and Tabellini (1994) provide supporting evidence for why high inequality might constrain growth.

In this article we show how the recent experience of East Asia also casts doubt on the conventional wisdom. East Asian economies have had rapid growth over three decades with relatively low levels of income inequality, and although evidence on changes in income inequality is fragile, these economies appear to have also achieved reductions in income inequality. Our argument is straightforward: policies that reduced poverty and income inequality, such as emphasizing basic education and augmenting labor demand, also stimulated growth. Moreover, low levels of income inequality may have stimulated growth.

The rapid growth in the East Asian region has been unparalleled. Eight "high-performing Asian economies"—labeled HPAEs in World Bank (1993)—are Hong Kong, Indonesia, Japan, the Republic of Korea, Malaysia, Singapore, Taiwan (China), and Thailand. Not only have the HPAEs outperformed the industrial economies since 1960, but they have grown at higher rates than other developing economies (see table 1). They have also grown at higher rates than the current industrial economies did during their earlier periods of most rapid growth: the annual average compound growth rate of GDP per capita of industrial economies was 1.1 percent for 1820–70 and 1.4 percent for 1870–1913 (Maddison 1982). The HPAEs also stand out for persistence of rapid growth: although other developing economies have experienced rapid growth for several years, few have managed to sustain such growth over three decades (Easterly and others 1993).

The decline in poverty in East Asia has been equally remarkable. The percentage of people below the poverty line has fallen far more rapidly in East Asia than in other regions. For example, in the period from 1970 to 1987, the percentage of people below the poverty line in Indonesia fell from 58 percent (among the highest of all developing economies for which measures are available) to 17 percent (among the lowest of all developing economies)(World Bank, various years b). As would be expected, the absolute number of poor people has also declined significantly in the HPAEs (World Bank, various years b). Although declines in income inequality are more difficult to document, measures such as the Gini coefficient show improvements in each of the HPAEs from 1965 to 1990 (World Bank 1993). Findlay and Wellisz (1993); Kuo, Ranis, and Fei (1981); and Leipziger and others (1992) also cite evidence suggesting that, unlike in many other developing economies during this period, income distribution in East Asian economies either improved or, at the very least, did not worsen. (However, for some years

Table 1. *Average Annual Growth in Per Capita GNP, Selected Economies, 1960–80 and 1965–90*

| *Country* | *1960–80* | *1965–90* |
|---|---|---|
| *High-performing East Asian economies* | | |
| Hong Kong | 6.8 | 6.2 |
| Indonesia | 4.0 | 4.5 |
| Japan | 7.1 | 4.1 |
| Korea, Rep. of | 7.0 | 7.1 |
| Malaysia | 4.3 | 4.0 |
| Singapore | 7.5 | 6.5 |
| Taiwan (China) | — | 7.2 |
| Thailand | 4.7 | 4.4 |
| Average | 5.9 | 5.5 |
| *Selected developing economies* | | |
| Argentina | 2.2 | −0.3 |
| Bangladesh | — | 0.7 |
| Botswana | — | 8.4 |
| Brazil | 5.1 | 3.3 |
| Cameroon | 2.6 | 3.0 |
| Chile | 1.6 | 0.4 |
| El Salvador | 1.6 | −0.4 |
| Guatemala | 2.8 | 0.7 |
| Ghana | −1.0 | −1.4 |
| India | 1.4 | 1.9 |
| Kenya | 2.7 | 1.9 |
| Mexico | 2.6 | 2.8 |
| Morocco | 2.5 | 2.3 |
| Nigeria | 4.1 | 0.1 |
| Pakistan | 2.8 | 2.5 |
| Panama | 3.3 | 1.4 |
| Peru | 1.1 | −0.2 |
| Senegal | −0.3 | −0.6 |
| Zimbabwe | 0.7 | 0.7 |

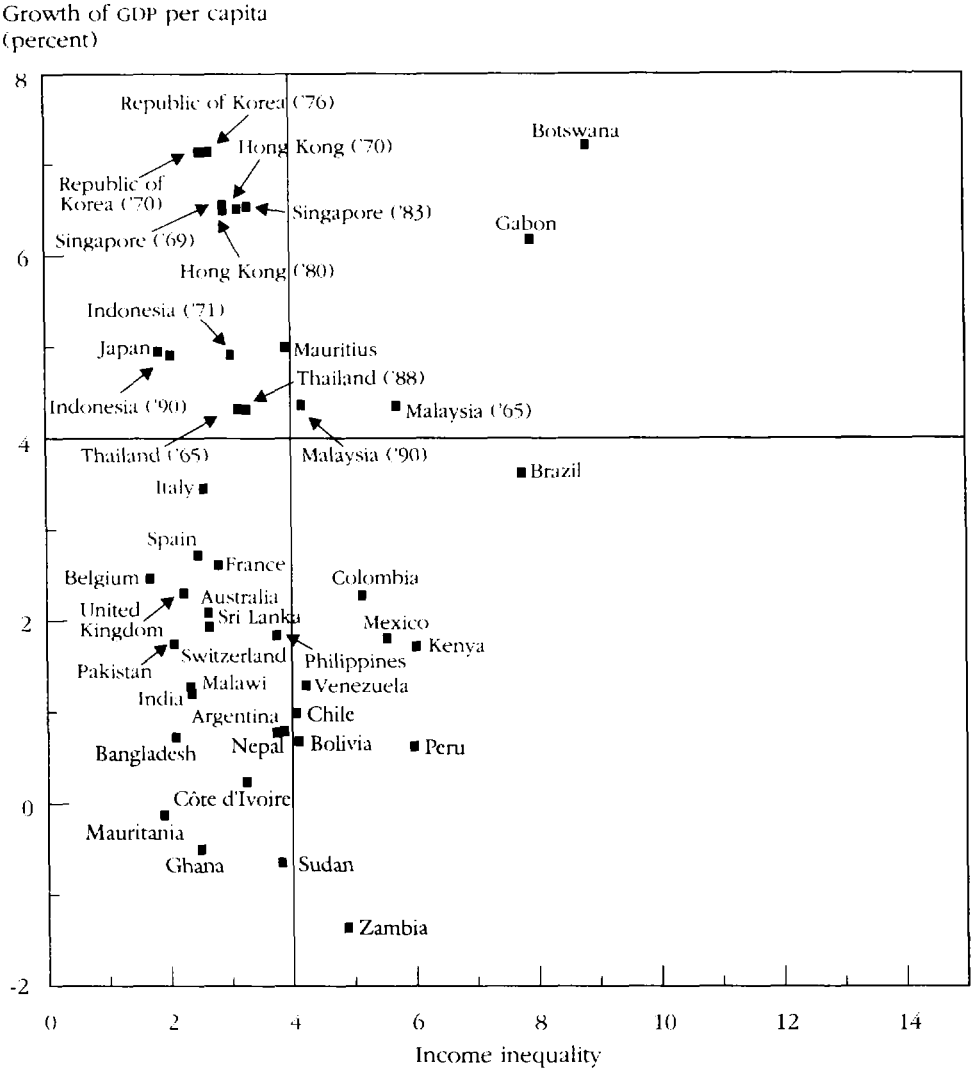
*Source:* World Bank (various years b) and Taiwan, China (1994).

within this period, some East Asian economies show evidence of worsening inequality. See, for example, Bruton 1992 and figure 1.)

Figure 1 shows the relationship across economies between percentage growth in GDP (1965–89) and average income inequality (1965–90) as measured by the ratio of the income share of the richest 20 percent to that of the poorest 40 percent. The East Asian economies stand alone in the northwest corner; they achieved a combination of rapid growth and low inequality over the twenty-five-year period.

In part, the association of rapid growth and low inequality is explained by policies and programs adopted by governments in East Asia that helped ensure widespread sharing in the benefits of economic growth. In addition to the post-war land reforms in Korea and Taiwan (China), a variety of other policies and programs had the effect of ensuring shared growth: public housing in Hong

Figure 1. *Income Inequality and Growth of GDP, 1965–89*



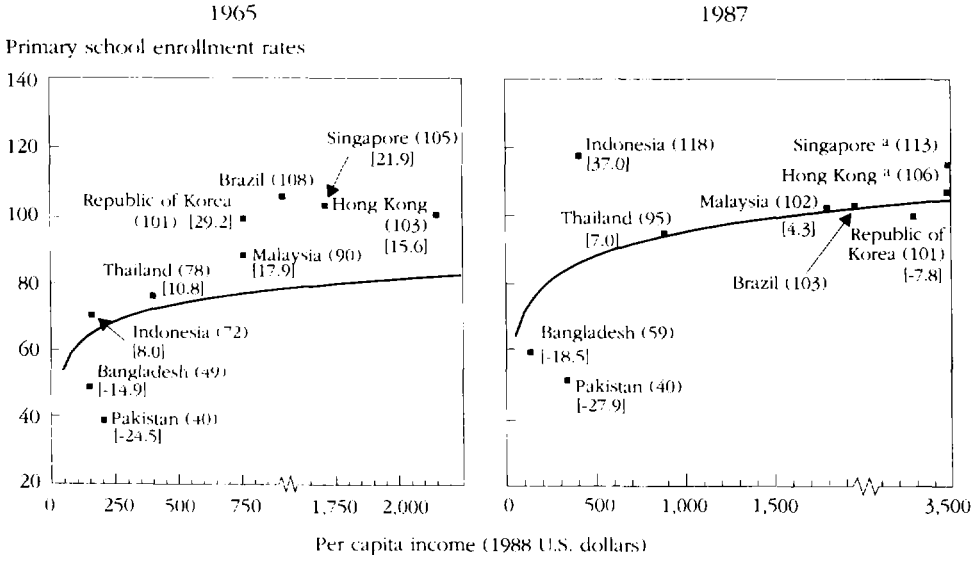

*Note:* Income inequality is measured as the ratio of the income shares of the richest 20 percent to the poorest 40 percent of the population. For the East Asian economies, the change in that ratio is shown using the earliest and the latest year for which that ratio is available. For all other economies, the average of that ratio is taken using all years in the period 1965–89 in which that ratio was available.
*Source:* World Bank (1994).

Kong and Singapore, extensive investment in rural infrastructure in Malaysia and Thailand, and, most common, widespread access to high-quality basic education and health services (World Bank 1993; Campos 1994). Equally important, the export-oriented, labor-demanding development strategy adopted in East Asia increased employment opportunities and wages. These policies and their

successful implementation were the result of some combination of historical circumstances, wisdom, political design, and good luck. With regard to their replicability elsewhere, it may not matter why they were implemented. What matters is that policies for sharing growth seem to have also stimulated growth.

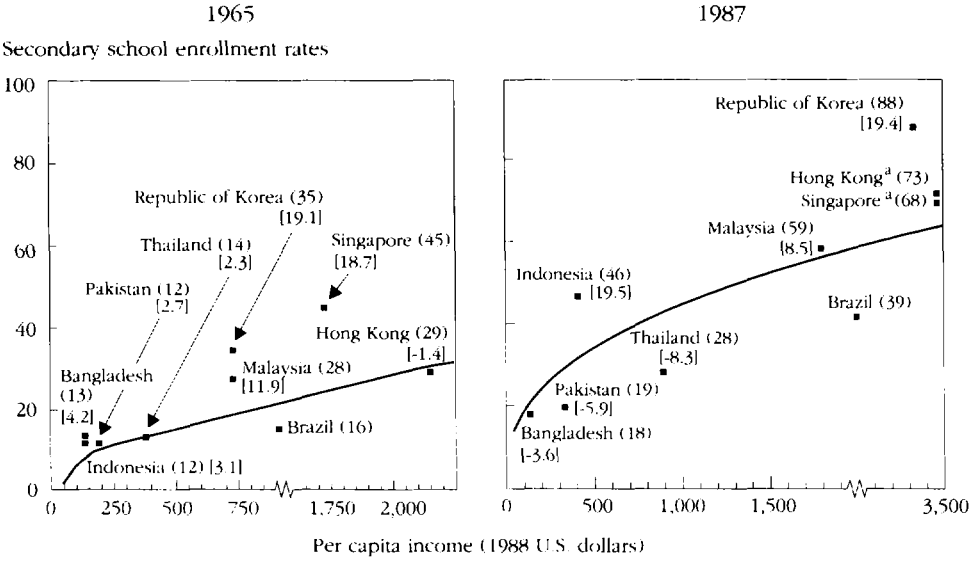
The educational dimension of this shared growth is the most easily measured. Figures 2 and 3 present a stylized summary of the results of regressing both primary and secondary school enrollment rates on per capita national income for more than ninety developing economies for the years 1965 and 1987. The figures indicate that, for the most part, economies in East Asia had significantly higher primary and secondary school enrollment rates than predicted by cross-economy comparisons. Where enrollment rates are low, poor children and girls are the least likely to be enrolled. Where enrollment rates are high, there is a corollary: a reduction in the inequality of access by socioeconomic background and by gender. In contrast to other developing regions, in which increases in the quantity of schooling resulted in the erosion of quality, particularly for the poor, in East Asia increases in quantity were associated with improvements in quality. Repetition and dropout rates—which are negatively correlated with school quality—have remained low in comparison with those in other developing regions, and East Asian children have consistently scored better in internationally com-

Figure 2. *Cross-Economy Regression for Primary School Enrollment Rates, 1965 and 1987*




*Note:* Figures in parentheses are enrollment rates; bracketed numbers show residuals.
a. Per capita income substantially exceeds US$3,500.
*Source:* Singapore and Hong Kong from World Bank (1994); all others from Behrman and Schneider (1992).

Figure 3. *Cross-Economy Regression for Secondary School Enrollment Rates, 1965 and 1987*




*Note:* Figures in parentheses are enrollment rates; bracketed numbers show residuals.
a. Per capita income substantially exceeds US$3,500.
*Source:* Singapore and Hong Kong from World Bank (1994); all others from Behrman and Schneider (1992).

parable tests than children from other developing economies (Birdsall and Sabot 1994).

In the first two sections of this article, we focus on two "virtuous circles" in East Asia. Section I describes a process of cumulative causation, in which education contributed to economic growth, which, in turn, stimulated investment in education. Section II shows how education contributed to low levels of income inequality which, in turn, stimulated investment in education. In section III we discuss how low inequality of income can have a direct, positive effect on economic growth, over and above its effect through education. We present cross-economy regression results that, contrary to the predictions of Kaldor and Kuznets, are consistent with a positive causal effect of low inequality on economic growth and with low inequality as an independent contributing factor to East Asia's rapid growth. We conclude that investment in education is a key to sustained growth, both because it contributes directly through productivity effects and because it reduces income inequality.

Throughout this article we invoke a variety of different results—from theory, microeconomic studies, and cross-economy studies—to piece together and illustrate our story. None of these alone would constitute an airtight case. In particular, the results of cross-economy statistical work must be interpreted with

caution;[1] the statistical difficulties of working with cross-economy regressions—a high level of aggregation, substantial measurement error, inconsistencies in data collection and definition, and specification and omitted-variable bias—are well documented (for example, Levine and Renelt 1992). However, only by looking across economies is it possible to generate the level of variation in key indicators of growth, inequality, and policy needed to capture the effects hypothesized here. The cross-economy regression results reported below complement, and are consistent with, theory and with the microeconometric evidence we present.

## I. Education Stimulates Growth, and Growth Stimulates Education

The accumulation of human capital, as measured by the educational attainment of the population, has consistently emerged as an essential feature of economic growth and development. But the direction of causality implied by the positive correlation between educational attainment and per capita output in a cross-section of economies is unclear: it could simply indicate that education is a luxury consumer good that is increasingly demanded as incomes rise. This concern has been eased by cross-economy regressions in which the characteristics of economies decades ago are used as predictors of subsequent rates of growth. A widely cited contribution to this burgeoning growth-regression literature is Barro (1991). The core regression from that paper, which we use as our base regression, appears in the first column of table 2. (Descriptive statistics for the variables used in table 2 appear in appendix table A-1.) The important contribution of education is among the most robust findings of these growth regressions, proving to be relatively insensitive to changes in either specification or sample composition (Levine and Renelt 1992).

These results are consistent with what human capital theory, the theory of investment in people, predicts. Education augments cognitive and other skills, which, in turn, augment the productivity of labor (Schultz 1961; Becker 1964). Moreover, human capital theory has been extensively tested in the microeconomic literature; Schultz (1988) and Strauss and Thomas (1995) review the evidence from dozens of studies, based on household- and firm-level data, showing that more-educated individuals receive higher wages and are more efficient in managing the health and nutrition of their households.

Endogenous growth theory also predicts that educational investments will enhance growth. A larger stock of human capital facilitates the production of new ideas and technological progress or, for an economy that is not on the

1. Cross-economy comparisons have been a prominent feature of the now fifty-year empirical search for uniformities in the process of the transformation of low-income economies into high-income economies. Clark (1940) and Kuznets (1966) were pioneers in this effort. Chenery and Syrquin (1975) provided a comprehensive description of the structural changes that accompany the growth of developing economies and analyzed their relations. More recently, Romer (1990), Barro (1991), and King and Levine (1993) have used cross-country regressions to test endogenous growth models.

Table 2. *Determinants of GDP Growth in a Cross-Economy Sample, 1960–85*

| *Variable* | *Basic regression* | *Basic regression with manufactured exports* | *Basic regression with income share* | *Basic regression with income share, excluding education variables* |
|---|---|---|---|---|
| Per capita GDP, 1960 (thousands of 1980 dollars) | −0.0075 | −0.0069 | −0.0075 | −0.0020 |
| | (−6.25) | (−5.135) | (−4.730) | (−1.801) |
| Primary school enrollment rate, 1960 | 0.0250 | 0.0271 | 0.0243 | |
| | (4.464) | (4.532) | (3.024) | |
| Secondary school enrollment rate, 1960 | 0.0305 | 0.0262 | 0.0366 | |
| | (3.861) | (1.723) | (2.427) | |
| Ratio of manufactured exports to GDP, 1965 | | 0.0007 | | |
| | | (1.539) | | |
| Secondary school enrollment–export interaction, 1960[a] | | 0.0005 | | |
| | | (0.324) | | |
| Income share ratio[b] | | | −0.0013 | −0.0018 |
| | | | (−1.897) | (−2.406) |
| Government consumption share of GDP, 1970–85[c] | −0.1190 | −0.0566 | −0.1229 | −0.1495 |
| | (−4.250) | (−2.419) | (−4.380) | (−4.760) |
| Average annual number of revolutions, 1960–85 | −0.0195 | −0.0168 | −0.0176 | −0.0268 |
| | (−3.095) | (−1.987) | (−1.867) | (−2.493) |
| Average annual number of assassinations, 1960–85 | −0.0333 | −0.0024 | −0.0055 | −0.0067 |
| | (−2.148) | (−0.738) | (−1.526) | (−1.626) |
| Absolute deviation in investment deflator, 1960[d] | −0.0143 | −0.0139 | −0.0086 | −0.0176 |
| | (−2.698) | (−2.136) | (−1.093) | (−2.012) |
| Constant | 0.0302 | 0.0202 | 0.0418 | 0.0706 |
| $R^2$ | 0.56 | 0.57 | 0.54 | 0.36 |
| Number of observations | 98 | 100 | 74 | 74 |

*Note:* The first column presents Barro's (1991) results; the other columns present new results. *t*-statistics are in parentheses.

a. Ratio of manufactured exports to GDP multiplied by 1960 secondary school enrollment rate.

b. Ratio of the total income of the top 20 percent to the total income of the bottom 40 percent for the first year for which data became available in each country.

c. Average annual ratio of real government consumption (exclusive of defense and education) to real GDP.

d. Magnitude of the deviation of the purchasing power parity value for the investment deflator (U.S. = 1.0) from the sample mean.

*Source:* Barro (1991) and authors' calculations based on data from Barro (1991); World Bank (various years a); Clarke (1995); and World Bank data.

technological frontier, relatively rapid adaptation of new ideas and acquisition of technological capability (Nelson and Phelps 1966; Romer 1990, 1993). Moreover, rates of return to human capital may actually be increasing over some range because of spillover benefits, that is, when one more-educated worker makes an entire group of workers more productive.

The magnitude of the contribution of investment in human capital to rates of economic growth can be assessed with counterfactual simulations derived from the basic regression model in table 2. Consider an economy characterized by the average value of each of the variables in the sample of ninety-eight economies in table 2. We simulated the changes in growth in per capita GDP for that economy during the period from 1960 to 1985 by assuming that the country had achieved primary and secondary school enrollment rates each half a standard deviation above or below the means of those variables for all countries. The one-standard-deviation difference in enrollment rates translates into a nearly 1.5-percentage-point difference in the annual per capita growth rate. The cumulative effect of this annual difference in growth rates over twenty-five years is large. The simulation indicates that a country with primary and secondary school enrollments half a standard deviation above the average in 1960 would have had a GDP per capita 40 percent higher than a country with 1960 enrollments half a standard deviation below the average. Similarly dramatic results emerge from simulations pairing actual countries. With Pakistan's 1960 enrollment rates, Korea's predicted growth rate falls, resulting in a per capita GDP for 1985 that is 39.6 percent lower than the one Korea actually attained.

Allowing the impact on growth of primary school enrollment rates to vary by gender (not shown) reveals no significant difference between the coefficient values for men and women, suggesting that increasing primary school enrollments for girls will be just as effective in stimulating growth as increasing enrollments for boys. This conclusion from cross-economy data is consistent with the microeconomic evidence that the private rates of return to education among wage earners are roughly the same for women as for men (Strauss and Thomas 1995; Summers 1992).

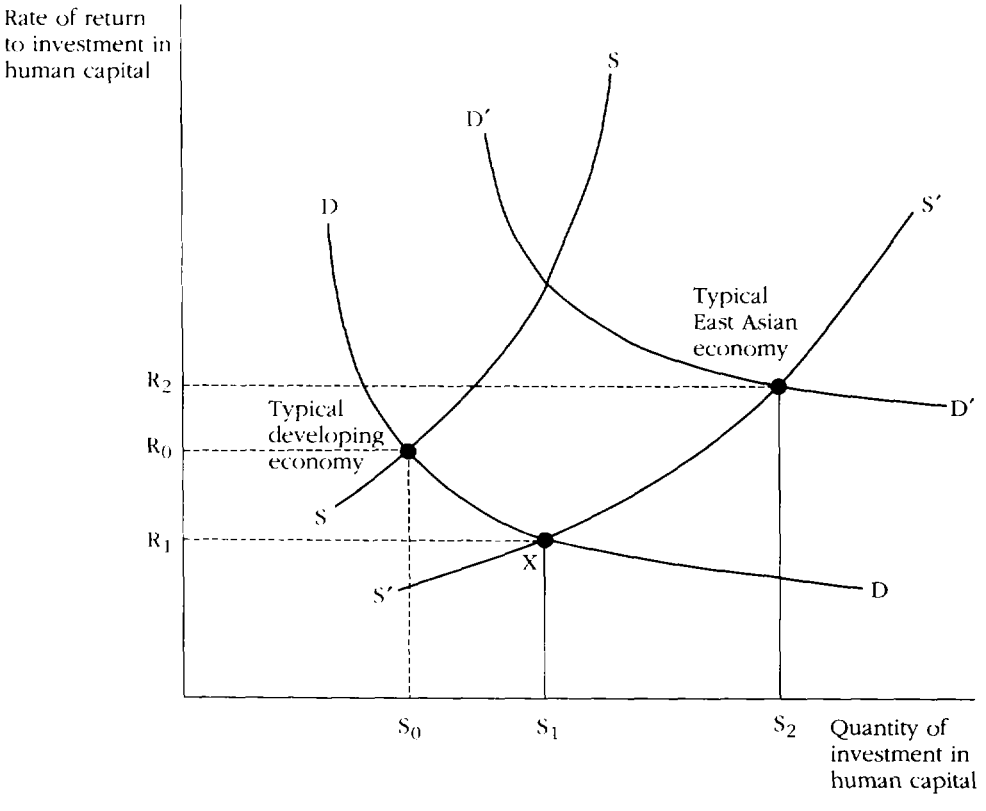
Of course, the rate of participation in the wage-labor market is much lower for women than for men. However, the economic payoff to educating girls is given not only by the increases in the productivity of wage labor but also by nonwage social benefits derived from changes in behavior within households. Fertility is a notable example: the evidence is clear that educated mothers have fewer children (Behrman 1993). Educated mothers also are more efficient users of health services for themselves and their children, send children to school who are better prepared to benefit from society's schooling investment, and are more likely to send their own daughters to school (see Birdsall, Ross, and Sabot 1993). In Asia the fertility rate for women with more than seven years of schooling is 54 percent of the rate for uneducated women (Summers 1992). Closing still another virtuous circle, the fertility decline in East Asia beginning in the mid-1960s resulted in a marked slowing of the growth of the school-age population in the 1970s. This contributed to a growth of public expenditures on basic education per eligible child that was more rapid than elsewhere, permitting more rapid increases in the quantity and improvements in the quality of schooling provided.

Differences among countries in the demand for skills, which have been neglected in regression studies of growth that test education's contribution, are

also important. The resulting omitted-variable bias in these regressions may be a factor in the substantial overprediction of rates of growth for some countries with higher-than-predicted rates of enrollment in primary and secondary schools in the 1960s. For example, countries such as Egypt, the Philippines, Sri Lanka, and the former Soviet Union, like East Asian economies, had greater human capital endowments than predicted for their initial levels of income. Weak demand for educated labor may help explain why these countries nevertheless have tended to underperform with respect to growth.

Figure 4 illustrates the link between the demand for skills in the labor market and education's contribution to growth. The horizontal axis measures the magnitude of investment in human capital, as proxied, for example, by enrollment rates and basic education. The vertical axis measures the rate of return to investment in human capital (schooling) and, implicitly, the contribution of investment to growth for a given level of investment in human capital. *S* and *D*

Figure 4. *Demand Shifts and the Returns to Human Capital, East Asian and Developing Economies*




*Note:* S is the skill supply; D is the demand function; S′and D′represent the same variables for a typical East Asian economy.

are, respectively, the skill supply and demand functions of the typical developing economy; $S'$ and $D'$ are the skill supply and demand functions of the typical East Asian economy. We show $S'$ as shifted to the right because, for example, there is greater public commitment to basic education in East Asia. (The supply function is more elastic because the distribution of income is more equal. There is less absolute poverty than in the typical developing economy, and families near the bottom of the distribution of income can better respond to increases in returns to investments in human capital.) $D'$ is shifted to the right on the premise that, for any given rate of return to skill, skilled workers have been in greater demand in East Asia than in the typical developing economy.

As drawn, in East Asia high levels of demand for skilled workers have offset the tendency for educational expansion to induce diminishing returns to investment in human capital. By contrast, although enrollment rates are higher in some developing economies than others, for example economy $X$, the returns to education will be lower. Educational expansion (from $S_0$ to $S_1$) reduces the rate of return to education (from $R_0$ to $R_1$) because economy $X$ has the human capital supply function of East Asia but the skill demand function of the typical developing economy. In contrast, in East Asia the net result of high demand for human capital as well as high supply is that, although skilled labor is more abundant than in the typical developing economy, the rate of return to investment in human capital (at $R_2$) is at least as high.

It is widely acknowledged that macroeconomic, agricultural, and, in particular, export-push policies contributed to a high demand for labor in East Asia (World Bank 1993).[2] Did these policies, by ensuring adequate demand for skilled as well as unskilled labor, also help ensure a high economic return to education? We assess this question using the regression shown in the second column of table 2. We control for the supply of education and test the relevance of the demand for skills using variables that measure the degree to which an economy is oriented toward manufactured exports.[3] Our prediction is that, in the standard growth rate function, there will be a positive effect on growth of the interaction between the degree of orientation of an economy toward manufactured exports and the level of educational endowments in 1960. That is, the more oriented toward manufactured exports the economy is, the greater the demand for skilled labor and the greater the impact of a given educational endowment on growth will be.

2. East Asia's emphasis on exports contrasted with the import-substituting, capital-intensive strategies that primarily served the elite and labor insiders in Latin America and elsewhere and did not generate a strong demand for labor (Banerji, Campos, and Sabot, forthcoming).

3. See, for example, Pack and Page (1994) and Wood (1994). Of course, a high rate of manufactured exports may also reflect the supply of skills—a function in part of human capital accumulation (Balassa 1984). In the second column of table 2, we control for the supply-side channel by direct inclusion of school enrollment rates. And just as we mitigate simultaneity between growth and human capital accumulation by using 1960 enrollment rates, we mitigate simultaneity between growth and export intensity by using the 1965 ratio of manufactured exports to GDP.

Table 3. *Public Expenditure on Basic Education in Kenya, the Republic of Korea, Mexico, and Pakistan, 1970–89*

| *Country and statistic* | *1970* | *1975* | *1985* | *1989* | *Percentage change, 1970–89* |
|---|---|---|---|---|---|
| *Kenya* | | | | | |
| Expenditure per eligible child (1987 dollars) | 38.6 | — | 46.6 | 53.4 | 38.3 |
| Expenditure as a percentage of GNP | 4.0 | — | 4.9 | 4.9 | 22.5 |
| Absolute expenditure index (1970 = 100)[a] | 100 | — | 220 | 286 | 186 |
| Number of children eligible for basic education (millions)[b] | 3.8 | 4.6 | 7.0 | 7.9 | 107.1 |
| *Korea, Rep. of* | | | | | |
| Expenditure per eligible child (1987 dollars) | 95.3 | 81.6 | 357.1 | 433.4 | 354.7 |
| Expenditure as a percentage of GNP | 3.1 | 1.9 | 3.8 | 2.7 | –12.9 |
| Absolute expenditure index (1970 = 100)[a] | 100 | 91 | 388 | 444 | 344 |
| Number of children eligible for basic education (millions)[b] | 10.1 | 10.7 | 10.4 | 9.8 | –3.0 |
| *Mexico* | | | | | |
| Expenditure per eligible child (1987 dollars) | 68.4 | 124.9 | 113.5 | 111.9 | 63.6 |
| Expenditure as a percentage of GNP | 1.6 | 2.6 | 2.0 | 2.0 | 25.0 |
| Absolute expenditure index (1970 = 100)[a] | 100 | 222 | 255 | 259 | 159 |
| Number of children eligible for basic education (millions)[b] | 16.2 | 19.7 | 24.9 | 25.6 | 58.6 |
| *Pakistan* | | | | | |
| Expenditure per eligible child (1987 dollars) | 7.9 | 9.4 | 13.4 | — | — |
| Expenditure as a percent of GNP | 1.1 | 1.6 | 1.6 | — | — |
| Absolute expenditure index (1970 = 100)[a] | 100 | 150 | 277 | — | — |
| Number of children eligible for basic education (millions)[b] | 21.0 | 26.6 | 34.4 | 42.2 | 101.4 |

— Not available.

a. Absolute expenditures on basic education in real 1987 U.S. dollars used to calculate indexes for absolute expenditures on education.

b. Calculated using enrollment rates and number of students in primary and secondary school.

*Source:* UNESCO (various years) and World Bank (various years c).

Because of substantial multicollinearity, the two variables we add in the second column of table 2—the ratio of manufacturing exports to GDP and the interaction of the secondary school enrollment rate and that measure of exports—are not significant at the 5 percent level. However, joint tests reject the null hypothesis that the coefficients are zero.[4] The results support the contention that the contribution of education to economic growth tends to be greater in countries in which manufactured exports are a higher proportion of GDP. The results are consistent with the view that in East Asia the stimulus that the greater supply of human capital gave to economic growth was augmented by the export orientation of those economies and the resulting skill-demanding growth paths they followed. Our finding is also consistent with the shift of East Asian exporters into more technologically sophisticated and more capital- and skill-intensive goods, as rapidly rising wages of unskilled labor eroded international competitiveness in labor-intensive manufactured goods.

Thus the combination of a greater supply of education and a greater demand for educated workers contributed to faster economic growth in East Asia than in other developing regions.

In the other half of the virtuous circle of education and growth, rapid economic growth and altered household behavior gave positive feedback to greater investment in education in East Asia, and resources for education expanded rapidly. In the ratio of public expenditures on basic education to the number of school-age children, in East Asia rapid economic growth increased the numerator while declining fertility (a result of earlier investments in education) reduced the denominator.

Table 3 indicates that in 1970, public expenditure on basic education per eligible child was not much higher in Korea ($95 in 1987 dollars) than in Mexico ($68). However, from 1970 to 1989 it more than quadrupled in Korea, to $433, whereas in Mexico it did not even double. As a consequence, in 1989, Mexican public expenditure on basic education per eligible child was only 26 percent of Korean expenditure per eligible child.

What accounted for this divergence? It was not government commitment: public expenditure as a percentage of GNP over this period was actually declining in Korea and rising in Mexico. For other East Asian economies as well, there is no evidence that greater government commitment to education produced the extraordinary performance in the provision of education.[5] In 1960 and in 1989, public expenditure on education as a percentage of GNP was not much higher in East Asia than in other regions. In 1960 the share of education in public expen-

4. The *F*-statistic for the restriction that secondary school enrollments, export share, and interaction coefficients are each zero is 6.42, which exceeds the critical value for 1 percent significance with three and ninety degrees of freedom. More important, the corresponding *F*-statistic for the export share and interaction coefficients (6.37) also exceeds the critical value for 1 percent significance.

5. Government expenditure on education, expressed as a percentage of GNP, was not significant in explaining expected years of schooling in a cross-economy regression covering fifteen Asian and Latin American economies. See Tan and Mingat (1992).

diture was 2.5 percent for East Asia, 2.4 percent for Sub-Saharan Africa, and 2.2 percent for all developing economies (Birdsall and Sabot 1994). Over the next three decades, all regions markedly increased the share of national output they invested in formal education, but in 1989, the share in Africa, 4.1 percent, was higher than the East Asian share of 3.7 percent, which barely exceeded the 3.6 percent average share for all developing economies (Birdsall and Sabot 1994).

The initial conditions, the colonial legacy, are also not decisive in explaining why enrollment rates have been so much higher in East Asia than elsewhere. Although Korea had much higher enrollment rates in 1950 than did most other developing economies, the roughly 50- and 70-percentage-point increases since then in, respectively, primary school and secondary school enrollment rates account for much of the current gap between Korea and other middle-income economies. Similar claims can be made for other East Asian economies.

Two factors permitted a faster expansion of educational systems and enrollment rates in East Asia: faster economic growth and a more rapid decline of fertility. In Korea the absolute level of expenditure on basic education rose faster than in Mexico in part because GDP growth was faster. For example, from 1965 to 1975 GDP growth averaged 9.5 percent in Korea and 6.5 percent in Mexico (World Bank, various years b). This implies that over that period, given a constant share of GDP allocated to education, the resources available to the education sector in Korea rose by almost 400 percent; they rose also in Mexico but less, by about 250 percent. (Rapid growth also raises the demand for labor, hence wages, and in particular, the wages of teachers. Because the pay of teachers accounts for a large proportion of recurrent expenditure on education in low-income economies, the tendency for rising costs to reduce the benefits of rapid growth would be strong except for an important mitigating factor. Rapid accumulation of human capital in one period increases the potential supply of teachers in the next, thereby reducing the relative earnings premium that teachers command. Although growth induces increases in average wages, the wages of more-educated workers, including teachers, tend to rise at a slower rate. For example, in part because of low relative salaries, the per-pupil operating costs of primary schools are about 13 percent of GNP per capita in Indonesia and Malaysia, whereas they are almost 30 percent in francophone Africa; Tan and Mingat, 1992.)

Over the two decades from 1970 to 1989 there was also a divergence between Mexico and Korea in the number of school-age children. Although in Mexico the number of children eligible for basic education increased by nearly 60 percent, in Korea the number actually declined by 2 percent (Birdsall and Sabot 1994). The difference in fertility rates, of which these diverging trends are a reflection, was in part caused by earlier differences in educational attainment, in particular, in the educational attainment of women.

Figure 4 illustrates another feedback mechanism. In the absence of the skill intensification of labor demand, educational expansion (from $S_0$ to $S_1$) reduces the rate of return to education (from $R_0$ to $R_1$) in the typical low-income economy.

This lower return then slows further investment in education. Investment in education by households is greater in East Asia because the demand for educated workers is greater, and consequently the returns to investment in schooling are higher. $S_2 - S_1$ is the difference between the East Asian economy and the typical developing economy in the level of investment in human capital induced by the greater demand for educated labor in the East Asian economy. A stronger demand for educated workers elicits a greater supply.

## II. Education Lowers Inequality, Which Stimulates Investment in Education

In a cross section of more than eighty economies, there is a clear correlation between basic-education enrollment rates and lower levels of income inequality (Birdsall and Sabot 1994). There is evidence that causality runs in both directions in this relationship—in a virtuous cycle of greater education as a cause and a consequence of lower inequality.

Because educational expansion increases the number of workers holding high-wage jobs, income inequality can increase rather than decline; the change in the educational composition of the labor force causes a disequalizing effect analogous to that identified by Kuznets. With educational expansion, however, the relative abundance of educated workers increases, and another factor kicks in: the scarcity rents that the educated earn are eroded. The result is a compression of the educational structure of wages. This compression effect can offset the composition effect, leading to a reduction in the inequality of pay and hence in total income inequality (Knight and Sabot 1983, 1991).

Differences between Brazil and Korea in educational opportunities in the 1970s and 1980s illustrate how, if education expands rapidly to all segments of a population, the inequality of pay can be reduced by the dominance of the compression effect. In 1980, Brazil and Korea had similar levels of per capita income. Both countries had achieved universal primary education, and in both countries secondary and tertiary enrollment had increased. However, there was a large and widening gap between Brazil and Korea in the extent to which enrollment rates for secondary and tertiary education grew. In Korea between the mid-1970s and mid-1980s the proportion of high school and postsecondary graduates in a large random sample of the wage-labor force increased so sharply that the proportion of workers with an elementary education or less declined from nearly 20 percent to under 8 percent. By contrast, for the Brazil sample, the comparable proportion fell from over 70 percent to just over 60 percent (Park, Ross, and Sabot, forthcoming).

Table 4 presents the results of estimating a standard Mincerian log wage regression on the Korean and Brazilian samples using 1976 and then 1985 data. In Korea, with the spread of high school and postsecondary education, the extra earnings of graduates at these levels fell. In 1976, workers with a high school education earned 47 percent more than primary school graduates; by 1986 that

Table 4. *Male Wage Structure in Brazil and the Republic of Korea, 1976 and 1985*

| | *Brazil* | | *Korea, Rep. of* | |
|---|---|---|---|---|
| *Variable* | *1976* | *1985* | *1976* | *1986* |
| Premium to primary schooling | 0.488 (55.68) | 0.449 (67.23) | — | — |
| Premium to secondary schooling | 0.958 (85.70) | 0.886 (110.53) | 0.176 (19.66) | 0.092 (7.54) |
| Premium to high school | — | — | 0.473 (48.19) | 0.296 (23.40) |
| Premium to higher education | 1.593 (100.22) | 1.508 (127.40) | 0.969 (71.48) | 0.655 (42.06) |
| Experience | 0.045 (64.97) | 0.048 (83.91) | 0.067 (61.90) | 0.078 (69.61) |
| Experience$^2$ | −0.0006 (61.41) | −0.0007 (79.27) | −0.001 (39.13) | −0.001 (50.27) |
| Constant | 1.149 | 7.043 | 10.231 | 11.779 |
| $R^2$ | 0.546 | 0.562 | 0.532 | 0.449 |
| Number of observations | 85,106 | 118,000 | 23,838 | 24,486 |
| Mean log of wages | 1.864 | 8.095 | 11.363 | 12.895 |

— Not available.

*Note:* Dummy variables were included to control for region, occupation, industry, and head of household (Brazil only); *t*-statistics are in parentheses.

*Source:* Authors' calculations.

premium had declined to 30 percent. Similarly, the extra earnings of workers with higher education fell from 97 percent to 66 percent. In Brazil, where the expansion of enrollment rates in secondary and tertiary education was less marked, the premia to workers with secondary and tertiary schooling barely changed.

In Korea the compression effect shown in table 4 offset the composition effect of increased education; the net effect of educational expansion was to reduce the log variance in wages by 22 percent. By contrast, in Brazil the composition effect dominated; the net effect of the smaller educational expansion over the decade was to increase that measure of wage inequality by 4 percent. By 1985, 25 percent of the gap between Brazil and Korea in the inequality of pay could be explained by differences in the educational attainment of the labor force (Park, Ross, and Sabot, forthcoming). The reduction in the wage premium to educated workers in Korea between the 1970s and 1980s and the lack of any such decline in Brazil suggest that the more rapid spread of education in Korea, and elsewhere in East Asia, contributed over time to the below-average levels of income inequality there compared with those in Latin America.

Closing the second virtuous circle, there has been a feedback effect from low inequality of income to high enrollment rates and thus greater education. Low inequality of income increases household demand for education and probably increases the public supply.

On the demand side, it is likely that budgetary constraints combined with poor access to capital markets mean that poor households do not invest in their

children's human capital even when the returns are high. The pressing need to use income simply to subsist crowds out high-return investments and constrains the demand for education. Table 5 pairs East Asian economies with other economies having similar levels of average per capita income but with considerably higher levels of income inequality and hence lower absolute incomes of the poor. For example, although the per capita income of Brazil (in 1983) slightly exceeded average income in Malaysia (in 1987), the bottom quintile received only 2.4 percent of total income in Brazil compared with 4.6 percent of total income in Malaysia. The per capita income of the bottom quintile in Brazil was thus only 54 percent of the per capita income of the bottom quintile in Malaysia. Given an income elasticity of demand for basic education of 0.50 (a conservative figure), if the distribution of income were as equal in Brazil as in Malaysia, enrollments among poor Brazilian children would be more than 40 percent higher.

Low levels of income inequality may have an influence on the supply side as well as the demand side of the market for education. For the government to provide subsidized basic educational opportunities for a large segment of the school-age population when the distribution of income is highly unequal, the tax burden on the rich has to be heavy. High-income families are likely to resist, for example, by trying to channel public spending for education into subsidies to higher education where their own children will be the beneficiaries (Birdsall and James 1993). If incomes are more equally distributed, the incidence of taxes to finance mass education need not be as concentrated, and resistance to such programs by high-income families is likely to be weaker.

Although public expenditure on education as a share of GNP is not higher in East Asia than in other developing regions, the share of public expenditure on education allocated to primary and secondary, as opposed to higher, education has been consistently larger in East Asia. Korea and Venezuela are extreme examples. In 1985, Korea allocated just 10 percent of its public education budget to higher education, but Venezuela allocated 43 percent (Birdsall and Sabot 1994). As a result, although both countries spent a similar percentage of GDP on education, Korea spent more than twice as much as Venezuela on primary and secondary education, a fact that surely helps to explain Korea's more abundant and higher-quality educational opportunities. Birdsall and Sabot (1994) present similar data for other East Asian economies.

Findings of analyses across economies are also consistent with the likelihood that more equality in the distribution of income leads to more investment in education. In a study of the determinants of secondary school enrollment rates, Williamson (1993) found that more egalitarian societies (measured using the ratio of the share of total income of the bottom 40 percent to the share of the top 20 percent) had higher secondary school enrollment rates. Using his estimated equation to decompose the difference in enrollment rates between Brazil and Korea, we found that none of the 27-percentage-point difference in enrollment rates could be explained by the difference in GNP per adult. Nearly all of that portion of the gap that could be explained was caused by the greater

Table 5. *Absolute Income Share of Lowest Quintile in Selected Economies and Years*

| *Economy* | *GNP per capita (U.S. dollars)* | *Population (millions)* | *Poorest 20 percent of households* | | | |
|---|---|---|---|---|---|---|
| | | | *Total GNP (millions of U.S. dollars)* | *Income share (percent)* | *Absolute income (U.S. dollars)* | *Per capita income (U.S. dollars)* |
| Botswana, 1986 | 840 | 1.1 | 924 | 2.5 | 23 | 115 |
| Brazil, 1983 | 1,880 | 129.7 | 243,836 | 2.4 | 5,852 | 226 |
| Costa Rica, 1986 | 1,480 | 2.6 | 3,848 | 3.3 | 127 | 254 |
| Indonesia, 1976 | 240 | 135.2 | 32,448 | 6.6 | 2,141 | 79 |
| Indonesia, 1987 | 450 | 171.4 | 77,130 | 8.8 | 6,787 | 251 |
| Kenya, 1976 | 240 | 13.8 | 3,312 | 2.6 | 86 | 31 |
| Korea, Rep. of, 1976 | 670 | 36.0 | 24,120 | 5.7 | 1,375 | 191 |
| Malaysia, 1987 | 1,810 | 16.5 | 29,865 | 4.6 | 1,374 | 416 |
| Philippines, 1985 | 580 | 54.7 | 31,726 | 5.5 | 1,745 | 160 |

*Source:* World Bank (various years b).

inequality in the distribution of income in Brazil. Were income distributed as equally in Brazil as in Korea, Korea's secondary school enrollment rate, instead of being 27 percentage points higher, would be only 6 percentage points higher.

## III. Low Inequality of Income Stimulates Growth

Levels of educational attainment above those predicted for economies having similar income levels help explain why, contrary to the conventional wisdom, we observe in East Asia both rapid growth and low levels of income inequality. Investment in education augments growth and reduces inequality and, closing a virtuous circle, rapid growth and low inequality induce higher investment in education. The adoption of a labor-demanding growth strategy that was in accord with factor endowments and comparative advantage also contributed to both rapid growth and low inequality (World Bank 1993; Banerji, Campos, and Sabot, forthcoming). Banerji, Campos, and Sabot (forthcoming) argue that in East Asia the incomes of those in the bottom half of the distribution of income were pulled up by the rapid growth of demand for wage labor. They show that in Korea from 1970 to 1990 wage employment in manufacturing grew at an annual rate of 19 percent and real wages in manufacturing increased at an annual rate of 9 percent. By contrast, in India over the same period both wage employment and real wages in manufacturing grew at less than 2 percent a year.

But does low inequality stimulate growth independently of its effect on education? In this section we first present some cross-economy evidence that it does. We then suggest some reasons why. We point out how public policy in East Asia counteracted the Kaldor and Kuznets effects, thereby weakening the tendency for rapid growth to be associated with high inequality.

### *Econometric Results*

Recent cross-economy studies of growth suggest a negative relationship between income inequality and average annual growth in per capita GDP (Alesina and Rodrik 1994; Persson and Tabellini 1994). For the period 1970–88, Clarke (forthcoming) found this relationship to be robust to the choice of (five) inequality measures and alternative specifications of the explanatory variables. We modified the Clarke regression because we wanted estimates for the period 1960–85 for comparability with our other growth regressions and because in Clarke's data set inequality observations for some countries are as recent as 1980. Current inequality, arguably, is simultaneously determined with growth. We assembled economy-by-economy observations from a variety of sources on the ratio of the income shares of the top 20 percent of population and the bottom 40 percent. We chose the earliest available observation and dropped observations where the measure postdated 1970. This procedure yielded a data set with seventy-four observations.[6]

6. The list of economies and earliest year for the inequality variable is available from the authors.

The addition of the inequality measure to the basic growth rate function reported in column 3 of table 2 does not much change the parameter estimates. The education variables remain significantly positive. The inequality variable is negative and significant (at the 10 percent level). Controlling for other determinants of growth, we find that inequality and growth are inversely related.

But might the negative correlation between initial inequality and growth be spurious? The correlation might reflect regional variation in omitted characteristics, as reflected in figure 1, with HPAEs concentrated in the northwest quadrant and Latin American countries in the southeast (Fishlow, forthcoming). Indeed, adding an HPAE or a Latin America dummy variable makes the inequality variable insignificant. Of course, the existence of regional effects does not in itself mean the relationship between inequality and growth is spurious; the use of dummy variables begs the nature of the heterogeneity they capture. In any event, given such problems as the comparability across economies of measures of inequality (Bruno, Ravallion, and Squire 1995), and the other weaknesses of cross-economy analysis noted above, we can only view the results as suggestive.

How much might low inequality stimulate growth? For our sample of low- and middle-income economies, the average annual growth in per capita GDP from 1960 to 1985 was 1.8 percent. A one-standard-deviation increase in primary school and secondary school enrollment rates raises growth rates by 0.62 and 0.34 percentage points, respectively. A one-standard-deviation decrease in the level of income inequality raises the predicted growth rate by 0.32 of a percentage point. The effect of reducing inequality is substantial; for example, other things being equal, after twenty-five years, GDP per capita would be 8.2 percent higher in an economy with low inequality than in an economy with inequality one standard deviation higher.

Korea's GDP grew at an average annual rate of 5.95 percent from 1960 to 1985 (World Bank 1993). We conducted two sets of simulations of would-be growth in Korea, with the inequality variables and then the enrollment variables varying from their actual values. Using the same equation that was used to calculate the regressions in the third column of table 2, we set the inequality and the enrollment rate variables first to the average for all economies in our sample, then to the average for all low- and middle-income economies, and finally to the level for Brazil, while holding other variables at the level for Korea. The predicted growth rates we thus obtained are presented in table 6.

If, in 1960, Korea had had Brazil's level of inequality, Korea's predicted growth rate over the following twenty-five years would have been reduced by 0.66 percentage points each year, implying that after twenty-five years, GDP per capita in Korea would have been nearly 15 percent lower. In 1960 the percentage gap between Korea's and Brazil's enrollment rates was smaller than the gap in their levels of inequality; as a consequence, the impact on Korea's predicted growth of substituting Brazilian enrollment rates is smaller than the impact of substituting Brazilian inequality.

Table 6. *Actual and Simulated Growth Rates for the Republic of Korea, 1960–85*

| *Simulation values* | *Inequality variable* | *Enrollment variables* |
|---|---|---|
| Actual for Korea | 5.95 | 5.95 |
| Average of all other countries | 5.76 | 5.68 |
| Average of all low- and middle-income countries | 5.67 | 4.77 |
| Brazil | 5.29 | 5.39 |

*Source:* Authors' calculations.

Moreover, low enrollment rates in Brazil probably in part resulted from the constraint on the demand for schooling imposed by high inequality. If in 1960 Korea had had Brazil's inequality, Korea's actual enrollment rates would have been lower, suggesting a still larger total constraint of high inequality on economic growth. Conversely, to the extent that low inequality stimulates growth by inducing increased investment in education, the coefficient on the inequality variable in the third column of table 2 captures only the direct effect of inequality on growth net of the effect of inequality on enrollments. This implies that deleting the education variables from the growth regression should increase the parameter value on the inequality variable. The fourth column of table 2 indicates that this prediction is correct: deleting the education variables raises the parameter value on the distribution variable by roughly 40 percent. In addition to its direct effects, an indirect effect of high inequality is that it appears to constrain growth by constraining investment in education.

### *Why Might Low Inequality Stimulate Growth?*

Using income transfers to reduce income inequality is unlikely to be good for growth for a number of reasons: transfers often result in the diversion of scarce savings from investment to the subsidization of consumption; the targeted group is often not the one to benefit from transfers, reducing the effectiveness of transfers as a means of raising the standard of living and hence the savings and investment rates of the poor; and transfers tend to distort incentives and reduce both allocative efficiency and x-efficiency—x-efficiency is a measure of workers' productivity holding constant all other inputs to the production process, including the workers' skills (Leibenstein 1966). However, policies that increase the productivity and earning capacity of the poor might be good for growth. Low inequality can stimulate growth in four ways: by inducing large increases in the savings and investments of the poor; by contributing to political and macroeconomic stability; by increasing the x-efficiency of low-income workers; and, with higher rural incomes, by increasing market demand for domestic producers.

*Increased savings and investments of the poor.* Low inequality can stimulate growth by inducing large increases in the savings and investments of the poor,

assuming there are not offsetting reductions in the savings of the nonpoor. Liquidity constraints can keep the poor from investing even when expected returns are high, but reducing inequality can increase investment in education by easing such constraints. Reducing inequality can have a similar impact on other dimensions of human capital investment, such as health and nutrition. Although the increased investment in health and nutrition is in part a consequence of education, particularly of mothers, part is independent of the education link. Lower income inequality implies higher absolute incomes for the poor for any given level of average country income per capita. Nutritious food and preventive and curative health care are superior goods and services. Therefore the impact of lower inequality on the consumption by the poor of these goods and services is likely to be larger than the impact of lower inequality on the incomes of the poor. Among the poor, the productivity of labor is adversely affected by inadequate nutrition and ill-health. Greater equality, through its effect on the health and nutrition of the poor, can imply higher-than-proportionate increases in productivity and thus in growth.

We are suggesting that the higher the absolute income of the poor, the smaller the negative impact on their investment rate of the capital market imperfection that prohibits borrowing to finance investment in human capital. This implies that where income inequality is low, the positive association between income and savings rates may not be as strong as Kaldor (1978) presumed or national income data indicate. The data capture only savings channeled through financial intermediaries, not the increased investments in human capital that result from eased liquidity constraints. (There may be a similar effect on other investments of the poor, for example in agriculture or small enterprises, although this effect is more likely to be reflected in measured investment rates.)

Of course, the higher nonfinancial savings rates of the poor where income inequality is low may be offset by lower financial and nonfinancial savings rates of the rich (who are in effect less rich for a given average country income); whether this is so warrants empirical study. However, note that the link of savings to growth depends on the efficiency of the resulting investment; if the savings of the poor are in the form of human capital and other investments they choose and closely monitor, we can expect the marginal returns to be relatively high compared with investments financed through traditionally intermediated savings. It may not be a coincidence that the fastest-growing East Asian economies have had not only unusually high rates of saving and private investment, but also high rates of return to investment compared with those of other developing regions (World Bank 1993).

*Political and macroeconomic stability.* Low inequality can stimulate growth by contributing to political and macroeconomic stability. For example, low inequality can reduce the tendency for fiscal prudence to be sacrificed to political expediency, discourage inappropriate exchange rate valuation, and accelerate the adjustment to macroeconomic shocks. The likelihood that fiscal, monetary,

and other economic policies will swing between the extreme of serving the narrowly defined and myopic interests of the elite and an equally myopic populist extreme is lower where lower levels of income inequality narrow the distance between the interests of the elite and the poor. In contrast, growth is constrained when policies at each extreme create pressures for higher fiscal expenditures that divert scarce public resources from high-return investments to consumption. Moreover, by increasing economic uncertainty, such policy swings, much more common in Latin America and Africa than in East Asia, are likely to discourage private investment.

The allocation of limited fiscal resources for tertiary education, so common in Latin America, is an example of a fiscal policy that reflects pressure for public spending on favored groups. Such spending, from which the children of the elite disproportionately benefit, contributes little to growth (Birdsall and James 1993; Birdsall and Sabot 1994). Because private returns to, and the resulting demand for, tertiary education are high, the private sector would have undoubtedly provided this schooling for the children of high-income families in the absence of government subsidies. The government thus enables the elite to increase their consumption by providing a service at subsidized rates for which they would otherwise be willing and able to pay full cost. For example, despite the fact that Korea spends a lower proportion of public funds on higher education than Venezuela does, enrollment rates for higher education are higher in Korea than in Venezuela (37 percent compared with 26 percent), in part because of higher enrollment in private universities in Korea (United Nations 1994).

At the other extreme, the provision by the government of make-work jobs in the public sector in an attempt to satisfy the excess demand for high-wage employment is an all too common example of populist policies. Because in these circumstances wages in the public sector are generally higher than the marginal product of labor, which may be zero or even negative, this excess employment is heavily subsidized and diverts scarce savings from high-return investments to consumption (Gelb, Knight, and Sabot 1991).

Low inequality can also contribute to macroeconomic stability by strengthening the resolve to avoid exchange rate overvaluation. A myopic, self-serving elite that is not concerned with the welfare of the poor will tend to overvalue the exchange rate (Sachs 1985). Overvaluation will reduce the price of imports (which the urban elite has a high propensity to consume) at the expense of agriculture (in which the poor are concentrated) and other export-oriented sectors. In addition, overvaluation is likely to exacerbate external imbalance, a common cause of macroeconomic instability. East Asian economies have maintained relatively stable exchange rates over the last two to three decades, avoiding the rapid appreciation so common elsewhere. Stable exchange rates eased the task of containing inflation and of limiting internal and external debt to manageable proportions.

The ability to respond quickly to unanticipated shocks may be another link between low levels of income inequality and macroeconomic stability. Lower inequal-

ity is likely to be associated with more policy flexibility in responding to a negative shock, because the negative effects on consumption will be seen as more widely shared and are less likely to erode the absolute incomes of the poor. In an economy where inequality has been declining, the short-run consequence of a negative shock for those in the bottom half of the distribution of income is more likely to be a decline in the rate of growth, rather than in the absolute level, of their incomes. Mazumdar (1993: 361) notes that as part of Korea's adjustment to the first oil shock "the unit cost of labor was reduced by a massive 25 percent," largely accounted for by currency devaluation. He goes on to observe that "even this amount of decline of the real share of labor did not imply a fall in the real wage. Rather, the wage increase in 1975 was held down to 1.4 percent, compared with the annual wage increase in excess of 10 percent from 1966 to 1973." If the burden of a negative shock is shared more equally across income groups, the government will be better able to adjust domestic absorption by reducing consumption while protecting investment. In contrast, an absolute reduction of the incomes of the poor as a means of reducing domestic absorption—more likely if inequality is rising—even if to promote long-term growth, may provoke a strong negative reaction and lead to political and economic disruption.

Low and declining levels of inequality are also likely to reinforce political stability directly. For example, when the incomes of the elite are extremely high or are increasing rapidly while the incomes of non-elites stagnate, the risk is that a large proportion of the population will become politically alienated. Low or declining inequality implies, in contrast, that non-elites are sharing in the benefits of economic growth. This reduces the risk of their political alienation, legitimizes the government in the eyes of the mass of the population, and helps build broadly based political support. A more stable political environment is conducive to economic growth: investment is likely to be higher where the risk of economically disruptive political upheaval and the expropriation of private assets is reduced.

*Increased x-efficiency of low-income workers.* High inequality may reduce growth by reducing the x-efficiency of low-income schoolchildren and low-income workers (Leibenstein 1966). High inequality in developing countries is often associated with real barriers to upward mobility for those at the bottom. Limited access to good schooling and credit and the racial and ethnic discrimination that are sometimes correlated with low income are examples of a nonlevel playing field that can discourage effort. In the competition for scarce places in the university system, children from low-income households learn from experience that no matter how great their effort, it will be insufficient to compensate for the low quality of the schools they attend. By contrast, poor children in high-quality schools who see tangible rewards for effort are more likely to make such an effort.

Similarly, extra effort is unlikely to be forthcoming from low-income workers or farmers who, because of policy biases, face economic arrangements that

do not reward effort. They are more likely to shirk and to become resentful and alienated. For example, in agricultural settings with a highly unequal distribution of land, workers (be they wage workers or sharecroppers) and owners cannot capture the full productivity effects of their activities. Productivity-enhancing behavior such as hard work and risk taking is impossible to specify contractually and is costly to monitor—a form of the principal-agent problem (Bowles and Gintis 1994). Although difficult to quantify, the increases in productivity associated with low inequality may, nevertheless, be large. Timmer (1993), for example, argues that the observed link between agricultural growth and increases in national total factor productivity is the result in part of the increased work effort and investments made by the rural poor in response to improved incentives to the agricultural sector. The work ethic for which East Asian children and labor are well known may be less an exogenous cultural trait than an endogenous response to incentives that reward effort.

Land reform is the most straightforward example of an agricultural policy, implemented in Korea and Taiwan (China), that both reduced inequality and increased productivity. Labor intensity and yield tend to increase as farm size decreases: value added per hectare on small farms (three hectares or less) tends to be three to five times greater than the average for large farms (500 hectares or more) (Squire 1981; Berry and Cline 1979).[7] This implies that the reduction in the inequality of landholding and in the average size of farms that resulted from the land reforms in Korea and Taiwan (China) also increased agricultural output and labor demand. In Indonesia, Korea, and Thailand, both the average size of farms and the Gini coefficient of farm size distribution are much smaller than in Argentina, Brazil, Mexico, and most other countries in Latin America (Squire 1981).

*Increased rural incomes.* Low levels of income inequality can stimulate growth by raising rural incomes, which limits intersectoral income gaps and the rent seeking associated with such gaps and increases the domestic multiplier effects of a given increase in per capita income. Societies with low levels of income inequality are more likely to have policies that contribute to, rather than sap, the dynamism of the agricultural sector. A larger share of public investment was allocated to rural areas in East Asian economies than in other low- and middle-income economies. Equally important, levels of direct and indirect taxation of agriculture were lower in East Asia than in other regions (World Bank 1993). Those societies will tend to have a smaller gap between rural income and urban income. A relatively better-off rural population will in turn generate less pressure for make-work jobs in the high-wage urban public sector. The small rural-urban income gap also implies a relatively weak Kuznets effect; differences in the

7. Benjamin (1995), however, considers the possibility that the relationship between productivity and labor intensity and farm size is a statistical artifact resulting from omitted land-quality variables. He provides supporting evidence from Indonesia.

productivity and incomes of labor in the low- and high-productivity sectors are smaller, so the intersectoral transfer of labor generates less income inequality.

Low levels of income inequality, associated with higher income in the rural and agricultural sectors, also mean higher demand for the agricultural inputs and consumer goods that can then stimulate the growth of labor-intensive nonagricultural output. In Taiwan (China), for example, in the 1950s and early 1960s, agriculture—not manufacturing for export—was clearly the leading sector, and roughly 60 percent of the increment to aggregate demand was domestic (Kuo, Ranis, and Fei 1981). More generally, among Asian economies there is a strong positive correlation between the rate of growth of the agricultural sector and the rate of growth of the nonagricultural sector (Mellor 1990). The relationship suggests that the multiplier effects of agricultural growth on manufacturing, construction, and services are large. A 1 percent increase in agricultural growth is associated with a 1.5 percent increase in the growth rate of the nonagricultural sector, implying that the faster agriculture grows, the faster its share of total output declines. With the exception of the East Asian city-states, those Asian economies with the fastest rates of growth of agricultural output over the last thirty years have tended to experience the biggest declines in the share of agricultural output in GNP. Ironically, a dynamic agricultural sector was on their path to industrialization.

Because the relatively simple manufactured inputs and consumer goods demanded by rural residents are generally more efficiently produced with labor-intensive techniques, the employment effects of these increases in demand are amplified (Ranis and Stewart 1987). Detailed study of these backward and forward linkages in the Muda River region of Malaysia provides microeconomic confirmation of the magnitude of this intersectoral multiplier (Bell, Hazell, and Slade 1982). By contrast, when the incomes of the urban elite increase, the tradables on which they spend their increased income tend to be capital-intensive goods.

It is not necessary to be a structuralist to believe that these demand and multiplier effects matter. Even in fully open economies, the multiplier effects of external demand, especially at early stages of development, may not be as great as the effects of local demand. Moreover, vibrant local markets and competition may be critical to producers' developing the expertise and efficiency required to enter external markets (Porter 1990). Strong domestic demand gave East Asia's early manufacturers a competitive advantage in international markets by giving them the opportunity to test-market labor-intensive goods and achieve economies of scale.

## IV. Conclusion

In East Asia, the system of basic education expanded rapidly in the postwar period and improved in quality. By increasing labor productivity, changing household behavior, and facilitating the acquisition of technological capability, the

rise in education levels stimulated growth. The positive effect of rising education on growth was reinforced by macroeconomic and trade policies, including an emphasis on manufactured exports, that generated demand for labor and increasingly for skilled labor. In a virtuous circle, rapid economic growth in turn led to a higher demand for, and supply of, education. At the same time, the growing abundance of educated labor eroded the scarcity rents of the more educated; thus, broadly based educational expansion reinforced low initial levels of inequality. Closing a second virtuous circle, lower levels of income inequality further increased the supply of, and demand for, education.

Educational performance alone does not explain the extraordinary combination of rapid growth and low inequality observed in East Asia. We have suggested that such other dimensions of the East Asian development strategy as promoting a dynamic agricultural sector and a labor-demanding, export-oriented growth path also stimulated growth and reduced inequality. Finally, we have suggested that low inequality may not only stimulate growth indirectly, by augmenting investment in education, but may also have a direct, positive effect on the growth rate. The direct effect would come from increasing investment in dimensions of human capital other than education, from increasing political and macroeconomic stability, and from reducing intersectoral income gaps and increasing the domestic multiplier effects of a given increase in income. Econometric evidence, crude though the data and methods may be, lends support to the possibility of this direct effect of lower inequality on growth. Other things being equal, economies with lower inequality near the start of the period 1960–85 grew faster. Moreover, our results suggest that the stimulus to growth of low inequality may be quite large.

In summary, the East Asian experience appears sufficient to reject the conventional wisdom of a necessary link between high levels of income inequality and rapid growth. Although our analysis has not been sufficient to confirm the opposite (and might be vulnerable to omitted-variable bias), we hope others now seriously consider the hypothesis that low inequality and well-designed inequality-reducing policies stimulate growth. This is not an argument for income transfers, which tend to be growth-constraining, but for policies that reduce inequality by eliminating consumption subsidies for the rich and by increasing the productivity of the poor. Whether by political design, luck, or foresight, East Asian leaders were successful in identifying and implementing such policies. It does not matter now what their motives were in doing so; what matters is that leaders in other regions now benefit from the lessons of the East Asian success.

*(Appendix table starts on the following page.)*

Table A-1. *Means and Standard Deviations of Variables Used in the Growth Analysis for the Cross-Economy Sample, 1960–85*

| Variable | Basic regression | | Basic regression with manufactured exports | | Basic regression with income share | | Basic regression with income share, excluding education variables | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Standard deviation | Mean | Standard deviation | Mean | Standard deviation | Mean | Standard deviation |
| GDP growth, 1960–85 | 0.02 | 0.019 | 0.0209 | 0.019 | 0.0230 | 0.018 | 0.0230 | 0.018 |
| Per capita GDP, 1960 (thousands of 1980 dollars) | 1.92 | 1.81 | 1.8104 | 1.745 | 2.0173 | 1.771 | 2.0173 | 1.771 |
| Primary school enrollment rate, 1960 | 0.78 | 0.31 | 0.7489 | 0.337 | 0.8254 | 0.288 | | |
| Secondary school enrollment rate, 1960 | 0.23 | 0.21 | 0.2078 | 0.207 | 0.2565 | 0.220 | | |
| Ratio of manufactured exports to GDP, 1965 | | | 4.2022 | 6.836 | | | | |
| Income share ratio[a] | | | | | 4.2650 | 2.458 | 4.2650 | 2.458 |
| Government consumption share of GDP, 1970–85[b] | 0.107 | 0.053 | 0.1776 | 0.064 | 0.1757 | 0.062 | 0.1757 | 0.062 |
| Average annual number of revolutions, 1960–85 | 0.18 | 0.23 | 0.1520 | 0.189 | 0.1466 | 0.191 | 0.1466 | 0.191 |
| Average annual number of assassinations, 1960–85 | 0.031 | 0.086 | 0.2164 | 0.458 | 0.2430 | 0.478 | 0.2430 | 0.478 |
| Absolute deviation in investment deflator, 1960[c] | 0.23 | 0.25 | 0.2065 | 0.223 | 0.2192 | 0.210 | 0.2192 | 0.210 |
| Number of observations | 98 | 98 | 100 | 100 | 74 | 74 | 74 | 74 |

*Note:* The first and second columns present Barro's (1991) results; the other columns present new results.

a. Ratio of the total income of the top 20 percent to the total income of the bottom 40 percent for the first year for which data became available for each country before 1970.

b. Average annual ratio of real government consumption (exclusive of defense and education) to real GDP.

c. Magnitude of the deviation of the purchasing power parity value for the investment deflator (U.S. = 1.0) from the sample mean.

*Source:* Barro (1991) and authors' calculations based on data from Barro (1991); World Bank (various years a); Clarke (1995); and World Bank data.

## References

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

Ahluwalia, Montek S. 1974. "Income Inequality: Some Dimensions of the Problem." In Hollis Chenery, ed., *Redistribution with Growth: An Approach to Policy.* New York: Oxford University Press.

Alesina, Alberto, and Roberto Perotti. 1994. "The Political Economy of Growth: A Critical Survey of the Recent Literature." *The World Bank Economic Review* 8(3):351–71.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics* 109(2):465–90.

Anand, Sudhir, and Ravi Kanbur. 1993. "Inequality and Development: A Critique." *Journal of Development Economics* 41(June):19–43.

Balassa, Bela. 1984. "A 'Stages Approach' to Comparative Advantage." *Asian Journal of Economics and Social Studies* 3(3):1–35.

Banerji, Arup, Jose Edgardo Campos, and Richard Sabot. Forthcoming. "The Political Economy of Formal Sector Pay and Employment in Developing Countries." In J. Edward Taylor, ed., *Labor Markets in Developing Countries.* Paris: Organisation for Economic Co-operation and Development (OECD).

Barro, Robert J. 1991. "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106(May):407–43.

Barro, Robert J., and Jong Wha-Lee. 1993. "International Comparisons of Educational Attainment." *Journal of Monetary Economics* 32(3):363–94.

Becker, Gary S. 1964. *Human Capital. A Theoretical and Empirical Analysis.* Princeton, N.J.: Princeton University Press.

Behrman, Jere. 1993. "Investing in Human Resources." In Inter-American Development Bank, *Economic and Social Progress in Latin America: 1993 Report.* Baltimore: Johns Hopkins University Press.

Behrman, Jere, and Ryan Schneider. 1994. "An International Perspective on Schooling Investments in the Last Quarter Century in Some Fast-Growing Eastern and Southeastern Countries." *Asian Development Review* 12(2):1–50.

Bell, Clive, Peter Hazell, and Roger Slade. 1982. *Project Evaluation in Regional Perspective.* Baltimore: Johns Hopkins University Press.

Benjamin, Dwayne. 1995. "Can Unobserved Land Quality Explain the Inverse Productivity Relationship?" *Journal of Development Economics* 46(February):51–84.

Berry, Albert R., and William R. Cline. 1979. *Agrarian Structure and Productivity in Developing Countries.* Baltimore: Johns Hopkins University Press.

Birdsall, Nancy, and Estelle James. 1993. "Efficiency and Equity in Social Spending: How and Why Governments Misbehave." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor.* Oxford: Oxford University Press.

Birdsall, Nancy, David Ross, and Richard Sabot. 1993. "Underinvestment in Education: How Much Growth Has Pakistan Foregone?" *Pakistan Development Review* 32(Winter):453–99.

Birdsall, Nancy, and Richard H. Sabot. 1994. "Virtuous Circles: Human Capital Growth and Equity in East Asia." World Bank, Policy Research Department, Washington, D.C. Processed.

Bowles, Samuel, and Herbert Gintis. 1994. "Efficient Redistribution: New Rules for Markets, Communities, and States." Discussion Paper. University of Massachusetts, Department of Economics, Amherst. Processed.

Bruno, Michael, Martin Ravallion, and Lyn Squire. 1995. "Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues." Paper prepared for the International Monetary Fund Conference on Income Distribution and Sustainable Growth, Washington, D.C. June 1–2. IMF, Washington, D.C.

Bruton, Henry J. 1992. *The Political Economy of Poverty, Equity, and Growth: Sri Lanka and Malaysia.* Oxford: Oxford University Press.

Campos, J. Edgardo. 1994. "Institutions, Leadership, and the East Asian Miracle: Making the Promise of Shared Growth Credible." World Bank, Policy Research Department, Washington, D.C. Processed.

Chenery, Hollis, and Moises Syrquin. 1975. *Patterns of Development, 1950–1970.* New York: Oxford University Press.

Clark, Colin. 1940. *The Conditions of Economic Progress.* London: Macmillan Press.

Clarke, George. Forthcoming. "More Evidence on Income Distribution and Growth." *Journal of Development Economics.*

Easterly, William, Michael Kremer, Lant Pritchett, and Lawrence H. Summers. 1993. "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks." *Journal of Monetary Economics* 32(3):459–83.

Fields, Gary. 1989. "Changes in Poverty and Inequality in Developing Countries." *The World Bank Research Observer* 4(2):167–85.

Findlay, Ronald, and Stanislaw Wellisz. 1993. *Five Small Open Economies.* Oxford: Oxford University Press.

Fishlow, Albert. Forthcoming. "Inequality, Poverty, and Growth: Where Do We Stand?" In *Proceedings of the World Bank Annual Conference on Development Economics 1995.* Washington, D.C.: World Bank.

Gelb, Alan, John Knight, and Richard Sabot. 1991. "Public Sector Employment, Rent-Seeking and Economic Growth." *Economic Journal* 101(408):1186–99.

Kaldor, Nicholas. 1978. "Capital Accumulation and Economic Growth." In Nicholas Kaldor, ed., *Further Essays on Economic Theory.* New York: Holmes and Meier Publishers, Inc.

King, Robert G., and Ross Levine. 1993. "Finance and Growth: Schumpeter Might Be Right." *Quarterly Journal of Economics* 108(3):717–37.

Knight, John, and Richard Sabot. 1983. "Educational Expansion and the Kuznets Effect." *American Economic Review* 73(5):1132–36.

——. 1991. *Education, Productivity, and Inequality.* New York: Oxford University.

Kuo, Shirley W.Y., Gustav Ranis, and John C. H. Fei. 1981. *The Taiwan Success Story: Rapid Growth with Improved Income Distribution in the Republic of China.* Boulder, Colo.: Westview Press.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45 (March):1–28.

———. 1966. *Modern Economic Growth.* New Haven, Conn.: Yale University Press.

Leibenstein, Harvey. 1966. "Allocative Efficiency vs. 'X-Efficiency.'" *American Economic Review* 56(June):392–415.

Leipziger, Danny M., David Dollar, Su-Yong Song, and Anthony F. Shorrocks. 1992.

*The Distribution of Income and Wealth in Korea.* EDI Development Study. Washington, D.C.: World Bank.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(September):942–63.

Maddison. 1982. *Phases of Capitalist Development.* New York: Oxford University Press.

Mazumdar, Dipak. 1993. "Labor Markets and Adjustment in Open Asian Economies: The Republic of Korea and Malaysia." *The World Bank Economic Review* 7(3):349–80.

Mellor, John. 1990. "Agriculture on the Road to Industrialization." Carl K. Eicher and John M. Staaz, eds. 2d ed. *Agricultural Development in the Third World.* Baltimore: Johns Hopkins University Press.

Nelson, Richard, and Edmund Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review Papers and Proceedings* 56:69–75.

Pack, Howard, and John M. Page. 1994. "Accumulation, Exports, and Growth in the High-Performing Asian Economies." *Carnegie-Rochester Conference Series on Public Policy* 40(June):199–250.

Park, Young-Bum, David Ross, and Richard Sabot. Forthcoming. "Educational Expansion and the Inequality of Pay in Brazil and Korea." In Nancy Birdsall and Richard Sabot, eds., *Opportunity Foregone: Education, Growth, and Inequality in Brazil.* Washington, D.C.: Inter-American Development Bank.

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *American Economic Review* 84(3):600–21.

Porter, Michael E. 1990. *The Competitive Advantage of Nations.* New York: Free Press.

Ranis, Gustav, and Frances Stewart. 1987. "Rural Linkages in the Philippines and Taiwan." In Frances Stewart, ed., *Macro-Policies for Appropriate Technology in Developing Countries.* Boulder, Colo.: Westview Press.

Robinson, Sherman. 1976. "A Note on the U-Hypothesis." *American Economic Review* 66(3):437–40.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(October):S71–S102.

———. 1993. "Two Strategies for Economic Development: Using Ideas and Producing Ideas." In *Proceedings of the World Bank Annual Conference on Development Economics 1992.* Washington, D.C.: World Bank.

Sachs, Jeffrey. 1985. "External Debt and Macroeconomic Performance in Latin America and East Asia." In William C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity.* Washington, D.C.: Brookings Institution.

Schultz, T. Paul. 1988. "Education Investment and Returns." In Hollis B. Chenery and T. N. Srinivasan, eds., *Handbook of Development Economics.* Amsterdam: North Holland.

Schultz, T. W. 1961. "Investment in Human Capital." *American Economic Review* 51(1):1–17.

Squire, Lyn. 1981. *Employment Policy in Developing Countries: A Survey of Issues and Evidence.* New York: Oxford University Press.

Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In T. N. Srinivasan and Jere Behrman, eds., *Handbook of Development Economics.* Vol. 3. Amsterdam: North-Holland.

Summers, Lawrence H. 1992. "Investing in All the People—Educating Women in Developing Countries." *Pakistan Development Review* 31(4):367–93.

Taiwan (China). 1994. *Statistical Abstract of National Income 1951–1993.* Taipei: Bureau of Statistics.

Tan, Jee-Peng, and Alain Mingat. 1992. *Education in Asia: A Comparative Study of Cost and Financing.* Washington, D.C.: World Bank.

Timmer, Peter. 1993. "Why Markets and Politics Undervalue the Role of Agriculture in Economic Development." Hibbard Memorial Lecture Series. University of Wisconsin, Department of Agricultural Economics, Madison. Processed.

UNESCO (United Nations Educational, Scientific, and Cultural Organization). Various years. *Statistical Yearbook.* New York.

United Nations. 1994. *Human Development Report.* New York: Oxford University Press.

Williamson, Jeffrey. 1993. "Human Capital Deepening, Inequality, and Demographic Events along the Asia-Pacific Rim." In Naohiro Ogawa, Gavin W. Jones, and Jeffrey Williamson, eds., *Human Resources in Development along the Asia-Pacific Rim.* Singapore: Oxford University Press.

Wood, Adrian. 1994. *North-South Trade, Employment and Inequality.* Oxford: Clarendon Press.

World Bank. 1993. *The East Asian Miracle.* New York: Oxford University Press.

———. 1994. *Social Indicators of Development.* New York: Oxford University Press.

———. Various years a. *IEC Foreign Trade Statistics.* Washington, D.C.

———. Various years b. *World Development Report.* New York: Oxford University Press.

———. Various years c. *World Tables.* Washington, D.C.

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3: 509-534

# Another Look at the East Asian Miracle

Gustav Ranis

*This article examines the main ingredients of the East Asian success story in each subphase of its post-war transition growth process. It focuses mainly on Taiwan (China), with comparative side glances at the Republic of Korea. The initial conditions facing the region were favorable, especially in Taiwan (China). The early import substitution subphase was unusually mild and contributed to the emergence of strong linkages between agricultural and nonagricultural activities. Subsequently, flexible labor markets, human capital policies, and major macro and structural policy reforms culminated in an increasingly export-oriented industrialization effort. Once labor surpluses were ultimately exhausted, government policies accommodated Taiwan (China)'s entry into the technological era of the 1970s and 1980s. During the past four decades of a signally successful transition growth effort, public policies in East Asia have consistently accommodated the changing needs of the economy, rather than guide its path directly.*

The key elements of East Asia's development success story have become shrouded in a controversy between supporters of the "Washington Consensus," generated by the World Bank and the International Monetary Fund (IMF), and its heterodox critics.[1] The controversy focuses almost exclusively on the importance of recent government industrial policy, mainly in the form of directed credit. The Bank has offered olive branches in the form of "market friendly interventions" followed by "reversible nonmarket contests" (World Bank 1991, 1993). These have not been accepted by either the heterodox critics or the Japanese government, and the argument continues to rage.

The main purpose of this article is to help move the discourse away from such ahistorical and relatively narrow industry-focused "markets versus government" arguments. Instead, it examines what I believe to have been the critical elements of the overall success of development in East Asia, especially Taiwan (China), over the longer run. Not surprisingly, there is no single key ingredient. Instead, what is critical and persuasive is the continuing flexibility of policymakers who have responded on a broad front to the changing demands of the economy dur-

1. The Washington Consensus is summarized in Williamson (1990). For examples of the heterodox critics, see Amsden (1989) and Wade (1990).

Gustav Ranis is with the Department of Economics at Yale University. An early version of this article was prepared as a background paper for the World Bank's *The East Asian Miracle* (World Bank 1993). The author gratefully acknowledges the research assistance of Ryan Schneider and the comments of three anonymous referees.

ing each of the identifiable subphases of the transition growth effort over the past four decades. As a result of this flexibility, the system was able to avoid running out of steam and to be in a position to shift onto new rails at the end of each subphase.

In the 1950s, East Asia, like virtually all developing regions, initially chose an import substitution package, but it was an exceptionally mild and brief version that was less destructive of agricultural incentives in comparison with versions adopted in other regions. Consequently, once the domestic market for labor-intensive nondurables had become exhausted in the early 1960s, it was much easier to overcome the resistance of vested interests to macro and structural policy changes aimed at facilitating an increasingly outward-oriented, labor-intensive industrialization effort. The next major challenge, the gradual exhaustion of that labor by the early 1970s, once again met an anticipatory response, this time in the form of public policies focused on human capital and technology-oriented investments. These policies enabled East Asia to flexibly overcome the external shocks of the 1970s and to participate as an ever more equal partner in the continuously changing product cycle of the industrial countries. Ultimately, the economies of East Asia themselves approached mature-economy status.

Each decade had its challenges, and each decade produced government policy responses that turned out, in Kuznets's terms, to have been accommodating rather than obstructive of changes required by mainly private agents. If there is one key to developmental success, it is avoiding the encrustation of ideas. This has been achieved by continuing policy reforms aimed at an ever-increasing reliance on the responsiveness of large numbers of dispersed decisionmakers.

Section I briefly describes the distinguishing initial conditions that set off East Asia's development efforts. Section II analyzes the unusually mild version of the import substitution subphase (1953–63) that permitted the system to avoid the neglect of its agricultural sector that usually occurs during such a phase. Section III highlights the classical behavior of Taiwan (China)'s dualistic labor market, which was critical for the generation of an initially vigorous domestic, subsequently export-oriented, process of balanced growth (1963–72). Section IV analyzes the importance of flexible policies with respect to the formation of human capital and technological infrastructure in the 1980s and 1990s. Section V presents conclusions and briefly explores the elements of this experience that may be relevant to other economies.

## I. Initial Conditions

A developing economy's policy space and its performance over time are constrained by the initial conditions bestowed upon it by nature and the economy's colonial past. Differences in these conditions yield important typological distinctions among developing countries and are instructive for analytical as well as comparative purposes (for more on this theme, see Ranis 1984).

Both Taiwan (China) and the Republic of Korea were relatively small societies with labor surpluses and poor natural resources in the early 1950s. Both had been part of the Japanese colonial network for which they produced mainly agricultural products (for example, 95 percent of Taiwan (China)'s sugar production and 52 percent of its rice production were exported to Japan [Wu 1972]). The colonial administration expended substantial resources on rural infrastructure (roads, drainage, irrigation, and power) as well as on institutional infrastructure (agricultural research, experiment stations, and farmers' associations). Those investments led to a significant, early "green revolution" in rice (based on the ponlai variety) yields that were raised by 4 percent a year between 1921 and 1937. By the early 1950s the multiple cropping index stood already at 137, and 33 percent of all households had electricity.

Taiwan (China) had also benefited from colonial policies ensuring unusually high literacy rates; in the early 1950s, 60 percent of the population was literate (for a full account of the colonial period in Taiwan (China), see Ho 1978). A colonial (1905) land reform was early on reinforced by another major reform (1949–53), and a similar substantial redistribution of land took place in Korea (1949) under U.S. prodding (see Kuo, Ranis, and Fei 1981 and United Nations, various years a).

The Japanese maintained ownership over large-scale industry. Taiwanese were not admitted into higher education, but the priority given to primary schooling encouraged their expanding participation in small and medium-size industrial and service activities. The already-respectable initial stock of human capital of independent Taiwan (China) was, moreover, substantially enhanced by the massive influx of small traders and entrepreneurs from China in 1949. Also at that time, some very large-scale activities that had been taken over from the Japanese, for example, mining, cement, and pulp and paper, landed in the public sector. Many of these activities were soon divested in connection with the land reform program; the continuing shift to predominantly private industrial activity on Taiwan (China) is illustrated in table 1. In Korea most large-scale industry had originally been located in the North; the South was preponderantly agricultural, with less public sector industrial activity at the outset and no evidence of an increase over time (tables 1 and 2).

Table 1. *Private and Public Sector Industrial Output in Taiwan (China), 1952–87, and Korea, 1963–72*
(percent)

| *Taiwan (China)* | | | *Korea* | | |
|---|---|---|---|---|---|
| *Years* | *Private* | *Public* | *Year* | *Private* | *Public* |
| 1952–61 | 48.7 | 51.3 | 1963 | 88.4 | 11.6 |
| 1962–71 | 64.2 | 35.8 | 1964 | 87.4 | 12.6 |
| 1972–81 | 80.8 | 19.2 | 1971 | 86.2 | 13.8 |
| 1982–87 | 83.8 | 16.2 | 1972 | 86.1 | 13.9 |

*Source:* Taiwan, China (various years a); for Korea, Jones (1975).

Table 2. *Average Annual Growth in Agricultural and Nonagricultural Output in Taiwan (China) and Korea, 1952–89*
(percent)

| Years | Agricultural output | | Nonagricultural output | |
|---|---|---|---|---|
| | Taiwan (China) | Korea | Taiwan (China) | Korea |
| 1952–60 | 4.2 | 2.2 | 11.7 | 9.7 |
| 1960–70 | 4.3 | 4.9 | 18.5 | 17.6 |
| 1970–80 | 2.6 | 3.4 | 12.9 | 19.2 |
| 1980–89 | –0.8 | 0.7 | 9.8 | 6.4 |

*Source:* Taiwan, China (various years a); for Korea, Das (1990).

Thus Taiwan (China) inherited relatively favorable initial conditions; so did Korea, although it had invested less in agricultural infrastructure during the colonial period. Moreover, in both countries there were relatively few landed or industrial vested interest groups to overcome. These factors, plus the perceived threat from China and the Democratic People's Republic of Korea, to Taiwan (China) and the Republic of Korea, respectively, gave these relatively autocratic governments unusual freedom of action in the sphere of economic policy.

## II. The Import Substitution Subphase, 1953–63

The import substitution syndrome as it played itself out in Taiwan (China) had some unusual features: government policies were not unduly discriminatory of the agricultural sector, encouraged rural nonagricultural activities as part of a balanced growth process, avoided intervention in labor markets, built on the foundations of substantial pre-existing human and physical overheads, and ensured early macroeconomic stabilization.

### *The Agricultural Sector*

A dynamic agricultural sector is not frequently encountered during the typical initial import substitution subphase in developing countries. Yet if the concept of an early "leading sector" has any validity, here it clearly was agriculture, especially in the case of Taiwan (China). In 1952 the agricultural sector generated 33 percent of the net domestic product, 60 percent of total employment, and 90 percent of still relatively modest exports. By 1988 these figures had declined to 6, 14, and 6 percent, respectively. Nevertheless, agriculture played a critical historical role, not only as a provider of savings, foreign exchange, and labor but also as a contributor to the dynamic nonagricultural sector in rural areas (Taiwan, China, various years a).

Agricultural output in Taiwan (China) had regained pre-war levels by 1951 and increased at roughly 4.2 percent a year during the 1950s (table 2). The output of traditional crops rose by about 50 percent and that of nontraditional crops, especially cotton, fruits, and vegetables, approximately doubled. Even

more significant, agricultural employment increased by 12 percent during the decade, worker-days deployed rose by 17 percent, and the multiple cropping index increased to 188 by 1964. Nontraditional crops were much more labor-intensive and more subject to changes in labor-using technology over time. Working days per hectare increased from 170 just after World War II to 260 in the early 1960s (Ho 1978).

Forty-five percent of the growth of agriculture during the 1950s can be attributed to a change in total factor productivity, largely a result of government-supported research and technology diffusion activities (Fei, Ranis, and Kuo 1979). Some of these activities focused on new strains of rice and sugar, but most of them focused on new crops—cotton, fruits, vegetables, and, during the 1960s, asparagus and mushrooms. Agricultural output growth in Korea was much more modest in this early period, but accelerated later on (table 2) as the government belatedly recognized its neglect of the rural economy and sought to redress it.

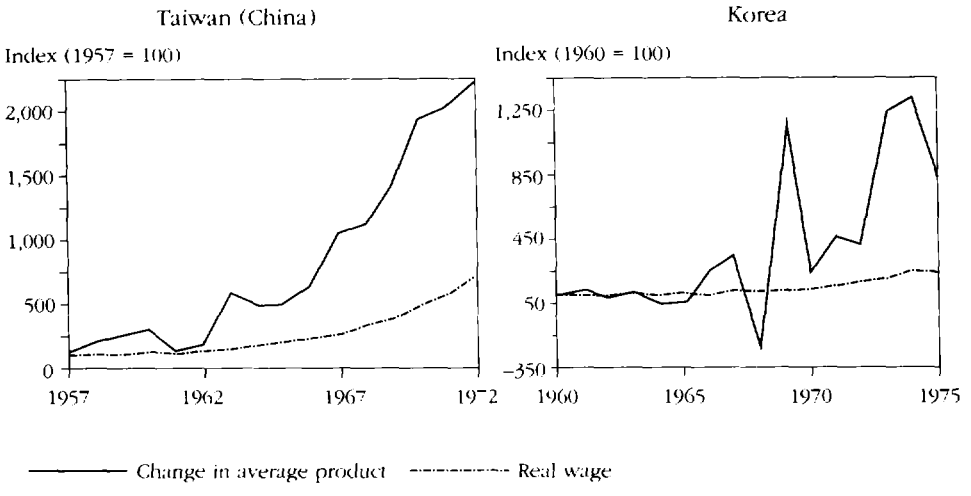
Arthur Lewis's classic concept of organizational dualism is empirically relevant to societies whose relatively preponderant agricultural sectors are initially characterized by heavy population pressure on scarce land (Lewis 1954, 1979). This means that, with the marginal product of agricultural labor at very low levels, real wages on largely family-operated farms are determined by an initial average product-related sharing mechanism, side by side with a commercialized, profit-maximizing nonagricultural sector in which wages (for unskilled workers) are tied to those in agriculture. Such extended family-related sharing has its roots in the agrarian tradition and serves as a form of insurance for those living not far from the margins of subsistence (see, for example, Scott 1976).

The increasing gap between changes in the average product of agricultural labor and agricultural wages (figure 1) permitted substantial agricultural surpluses to be generated, especially in Taiwan (China). Although labor productivity grew by 3.7 percent a year between 1955 and 1964 in Taiwan (China) (Kuo 1983), in Korea it grew at a more modest 2.4 percent a year over the same period (Das 1990). Growth was lower in Korea for several reasons. The initial conditions, including climate and soil fertility, were less favorable. Infrastructural investments during the Japanese period were less substantial. And the Korean government initially paid less attention to rural development.

### *Nonagricultural Activities*

Taiwan (China)'s agricultural surpluses went partly into the coffers of the government through land taxes, a fertilizer-rice barter program that represented a hidden tax on agriculture, and procurement by the armed forces at below-market prices. But an increasing proportion of the surplus that found its way into rural nonagricultural activities was contributed by private household savings. The household savings resulted from a combination of fast growth, the absence of a need to spend privately on education, the growth of financial institutions, and a major interest rate reform that took place in the 1950s. It is estimated that 15 to 20 percent of total agricultural output was transferred in the

Figure 1. *Real Wages and the Change in Average Product of Agricultural Labor in Taiwan (China), 1957–72, and Korea, 1960–75*




*Source:* Das (1990); Mason and others (1980); Government of Korea (1978); United Nations (various years b); Kuo (1983); Taiwan, China (various years a).

1950s and 1960s, 3 percent of which was in the form of private savings in the 1950s, and raised to 12 to 15 percent in the 1960s (Ho 1978).

Given the equal distribution of land on Taiwan (China), plus the shift toward more labor-intensive crop mixes and technologies, the demand for nonagricultural products was increasingly directed toward small- and medium-scale enterprises. These enterprises were largely in the rural areas, thus strengthening the mutual backward and forward linkages between changes in agricultural and nonagricultural output (see Ranis and Stewart 1993). The preponderance of small- and medium-scale firms clearly represented a central feature of the landscape, helping to explain the multidimensional success of the system in terms of growth, employment generation, and enhanced equity.

In employment terms, for example, a remarkable shift occurred in the allocation of rural household labor from agricultural to nonagricultural activities: the latter made up 29 percent of total employment in 1956, 47 percent in 1966, and 67 percent in 1980 (Taiwan, China, various years a). In income terms, the proportion earned by rural households outside of agriculture rose from 25 percent in 1962 to 43 percent in 1975 to 60 percent in 1980 (Fei, Ranis, and Kuo 1979). Taiwan (China)'s industrialization pattern thus clearly avoided the customary relative (sometimes absolute) decline in postcolonial rural nonagricultural activities. Rural nonagricultural activities were not—as so often occurs—competed out of existence by favored urban industry and services. In Korea, agriculture was not as dynamic and, partly as a consequence, industrialization was substantially more urban-oriented and large scale.

As a result of land reform and the increasingly labor-intensive mix of agricultural outputs, the distribution of rural incomes in Taiwan (China) also improved dramatically, with the rural income Gini declining from close to 0.50 in 1950 to 0.35 in 1957, 0.35 in 1966, and 0.31 in 1970.[2] As incomes rose, farmers' demand for both domestically produced consumer goods (forward linkages) and intermediate inputs (backward linkages) rose. At the same time, farmers' ability and willingness to take advantage of new investment opportunities outside of agriculture increased. Both trends further encouraged the growth of rural, small- and medium-scale industries and services.

Taiwan (China)'s nonagricultural output increased at the rate of about 12 percent during the 1950s (table 2). Korea's nonagricultural growth rate in the same period was not very different, but more of it was urban and financed by a combination of industrial reinvestment and foreign capital. Moreover, although exports began to play an increasing role in both systems, in Taiwan (China) 60 percent of overall growth between the early 1950s and mid-1960s was still domestically oriented, a fact that is often overlooked when the East Asian experience is cited (Taiwan, China, various years a). During this relatively mild import substitution subphase, the close contact between agriculture and rural industry and services induced mutually reinforcing changes in productivity.

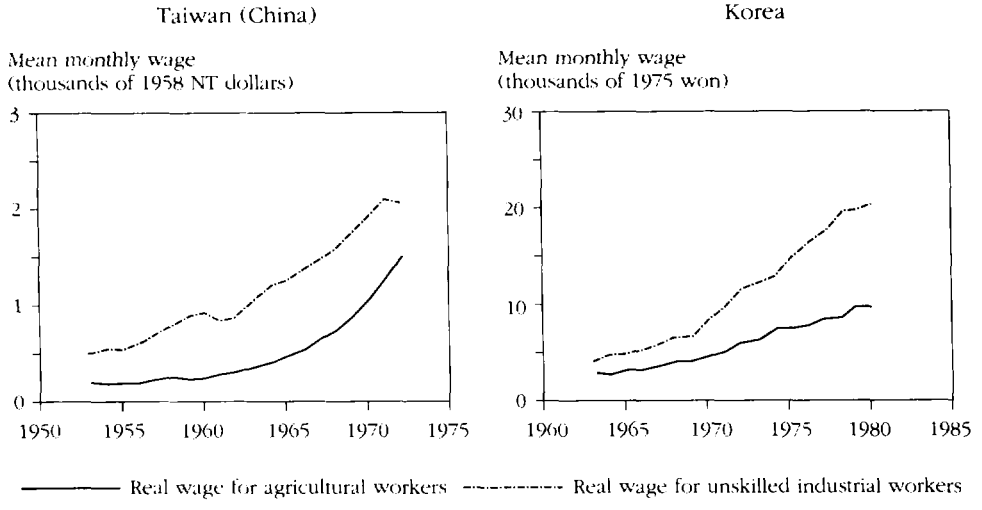
### *Labor Markets*

As long as labor was in surplus, nonagricultural wage rates in Taiwan (China) and Korea were at a modest margin above agricultural wage rates. Both rates moved up only slowly throughout the 1950s and 1960s (figure 2). Nonagricultural wages are represented by unskilled wages rather than average wages, which would distort the picture by incorporating changes in the output and therefore the skill mix. Capital accumulation and changes in technology in a labor-using direction rapidly raised industrial employment and incomes. Unskilled wages, tied to agricultural incomes, rose only modestly, and substantial nonagricultural profit margins also became available for savings and reinvestment.

Employment increased at more than 3 percent a year, even during the 1950s and early 1960s (Taiwan, China, various years a). Thus, the nonagricultural wage bill did not decline, in spite of the fact that the ample supply of labor kept wages from rising very much. This fact, plus the improvement in the distribution of agricultural incomes following land reforms and the marked changes in the composition of agricultural output in the direction of greater labor intensity, permitted the overall distribution of income to improve continuously (table 3). Private savings, generated largely by households and small businesses, contributed substantially to the high, and rising, national savings rate (table 4) as early as the 1950s and 1960s. The contrast with Korea's substantially lower domestic savings rate and consequently greater reliance on foreign savings is marked (table 5).

2. The 1950 and 1957 figures are based on farm size as a proxy for income (see Fei, Ranis, and Kuo 1979) and are therefore less reliable than the post-1964 figures based on data from Taiwan, China (various years b).

Figure 2. *Real Wages for Agricultural and Unskilled Industrial Workers in Taiwan (China), 1950–1975, and Korea, 1960–85*




*Source:* Kuo (1983); Taiwan, China (various years a); Das (1990); Mason and others (1980); Government of Korea 1978.

The gap between the real wages of industrial skilled and unskilled workers in Taiwan (China) and Korea (figure 3) was relatively modest by the standards of most developing countries, especially those in Latin America and Africa. There is little evidence of government intervention in the nonagricultural labor market through minimum wage legislation, public sector wage setting, or support of union activities. "Repressed" labor markets, as they are sometimes erroneously

Table 3. *Comparative Performance in Overall Development, Korea and Taiwan (China), 1953–90*

| | *Average annual growth in real GNP per capita (percent)* | | *Equity (Gini coefficient)*[a] | | *Long-term foreign debt*[b] *(millions of dollars)* | |
|---|---|---|---|---|---|---|
| *Years* | *Korea* | *Taiwan (China)* | *Korea* | *Taiwan (China)* | *Korea* | *Taiwan (China)* |
| 1953–60 | 1.80 | 4.63 | — | .440 | 25 | — |
| 1960–70 | 8.71 | 9.61 | .357 | .330 | 447 | 226 |
| 1970–80 | 1.46 | 10.92 | .313 | .289 | 10,681 | 2,864 |
| 1980–90 | 9.12 | 8.85 | .301 | .296 | 20,239 | 11,315 |

— Not available.

a. Data for Korea are for 1967, 1975, and 1984. Data for Taiwan (China) are for 1959, 1966, 1976, and 1986.

b. Data are for the end of the decade, for example 1960 for Korea for the 1953–60 period.

*Source:* Taiwan, China (various years a); Government of Korea (various years); United Nations (various years b); Mason and others (1980).

Table 4. *The Rate and Composition of Savings in Taiwan (China), 1952–88*
(percent)

| Year | Gross domestic savings as a percentage of GNP | Source | | |
|---|---|---|---|---|
| | | Private sector | General government | Public enterprises |
| 1952 | 15.3 | 50.93 | 35.58 | 13.49 |
| 1960 | 17.8 | 54.72 | 22.45 | 22.83 |
| 1965 | 20.7 | 69.55 | 12.64 | 17.80 |
| 1970 | 25.6 | 68.03 | 13.94 | 18.02 |
| 1980 | 32.3 | 60.35 | 24.49 | 15.16 |
| 1985 | 32.6 | 65.61 | 16.81 | 17.58 |
| 1988 | 34.9 | 64.22 | 22.76 | 13.02 |

*Source:* Taiwan, China (various years a).

Table 5. *The Rate and Composition of Savings in Korea, 1945–88*
(percent)

| Year | Gross domestic savings as a percentage of GNP | Source | |
|---|---|---|---|
| | | Private sector | Public sector |
| 1945 | 4.63 | 92.27 | 7.73 |
| 1960 | 6.79 | 63.90 | 36.10 |
| 1965 | 7.40 | 76.78 | 23.22 |
| 1970 | 17.30 | 62.37 | 37.63 |
| 1975 | 19.10 | 78.48 | 21.52 |
| 1980 | 21.90 | 68.94 | 31.06 |
| 1984 | 27.40 | 76.14 | 23.86 |
| 1988 | 30.00[a] | — | — |

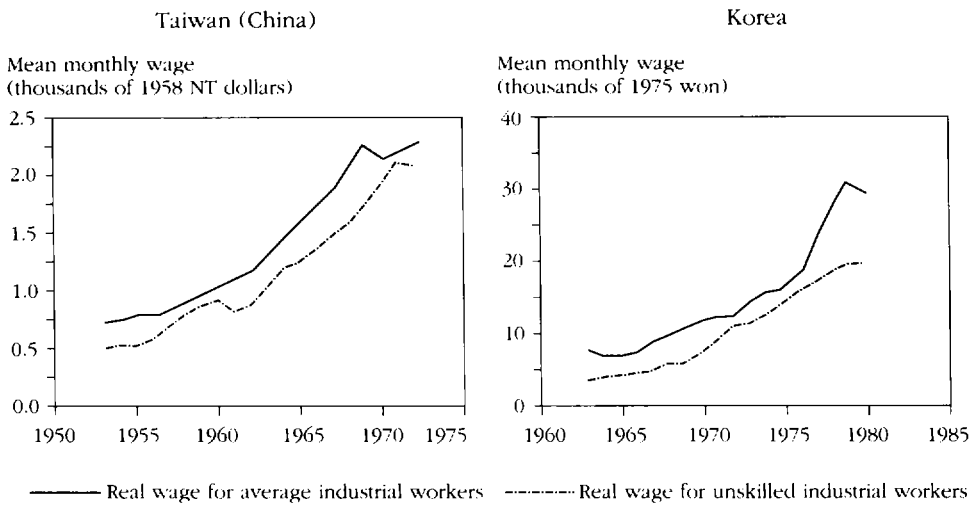
— Not available.
a. Percentage of GDP.
*Source:* Government of Korea (various years); United Nations (various years c).

referred to, are really no more than an unwillingness by the government to support an elite organized labor force at the expense of foregone employment opportunities for the rest. The fact that wage bills held steady or increased, while wage rates continued to reflect the underlying labor surplus condition during the 1950s and 1960s, actually helped improve the relative position of the working poor.

### *Building on Pre-existing Physical and Human Overheads*

Taiwan (China) and, once again to a lesser extent, Korea continued to build on the favorable rural infrastructure left by the Japanese. Taiwan (China)'s rural road network was further extended, and access to infrastructure climbed rapidly. Electricity coverage, especially critical for decentralized industrial growth in the rural areas, increased from 33 percent of households in 1952 to 99.7 percent in 1979. In 1952, 2 percent of households had telephones and 14.4 percent had piped water, compared with 71.1 percent and 63.8 percent, respectively, in 1979 (Taiwan, China, various years c). Moreover, with respect to the

Figure 3. *The Industrial Wage Gap in Taiwan (China), 1950–1975, and Korea, 1960–85*




*Source:* Kuo (1983); Taiwan, China (various years a, b); Government of Korea (1978); Kim (1990).

institutional infrastructure, Taiwan (China)'s Joint Commission on Rural Reconstruction (JCRR), supported by U.S. aid, was deployed to provide financial and technical assistance to decentralized farmers' associations left behind by Japan. Very similar associations were initially used for "top-down" instruction and taxation by the more centralized government in Korea. Both this and the fact that Korea let its agricultural terms of trade deteriorate were among the reasons for not only the referenced differential performance of agriculture in the two economies but also for a very different pattern of nonagricultural growth.

In Taiwan (China) most farmers' associations were focused on rice and sugar production, others specifically on irrigation, and still others on fishing. In addition to providing all-important research and extension services, the JCRR assisted with purchasing, marketing, warehousing, and processing. Farmers' associations organized cooperative banks and used positive real interest rates to induce members to deposit their own savings and receive locally supervised loans. Over time, the JCRR provided information and technical assistance on new crops and on linked nonagricultural activities in rural areas. Food processing, for example, came to represent the fastest-growing nonagricultural activity by the early 1960s.

Korea started with a less favorable initial base and neglected its agricultural sector during the 1950s and early 1960s in both the infrastructural and institutional dimensions. The relative contribution of total factor productivity to the lower growth of agricultural output in Korea was 36 percent in 1954–65 com-

Table 6. *Population by Level of Education in Taiwan (China), 1952–88, and Korea, 1944–85*
(percentage of population)

| Year | *Higher education* | *Secondary education* | *Primary education* | *Illiterate* | *Others*[a] |
|---|---|---|---|---|---|
| *Taiwan (China)* | | | | | |
| 1952 | 1.4 | 8.8 | 43.5 | 42.1 | 4.2 |
| 1960 | 1.9 | 12.4 | 54.1 | 27.1 | 4.5 |
| 1966 | 2.5 | 15.8 | 54.8 | 23.1 | 3.8 |
| 1970 | 3.7 | 26.5 | 51.8 | 14.7 | 3.3 |
| 1974 | 4.8 | 29.1 | 49.9 | 13.3 | 2.9 |
| 1980 | 7.1 | 36.9 | 43.3 | 10.3 | 2.4 |
| 1988 | 10.1 | 44.2 | 37.0 | 7.4 | 1.3 |
| *Korea* | | | | | |
| 1944 | 0.3 | 1.7 | 11.3 | 86.7 | n.a. |
| 1960 | 2.6 | 17.5 | 36.2 | 43.7 | n.a. |
| 1966 | 4.5 | 24.7 | 40.0 | 30.8 | n.a. |
| 1970 | 5.6 | 31.8 | 39.2 | 23.4 | n.a. |
| 1974 | 5.7 | 38.0 | 36.0 | 20.3 | n.a. |
| 1985 | 7.4 | 48.7 | 32.3 | 11.6 | n.a. |

n.a. Not applicable.
a. Includes people who have not completed primary education.
*Note:* Data are end-of-year values for percentage of population age six and over.
*Source:* Taiwan, China (various years a); for Korea, McGinn and others (1980) and Government of Korea (1985).

pared with 49 percent in Taiwan (China) in 1957–67. Until as late as 1970, Taiwan (China) spent twice as much as Korea on agricultural research as a percentage of agricultural output. The situation was reversed with the advent of the "New Village Movement" in the 1970s. The contribution of total factor productivity rose to 86 percent in 1970–77 in Korea and fell to 18 percent in 1973–77 in Taiwan, China (Chi-Ming and Yu 1982).

As for human capital formation, the emphasis in both countries during this period continued to be mainly on further improving literacy through compulsory universal primary education, as table 6 indicates. Total enrollment rates rose especially rapidly between 1953 and 1965, faster for women than for men. It is worth noting that Taiwan (China) not only had a substantial advantage in primary education and literacy early on but also that it held on to it over time.

### *Early Macroeconomic Stabilization*

Government policies during this subphase not only moved both systems toward positive real interest rates, breaking the back of inflation with the help of foreign aid inflows, but also began to gradually open them to foreign trade and investment. Quantitative controls began to be replaced by tariffs, and as early as 1956 a system of customs duty rebates for exports was instituted on Taiwan (China), followed by a major devaluation, unification of the exchange rate, and

maintenance of a relatively realistic rate after 1960. Indeed, by the early 1960s, the well-known "Nineteen Points" of macroeconomic reform had liberalized much of Taiwan (China)'s economy, facilitating its entry into the next, export-oriented subphase of development. A similar package of fairly orthodox macroeconomic and structural reforms was introduced a few years later in Korea.

In summary, Taiwan (China) began to get its macroeconomic budgetary and financial house in order in the 1950s, with the help of a temporary ballooning in foreign aid. Its relatively mild import substitution phase meant avoidance of the customary severe discrimination against agriculture and permitted a remarkably balanced process of rural growth. The major, foreign trade–oriented liberalization package of the early 1960s was once again helped by U.S. aid commitments. In the political economy sense, it was also assisted by the (1961) announcement that the United States intended to terminate such assistance by 1965. The combination of an intrinsic natural resources poverty and the impending end to economic aid by its major donor undoubtedly helped focus the minds of Taiwan (China)'s policymakers on the famous Nineteen Points of reform, permitting the economy to shift relatively smoothly into its competitive export-oriented subphase of development.

In the Republic of Korea these macro and structural policy shifts took place a little later. The aftermath of the Korean War, the continued confrontation with the Democratic People's Republic of Korea, plus the fact that the U.S. economic as well as military presence remained a feature of the landscape reduced the pressure on the Republic of Korea to undertake thoroughgoing reforms at the same pace. These factors led to a relatively greater neglect of the rural sector early on and to a greater concentration on large-scale urban industry partly financed with private foreign capital. However, in spite of these differences, and certainly in comparison with the standards in other developing countries, both these East Asian systems had been eminently successful by the mid-1960s in making a politically difficult transition to a vigorous, labor-intensive, industrial export subphase.

## III. The Labor-Intensive Export Substitution Subphase, 1963–72

The decade 1963–72 in Taiwan (China) was witness to a remarkable performance of rapid growth coupled with enhanced equity. Such a performance was made possible by the complementary effects of balanced, domestic, intersectoral growth and an increasingly powerful export drive. The contours of this new and more competitive development subphase and the policies that accommodated it are analyzed in this section.

The overall economic record is quite clear: during the 1960s, rapid growth, heavily supported initially by agriculture and then fueled increasingly by labor-intensive industrial expansion, resulted in a rapid generation of employment and an increasingly favorable income distribution. Savings and growth rates

soared (not only in Taiwan, China, but also in Korea), and total factor productivity rates in industry increased with enhanced openness. The economy witnessed a rapid rise in the importance of exports as well as rapid change in their composition. Emerging nonagricultural output mixes and technologies now permitted the reallocation at a 6 percent annual rate of the relatively unskilled, although increasingly literate, numerate, and therefore high-quality agricultural labor force. As a consequence, by the early 1970s even the combination of increased, especially female, participation rates (see Liu 1992), offshore (Southeast Asian) investment sourcing, and some (largely illegal) immigration of Southeast Asian workers could not postpone the advent of a labor shortage.

*Changes in Output, Labor Allocation, and the Distribution of Income*

Taiwan (China)'s agricultural output increased at an average rate of 4.3 percent during the 1960s (table 2). Labor productivity rose even faster, with the total agricultural labor force beginning to decline absolutely after 1965 (Taiwan, China, various years a). Most of the increase in agricultural output came through higher yields at the intensive margin because of a burst in agricultural research combined with dramatic changes in the composition of agricultural output, illustrated in table 7.

Nonagricultural output boomed in both Taiwan (China) and Korea (table 3). In Taiwan (China) the rate of increase in nonagricultural employment doubled to 6 percent annually (Das 1990; Taiwan, China, various years a). The role of agricultural savings in helping to finance this rapid expansion of nonagricultural output remained important, especially in Taiwan (China). With 15 to 20 percent of its agricultural output transferred as agricultural surplus, the sector continued to be a major, although gradually declining, source of fuel for the overall growth of the economy during the 1960s. By 1972 its contribution to gross national product (GNP) had fallen to 14 percent, from 36 percent in 1952, and its contribution to employment to 33 percent, from 56 percent. Manufacturing output had increased from 11 percent in 1952 to 32 percent, and employment in manufacturing from 12 percent to 25 percent (Taiwan, China, various years a).

This was clearly the decade when export-oriented industrialization—concentrated initially on processed agricultural goods and later increasingly on labor-intensive commodities based on imported raw materials—began to soar. The

Table 7. *Composition of Agricultural Output in Taiwan (China), 1960–88*
(percent)

| *Year* | *Rice* | *Other traditional crops* | *Nontraditional crops* |
|---|---|---|---|
| 1960 | 57.0 | 15.8 | 27.3 |
| 1970 | 45.0 | 12.9 | 42.1 |
| 1980 | 41.8 | 7.4 | 51.0 |
| 1988 | 27.1 | 7.9 | 65.0 |

*Note:* Numbers may not add to 100 because of rounding.
*Source:* Mao and Schive (1992, table 2.6).

1960s embody the best example of a mutually reinforcing relationship between the two blades of successful development: domestic balanced growth and labor-intensive industrial exports. Taiwan (China)'s exports as a percentage of a rapidly rising GNP increased from 10 percent in the early 1960s to more than 23 percent by 1972 (Taiwan, China, various years a). Table 8 provides evidence of the continued dramatic shift in the composition of these rapidly growing exports, a shift from more than 90 percent agricultural to more than 90 percent nonagricultural, in less than three decades (Taiwan, China, various years a).

Much of the labor reallocation in Taiwan (China) consisted of people not actually leaving rural households. Only 17 percent of the so-called migrants actually departed from the rural areas in 1963, and even as late as 1968 only 26 percent of newly absorbed industrial workers had actually physically migrated to urban centers (Ho 1976). This feature underlines the continued, remarkable expansion, not only absolutely but also relatively, of rural industry and services. As table 9 shows, the percentage distribution of employed persons by locality in manufacturing and services actually shifted in favor of the rural areas.

The two main factors that rendered such rapid growth in the 1960s consistent with an increasingly equitable distribution of income were already at play earlier and continued to assume an important role. One was that the smallest, that is, the poorest, farmers participated more than proportionally in the economy's booming rural industrial and service activities. The other was that these rural industries and services were increasingly labor-intensive, especially in the mid-1960s, when food processing still played a prominent role. As a consequence, the Simon Kuznets–Arthur Lewis prediction—that during a period of labor surplus the profit share would have to rise and the labor share fall and that the size distribution of income would therefore worsen—was not borne out. Employment increases in Taiwan (China) made up for the continued low level of unskilled industrial wages and permitted the wage share to rise and the Gini coefficient to fall into the 0.29 range during a period of very rapid growth (Fei, Ranis, and Kuo 1979). Counter to the usual assumptions, the shift that took

Table 8. *Composition of Exports in Taiwan (China) and Korea, 1952–88*
(percent)

| | *Taiwan (China)* | | *Korea* | |
|---|---|---|---|---|
| *Year* | *Agricultural* | *Industrial* | *Agricultural* | *Industrial* |
| 1952 | 91.9 | 8.1 | — | — |
| 1960 | 67.7 | 32.3 | 85.8 | 14.2 |
| 1970 | 21.4 | 78.6 | 22.6 | 77.4 |
| 1980 | 9.2 | 90.8 | 9.5 | 90.5 |
| 1988 | 5.5 | 94.5 | 6.2 | 93.8 |

— Not available.

*Note:* Agricultural exports are raw and processed agricultural products. Numbers may not add to 100 because of rounding.

*Source:* Taiwan, China (various years a); for Korea, Bank of Korea (various years) and Government of Korea (1980).

Table 9. *Distribution of Employed Persons by Locality in Taiwan (China), 1956 and 1966*
(percent)

| Sector | 1956 | | | 1966 | | |
|---|---|---|---|---|---|---|
| | *Cities* | *Towns* | *Rural areas* | *Cities* | *Towns* | *Rural areas* |
| Manufacturing | 42.68 | 10.06 | 47.26 | 37.02 | 10.80 | 52.18 |
| Construction | 44.84 | 11.45 | 43.71 | 43.92 | 12.27 | 43.81 |
| Utilities | 51.30 | 9.74 | 38.96 | 47.25 | 9.89 | 42.86 |
| Commerce | 44.36 | 9.98 | 45.66 | 42.34 | 10.56 | 47.10 |
| Transport and communications | 54.27 | 8.92 | 36.81 | 50.74 | 9.50 | 39.76 |
| Services | 41.16 | 9.73 | 49.11 | 34.44 | 9.38 | 56.17 |

*Note:* Cities are the seven largest cities. Towns are the nine largest towns. Rural areas are all other locations.

*Source:* Ho (1976, table 4) and data from the 1956 and 1966 population censuses.

place was not from an egalitarian agricultural sector to a less egalitarian nonagricultural sector, but from an egalitarian agricultural sector to an even more egalitarian nonagricultural sector.

The situation in Korea was not quite as favorable. Less success in mobilizing agriculture meant a lower savings rate, an earlier resort to food imports, and a greater reliance on foreign savings. The nonagricultural sector was more urban-oriented, large-scale, and capital-intensive. Thus, although Korea clearly still figures prominently as one of the success stories, both its Gini coefficients and its foreign debt levels were substantially higher than those of Taiwan (China) (table 3).

### *The Role of Government Policy*

As a semipublic, semiautonomous agency of the government of Taiwan (China), the JCRR was relatively isolated from political pressures. It was able to respond to local needs in research, extension, credit, and marketing at provincial, county, and township levels. The government also continued to build on the initial conditions by making substantial investments in power, transport, and communications throughout the island. Infrastructural investments were not directed to the support of particular regions or industries; indeed, there appears to have been a conscious effort by the government to be evenhanded and to permit fairly decentralized decisions in setting priorities. Rural and urban electricity rates were maintained at parity, so the power system, although not subsidized overall, was subject to internal cross-subsidization. After 1960 the government began to establish industrial districts, providing infrastructure, other overhead facilities, and horizontal linkages specifically to encourage decentralized industrial activities. By 1986, eighty-eight such districts were operational.

Booming nonagricultural growth was still heavily rural but also increasingly export-oriented. It was further facilitated by such policy inducements as the streamlining of previously established tariff rebates for exports (in lieu of fuller import liberalization) as well as the construction of several export-processing

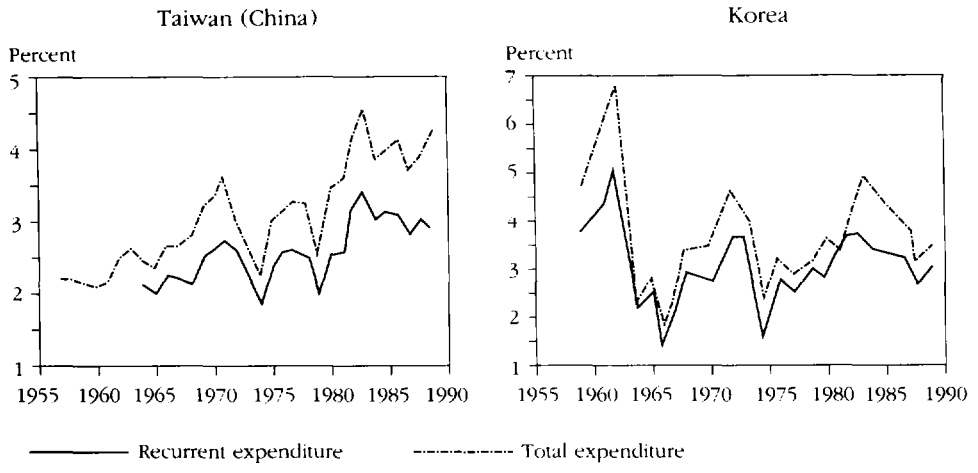
zones in the mid-1960s, supplemented by bonded factories dotting the landscape. In support of Lucas's (1993) theoretical note on Taiwan (China)'s success, this rapid change in East Asian exports and export composition was the result of an accelerating learning-by-doing process.

These useful transition devices for economies moving from a domestic to an external orientation proved especially relevant in the electronics and other light assembly industries that were part of the rapid output and employment generation spurt of the 1960s. Over the years the zones provided no more than 1 percent of total employment and 3 percent of investment, but they contributed 6 percent of exports (Fei, Ranis, and Kuo 1979). More important, they permitted entrepreneurs to compete abroad while the bulk of the domestic economy continued substantially protected. Gradually, as competence grew, additional forward or finishing stages were added, intermediate inputs were procured from, and final sales made into the domestic economy. Once labor surplus had come to an end, by the late 1960s and early 1970s, the historical mission of the zones had been accomplished and they began to atrophy. Thus, in a number of specific ways, beyond the accommodating shift in macro policies, the government provided direct institutional support. This support permitted Taiwan (China) both to explore domestic rural niche markets and to successfully enter growing international markets.

Taiwan (China)'s unusually decentralized industrialization effort was helped by its initially high literacy rates and early in-migration of a substantial number of small entrepreneurs from China. Then, as the economy moved through the 1960s and the industrial output mix changed, the government's educational emphasis shifted from compulsory primary to compulsory secondary education. Overall, expenditures on education rose from 2.1 percent of GNP (11 percent of the budget) in 1955 to 4.6 percent of GNP (20 percent of the budget) in 1970 (figure 4). In Korea, expenditures on education first fell and then rose, going from a low of 1.7 percent of GNP in 1965 to a high of 3.5 percent in 1970 (figure 4). As for quality, in both countries expenditures per student increased consistently over time for each level of education. Taiwan (China)'s expenditures on education per student increased sixfold between 1960 and 1975 (Taiwan, China, various years b). Those of Korea almost doubled over the same period (Mason and others 1980; see also Taiwan, China, various years b, and United Nations, various years a). Recurring expenditures were unusually high in relation to nonrecurring expenditures, as indicated in figure 4.

East Asian governments clearly understood the importance of human capital and, even more important, were flexible in shifting the emphasis of the educational structure as changes in the economy demanded. Although general literacy and numeracy objectives seem to have dominated at the outset, the six years of compulsory primary education in Taiwan (China) were extended to nine years in 1968, just when scarcity of unskilled labor, indexed by the first pronounced rise in unskilled real wages, appeared imminent. In addition, female participation rates increased sharply after 1967. This "turning point" also coincided with

Figure 4. *Total and Recurrent Expenditures on Education in Taiwan (China) and Korea, 1955–90*




*Source:* Taiwan, China, (various years b); United Nations (various issues a).

a much increased emphasis given to vocational as opposed to academic training at the secondary level. Vocational training increased sixfold between 1966 and 1974, a time when the nonagricultural labor force increased by 80 percent (Taiwan, China, various years a). Only 40 percent of Taiwan (China)'s high school students were in the vocational track in 1963, but by 1972 that percentage was 52 percent and by 1980 almost 70 percent (Taiwan, China, various years a). This pattern was much less pronounced in Korea, for both level and trend.

Approximately half the vocational education in Taiwan (China) was administered by private enterprises and half by government agencies. Public sector policies influenced the mix between academic and vocational or technical high schools by using differential tuitions as a carrot. Within the vocational schools, agriculture, commerce, and the industrial arts were combined. Offerings, unlike those in many other developing countries, were highly diversified, flexible, and responsive to changing market demands.

Both Taiwan (China) and Korea were able to count on a cheap and unskilled, yet efficient and literate labor force, extremely important ingredients for any competitive, export-oriented development drive. The percentage of population of ages sixteen to twenty-four with primary education in Taiwan (China) rose from 40 percent in 1953–54 to 60 percent by the early 1970s. The increase was even faster for women; the average number of years of schooling for women over age fifteen rose from 2.8 to 6.9 (Taiwan, China, various years a). Similar figures can be cited for Korea.

Education as a public sector responsibility represents a very strong cultural tradition in East Asia. Even as private sector schooling above the primary level—both secondary and higher education—increased in relative importance (from

10 percent in 1950 to 60 percent in 1980), demand continued to be strongest for entry into the higher-quality elite public institutions. Expenditures per public school student ran at 2.5 times those in private schools, and in public universities the student-teacher ratio was twelve to one, compared with twenty-nine to one in private universities (Taiwan, China, various years b; Mason and others 1980; Taiwan, China, various years a). Moreover, parents had to pay 50 percent of total costs in private schools, but only 7 percent in public high schools and universities (Taiwan, China, various years b; Mason and others 1980). Throughout the public school system, competitive examinations—part of the cultural heritage derived from the imperial examinations system of China—rather than ability to pay or family connection, determined entry.

### *The Labor Market*

As indicated in figure 1, agricultural wages continued to lag considerably behind the increase in agricultural productivity, thus setting the baseline for the remuneration of unskilled labor as long as the labor surplus persisted. Also, as shown in figure 2, unskilled nonagricultural wages, relatively undistorted by minimum wage legislation or union activity, continued at a reasonably small margin above agricultural wages. Unions concentrated heavily on working conditions, not collective bargaining on wages. Strikes were either illegal or officially frowned upon, and work stoppages were rather rare.

Unionization has had little effect on labor market conditions—at least until very recently. The percentage of the nonagricultural labor force unionized in Taiwan (China) (28 percent in 1988, from 18 percent in 1973) was consistently larger than in Korea (where it peaked at 16 percent), even though unions have to date played a substantially more prominent political role in Korea (Taiwan, China, various years a; United Nations 1990). In Taiwan (China), minimum wage legislation has been on the books since the 1950s, but these wages have been consistently at 40 to 60 percent of the actual manufacturing wage, that is, not "biting," even for unskilled labor (Taiwan, China, various years a). In both Korea and Taiwan (China), wages across different skill levels seem to have converged more as a result of supply-side changes brought about by more education and increased access rather than as a consequence of demand-induced changes in the industrial structure. The wage gap between skilled and unskilled labor in Taiwan (China) was only about 20 percent in 1953, approximately 30 percent in the mid-1960s, and once again 20 percent in 1970, very low by developing-economy standards (figure 3; Kuo 1983; Taiwan, China, various years a, b).

## IV. The Science and Technology Subphase—the Mid-1970s, 1980s, and Beyond

The latest subphase of development in Taiwan (China) and Korea may be characterized as dominated by three major phenomena. First, agriculture, having performed its mission—although less well and more belatedly in Korea—is

no longer a major catalyst and source of saving, and has instead become a subsidized appendage of the economy. Second, the increasing shortage of labor has caused a marked shift in the industrial output and export mix. Electronics, information, science, and technology-intensive exports soar and now substantially exceed those of textiles, garments, and other light industry. Third, the government is paying vastly increased attention to infrastructure related to science and technology and to institutional changes encouraging additional activity in research and development (R&D).

It is, of course, difficult to assess the impact of the combination of human capital accumulation, institutional construction, and R&D activity on bottom-line performance. In the mixed economy much depends on the competitiveness of the private sector and the support of public good externalities. In the absence of workably competitive conditions, the existence of large, unearned rents can be expected to lead to satisficing behavior and a reduced incentive for product and process innovations (see, for example, Otsuka, Ranis, and Saxonhouse 1988). Where workable competition does exist—more true for Taiwan (China) than for the *chaebol*-dominated industries of Korea—the externalities provided by the government can have a substantial payoff. One way to assess this issue is to examine the contribution of total factor productivity to growth. As expected, East Asia has done markedly better in this respect than, say, Latin America, and Taiwan (China) levels exceed those of Korea (see Page 1994; Young 1993; Pack 1992; Kim and Roemer 1979). Counter to current revisionism, high investment rates alone do not tell the full story.

As clearly demonstrated in table 2, the underlying, longstanding poverty of land and natural resources in Taiwan (China) and Korea has gradually reasserted itself. The "New Village Movement" in Korea gave a somewhat belated spurt to the agricultural sector there, but both Taiwan (China) and Korea are now net importers of agricultural goods. Farmers are being subsidized rather than taxed. Although 15 to 20 percent of Taiwan (China)'s agricultural output was transferred into nonagricultural investments during the 1950s and 1960s, the direction of resource flows reversed after 1973 (Ho 1978). A second land reform, involving land consolidation, increased mechanization, and the advent of specialized-cultivating firms, represented rational policy responses to the changing role of agriculture. Policies protecting and subsidizing farmers, by contrast, reflect political pressures of the kind familiar to most advanced industrial countries and represent an equally predictable but less rational response.

East Asia is still trying to prolong the life of its labor-intensive industries by investing in such low-wage neighbors as China, Indonesia, the Philippines, and Thailand. It is also closing its official eye to a limited number of unskilled, largely illegal, immigrants from these countries. An absolutely large volume of labor-intensive exports continues to flow predominantly to the industrial countries. As international trade theory predicts, however, an increasing proportion of East Asia's rapidly expanding high-tech and capital-intensive exports are destined for neighboring developing countries.

Government policy has accommodated shifts in the output mix in four main ways: first, by placing marginally greater reliance on public enterprises, given the still relatively underdeveloped domestic financial markets;[3] second, by focusing public policy on creating the necessary externalities on the human capital side; third, by considering institutional investments in science and technology as a high-priority public good; and fourth, by providing directed credit to export-oriented industries on an across-the-board basis.

Human capital formation has accommodated the accelerating shift in nonagricultural output mixes and technologies. By 1988 total expenditures on education in Taiwan (China) had risen to a new high of 5.2 percent of GNP. Moreover, in the 1950s, 50 percent of the total expenditures on education were spent on primary education, 35 percent on secondary education, and 15 percent on higher education; more recent figures are 30, 40, and 30 percent, respectively (Taiwan, China, various years b; Mason and others 1980). In addition, by 1989 fully 70 percent of the high schools in Taiwan (China) were vocational, representing a marked acceleration of earlier trends (Taiwan, China, 1989).

In both Taiwan (China) and Korea, higher education has continued to shift toward engineering and the natural sciences and away from the traditional humanities and agricultural concentrations. The human capital foundation for this shift clearly had been laid earlier. International comparisons for 1990, provided by the International Assessment of Educational Performance, indicate that Taiwanese and Korean students already occupied the two top positions in the sciences and mathematics at the age of thirteen.

The government intervened at the university level partly by shifting entry quotas, but partly also by using financial support as incentives favoring such new departments as industrial engineering, industrial design, automatic controls, and petro-chemistry. As a result, enrollments in graduate engineering and material science programs in Taiwan (China) virtually doubled (to 40 percent of total graduate enrollments) between the mid-1960s and late 1980s (Taiwan, China, 1989). It seems that policymakers recognized that more engineering and science-oriented investments were now required. The countries were no longer able to depend for their total factor productivity increases either on the shift from agriculture to nonagriculture or on the adaptation of relatively simple imported technologies and designs associated with labor-intensive output mixes (see, for example, Westphal 1985). The increased need to shift to domestically generated Schumpeterian-type innovative responses meant an accelerating shift toward more high-tech labor as well as R&D.

Every year from 1960 to 1967 approximately 15 percent of Taiwan (China)'s college graduates went to the United States for postgraduate study, two-thirds of them in science and engineering, but only 4.5 percent of them returned each year. After 1973 the government initiated a major effort to encourage the re-

3. This reliance continues to be reflected in Taiwan (China)'s current six-year plan, which permits the small savings of many individuals to be channeled into large investment projects.

verse flow of this high-talent labor by actively recruiting in such locations as California's Silicon Valley and offering Taiwanese graduates substantial financial incentives to return. Although only a handful of highly trained graduates returned before 1970, by 1986 more than 90 percent did so. It is estimated that in recent years almost 10,000 Taiwanese graduates of U.S. universities have returned to take up positions related to science and technology. Many have become venture capitalists in their own right (Taiwan, China, various years b).

The construction of complementary science and technology infrastructure in Taiwan included the establishment of a Cabinet Committee on R&D (1965), the National Science Council (1967), the Chung Shan Institute of Science and Technology (1965), the Industrial Training and Research Institute (ITRI; 1973), the Institute for Information Industry (1979), and the Shin Chu Science-Based Industrial Park (1979). Most prominent of these has been the last mentioned, which has encouraged new ventures in high-tech export-oriented areas by providing public facilities to private firms on favorable terms. These terms include five-year tax exemptions, a ceiling on taxes at 22 percent thereafter, subsidized rent and credit facilities, and other amenities. The Industrial Park, moreover, guaranteed close physical and intellectual contacts with academic and private commercial interests; in recent years fully 70 percent of the companies in the park have been managed by repatriated Chinese.

ITRI, located nearby, carries on basic industrial research in areas that might otherwise be underrepresented because of the continued predominance of small- and medium-scale firms. ITRI's $400 million annual budget permits a substantial volume of in-house product and process research in adapting foreign technology through both reverse engineering and the generation of original patents. ITRI also sells technical assistance and market information to small- and medium-scale firms and, although it is still heavily focused on government basic research, private contracts are growing gradually.[4]

Overall, 60 percent of Taiwan (China)'s industrial R&D is carried out by the private sector and 40 percent within various public sector institutes. The government funds 60 percent and the private sector 40 percent of the total expenditure on R&D (Taiwan, China, various years a). In other words, formal private sector R&D is generally carried out within large firms whereas the government supports R&D and training-related activities of benefit to the large population of small- and medium-scale firms.

During the 1980s, both Korea and Taiwan (China) increased their formal R&D expenditures from approximately 1 percent to 2.5 percent of GNP. That was a substantial expansion symptomatic of the new subphase of growth they had entered, although still substantially less than what mature economies allocate. Moreover, these figures do not capture the informal, blue-collar tinkering

4. A recent visit to ITRI convinced me that a major explicit contribution of the institute, in addition to providing information and venture capital to medium- and small-scale entrepreneurs, has been the generation in-house of a group of technologically oriented venture capitalists, who, given ITRI's remarkably high staff turnover rate, make their way each year into the private sector.

type of R&D that takes place in the machine shops and on the factory floors of many small- and medium-scale establishments. Basic research is still carried on largely by the public sector and applied research by the private sector. But product development, which comprised only 40 percent of R&D in 1970 and is virtually exclusively the province of private firms, now constitutes more than 70 percent of the total R&D expenditure. New patents, although an admittedly imperfect indicator of innovative activity, increased from 2,770 in 1980 to 10,379 in 1990 (Taiwan, China, various years a). Over the same decade there has been an almost tenfold increase in paper citations, a measure of the quality of the basic research output of government-sponsored institutes and universities.

## V. Some Conclusions

If any common ground exists between adherents of the Washington Consensus, as purveyed by the IMF and the World Bank, and its heterodox critics, it is that developing economies, to be successful, must get their "fundamentals" right. That is, they must practice fiscal and monetary restraint and maintain a reasonably realistic exchange rate. By contrast, disagreements have focused on industrial policy, including what the appropriate role for governments and markets should be and on how micro structural reforms should be sequenced. What is still required, in my view, is a broader multisectoral view with typological and historical sensitivity, that is, concern with a country's initial conditions as well as the subphase of the transition growth process that has been reached.

The East Asian economies were, at the outset, relatively small, poor in natural resources, rich in human resources, and endowed with relatively good physical and institutional infrastructure. All this was as important as the fact that their newly independent governments continued to actively build on that base. It is also important that the typically interventionist, import substitution subphase was relatively mild and relatively short-lived. What proved critical, especially for Taiwan (China)'s success, was the early avoidance of agricultural neglect, the encouragement of the agricultural sector's mutually beneficial interactions with a dispersed nonagricultural sector, and the timely shift from land-based to human resources. The controversy swirling around *The East Asian Miracle* study (World Bank 1993) focused unduly on recent vintage industrial credit policy.

Korea's long-run pattern was broadly similar to Taiwan (China)'s, but there were also important and instructive differences. For example, Korea paid relatively less attention to the agricultural sector and to rural industry early on and was more tempted by the notion that large-scale, urban, export-oriented activities could pull the economy along. High growth rates comparable to Taiwan (China)'s were indeed achieved. However, because the financing of that growth received less support from agricultural savings, it required a larger volume of foreign capital, leading to a more capital-intensive industrialization path and a much larger foreign debt than was the case in Taiwan (China) (table 3).

Both Taiwan (China) and Korea initially highlighted literacy and numeracy by concentrating on primary education. They then moved flexibly to emphasizing vocational and, still later, graduate-level science and technology-oriented education—always in line with the system's continuously changing requirements. Labor markets were permitted to behave competitively; that is, there was very little intervention either through government support of unions or through minimum wage legislation with any "bite." Public policy was instead clearly focused on easing the path of an early transition from inward- to outward-oriented development.

The East Asian experience, moreover, confirms the usefulness of government actions in support of agricultural research and small- and medium-scale industrial development. There were critical government interventions in fashioning the appropriate organizational and institutional infrastructure, reducing transactions costs, and providing the necessary facilities and policies to ease the transition. Finally, there was the need to create temporary innovation profits in the hands of private industrialists through R&D and the strengthening of patents and other intellectual property rights.

Moreover, success did not lead the government to become excessively active in promoting growth. The importance of prudence in sequencing, and flexibility in implementing, reforms becomes even more evident when put to a stronger test during times of adversity. Unlike many other developing economies, Taiwan (China) was willing to accept temporarily lower growth rates in the face of negative external shocks such as the oil crisis, but maintained the direction of reforms without major zigzags in the various policy arenas. The avoidance of periods of relatively doctrinaire interventionism alternating with equally doctrinaire free market episodes was a key ingredient of its success.

Both macro and structural policies tended to oscillate more in Korea, although around an overall favorable trend, partly as a consequence of greater political instability, but largely because of a less consistent set of domestic policies. There were indeed times when Korea's behavior temporarily resembled that of the more "typical" developing economy, characterized by growth activism, resort to large-scale monetary expansion, deficit financing, and high levels of inflation. However, such deviations were temporary, with the system always able to return to its underlying trend toward liberalization. Consequently, the long-run performance of Korea, like that of Taiwan (China), has nevertheless been outstanding by any international standard—even if the results, in terms of foreign debt and income distribution, were not as favorable. What clearly marks off East Asia from the average developing-economy case is the remarkable ability of the authorities to recognize mistakes and quickly take appropriate corrective action.

In brief, public policies in East Asia consistently accommodated the changing needs of the economy, rather than specifically directing its path. Over time, East Asia has shown the necessary entrepreneurial and policy flexibility to weather inevitable exogenous shocks and turning points and thus managed to "stay the course" en route to economic maturity. It refused to retreat to stop-go restric-

tionist regimes at times of crisis and continuously adopted accommodating policy mixes that recognized the ever more critical reliance on human capital and technology in navigating the transitional growth process.

## REFERENCES

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

Amsden, Alice. 1989. *Asia's Next Giant: South Korea and Late Industrialization.* New York: Oxford University Press.

Bank of Korea. Various years. *Economic Statistics Yearbook.* Seoul.

Chi-Ming, Hou, and T. S. Yu, eds. 1982. *Agricultural Development in China, Japan, and Korea.* Taipei, Taiwan (China): Academia Sinica.

Das, Dilipk. 1990. *Korean Economic Dynamism.* Paris: Macmillan.

Fei, John, Gustav Ranis, and Shirley Kuo. 1979. *Growth with Equity: The Taiwan Case.* Oxford: Oxford University Press.

Government of Korea. 1978. *Major Statistics of the Korean Economy.* Seoul: Economic Planning Board.

_____. 1980. *Statistical Yearbook of Foreign Trade.* Seoul: Office of Customs Administration.

_____. 1985. *Supplementary Tables, Census of Population.* Taiwan: National Bureau of Statistics, Economic Planning Board.

_____. Various years. *Korean Statistical Yearbook.* Economic Planning Board, Seoul.

Ho, Samuel. 1976. "The Rural Non-Farm Sector in Taiwan." World Bank Studies in Employment and Rural Development. Working Paper 32. World Bank, Washington, D.C. Processed.

_____. 1978. *Economic Development of Taiwan, 1860–1970.* New Haven, Conn.: Yale University Press.

Jones, Leroy. 1975. *Public Enterprise and Economic Development.* Seoul: Korea Development Institute.

Kim, C. 1990. "Labor Market Development." In Jene K. Kwow, ed., *Korean Economic Development.* New York: Greenwood Press.

Kim, Kwang Suk, and Michael Roemer. 1979. "Growth and Structural Transformation." In *Studies in the Modernization of the Republic of Korea: 1945–1975.* Cambridge, Mass.: Harvard University Press.

Kuo, Shirley. 1983. *The Taiwan Economy in Transition.* Boulder, Colo.: Westview Press.

Kuo, Shirley, Gustav Ranis, and John Fei. 1981. *The Taiwan Success Story: Rapid Growth with Improved Distribution in the Republic of China, 1952–1979.* Boulder, Colo.: Westview Press.

Lewis, W. A. 1954. "Economic Development with Unlimited Supplies of Labor." *The Manchester School* 22(2):139–91.

_____. 1979. "The Dual Economy Revisited." *The Manchester School* 47(3):211–29.

Liu, Peter. 1992. "Science, Technology, and Human Capital Formation." In Gustav Ranis, ed., *Taiwan: From Developing to Mature Economy.* Boulder, Colo.: Westview Press.

Lucas, Robert. 1993. "Making a Miracle." *Econometrica* 61:251–72.

Mao, Y. K., and Chi Schive. 1992. "Agricultural and Industrial Development of the Republic of China." Council of Agriculture, The Executive Yuan, Taipei, Taiwan (China). Processed.

Mason, Edward, M. Kim, Dwight H. Perkins, Kwang Suk Kim, and David Cole. 1980. *The Economic and Social Modernization of the Republic of Korea.* Cambridge, Mass.: Harvard University Press.

McGinn, Noel F., Donald Snodgrass, Yung Bong Kim, Shin-Bok Kim, and Quee-Young Kim. 1980. *Education and Development in Korea.* Cambridge, Mass.: Harvard University Press.

Otsuka, Keijiro, Gustav Ranis, and Gary Saxonhouse. 1988. *Comparative Technology Choice: The Indian and Japanese Cotton Textile Industries.* London: Macmillan.

Pack, Howard. 1992. "New Perspectives on Industrial Growth in Taiwan." In Gustav Ranis, ed., *Taiwan: From Developing to Mature Economy.* Boulder, Colo.: Westview Press.

Page, John. 1994. "The East Asian Miracle: An Introduction." *World Development* 22(April):615–25.

Ranis, Gustav. 1984. "Typology in Development Theory: Retrospective and Prospects." In Moshe Syrquin, Lance Taylor, and Larry Westphal, eds., *Economic Structure and Performance: Essays in Honor of Hollis B. Chenery.* Orlando, Fla.: Academic Press.

Ranis, Gustav, and Frances Stewart. 1993. "Rural Nonagricultural Activities in Development: Theory and Application." *Journal of Development Economics* 40(1):75–101.

Scott, J. C. 1976. *The Moral Economy of the Peasant.* New Haven, Conn.: Yale University Press.

Taiwan (China). 1989. *Educational Statistics of the Republic of China.* Taipei: Ministry of Education.

_____. Various years a. *Taiwan Statistical Data Book.* Taipei: Ministry of Finance and Directorate-General of Budget, Accounting, and Statistics, Council for Economic Planning and Development.

_____. Various years b. *Statistical Yearbook of the Republic of China.* Taipei: Directorate-General of the Budget, Accounting, and Statistics.

_____. Various years c. *Taiwan Statistical Abstract of Transport and Communication.* Taipei: Ministry of Transportation and Communications.

United Nations. Various years a. UNESCO *Statistical Yearbook.* New York.

_____. Various years b. *United Nations Statistical Yearbook.* New York.

_____. Various years c. *National Income Accounts Statistics Yearbook.* New York.

Wade, Robert 1990. *Governing the Market: Economic Theory and the Role of the Government in East Asian Industrialization.* Princeton, N.J.: Princeton University Press.

Westphal, Larry. 1985. "Reflections on the Republic of Korea's Acquisition of Technological Capability." In Nathan Rosenberg and Carlos Frischtak, eds., *International Technology: Concepts, Measures, and Comparisons.* New York: Praeger.

Williamson, John. 1990. "What Washington Means by Policy Reform." In John Williamson, ed., *Latin American Adjustment: How Much Has Happened?* Washington, D.C.: Institute for International Economics.

World Bank. 1991. *World Development Report 1991.* New York: Oxford University Press.

_____. 1993. *The East Asian Miracle: Economic Growth and Public Policy.* New York: Oxford University Press.

Wu, Rong-I. 1972. "Some Aspects of Export Trade and Economic Development in Taiwan." University of Louvain, Department of Economics, Louvain, Belgium. Processed.

Young, Allyn. 1993. "Lessons from the East Asian NIC's: A Contrarian View." NBER Working Paper 4482. National Bureau of Economic Research, Cambridge, Mass. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3 535–536

# Cumulative Index of Authors

## INDEX TO VOLUME 9 (1995)

THE WORLD BANK ECONOMIC REVIEW, VOL. 9, NO. 3 537-538

# Cumulative Index of Titles

## INDEX TO VOLUME 9 (1995)

# *World Bank Publications Order Coupon*

**CUSTOMERS IN THE UNITED STATES:**
Complete this coupon and return to

**The World Bank**
**Box 7247-8619**
**Philadelphia, PA 19170-8619**
**U.S.A.**

Charge by credit card by calling (202) 473-1155 or fax this completed order coupon to (202) 522-2627.

**CUSTOMERS OUTSIDE THE UNITED STATES:**
Contact your local World Bank Publications distributor for information on prices in local currency and payment terms. (See next page for a complete list of distributors.) If no distributor is listed for your country, use this order form and return it to the U.S. address. **Orders that are sent to the U.S. address from countries with distributors will be returned to the customer.**

| Quantity | Title | Stock Number | Price | Total |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

* If a purchase order is used, actual postage will be charged. If payment is by check or credit card, postage and handling charges are US$5.00 per order. For air mail delivery outside the U.S., include US$8.00 for the first item and US$6.00 for each additional item.

Subtotal US$ _____

Postage and handling* US$ _____

Total US$ _____

## CHECK METHOD OF PAYMENT

❑ Enclosed is my check payable to The World Bank.

❑ Charge my ❑ VISA ❑ MasterCard ❑ American Express

_____
Credit card account number

_____ _____
Expiration date Signature

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

**PLEASE PRINT CLEARLY**

Name _____

Firm _____

Address _____

City _____ State _____ Postal code _____

Country _____ Telephone _____

# *Distributors of World Bank Publications*

*Prices and terms vary by country*

**ARGENTINA**
Carlos Hirsch, SRL
Galeria Guemes
Florida 165, 4th Floor-Ofc.
453/465
1333 Buenos Aires

Oficina del Libro Internacional
Alberti 40
1082 Buenos Aires

**AUSTRALIA, PAPUA NEW GUINEA, FIJI, SOLOMON ISLANDS, VANUATU, AND WESTERN SAMOA**
D.A. Information Services
648 Whitehorse Road
Mitcham 3132
Victoria

**AUSTRIA**
Gerold and Co.
Graben 31
A-1011 Wien

**BANGLADESH**
Micro Industries Development Assistance Society (MIDAS)
House 5, Road 16
Dhanmondi R/Area
Dhaka 1209

**BELGIUM**
Jean De Lannoy
Av. du Roi 202
1060 Brussels

**BRAZIL**
Publicacoes Tecnical Internacionais Ltda.
Rua Pexoto Gomide, 209
01409 Sao Paulo, SP

**CANADA**
Le Diffuseur
151A Boul. de Mortagne
Boucherville, Québec
J4B 5E6

Renouf Publishing Co.
1294 Algoma Road
Ottawa, Ontario K1B 3W8

**CHINA**
China Financial & Economic Publishing House
8, Da Fo Si Dong Jie
Beijing

**COLOMBIA**
Infoenlace Ltda.
Apartado Aereo 34270
Bogota D.E.

**COSTA RICA, BELIZE, GUATEMALA, HONDURAS, NICARAGUA, PANAMA**
Chispas Bookstore
75 Meters al Norte del Hotel Balmoral en calle 7
San Jose

**COTE D'IVOIRE**
Centre d'Edition et de Diffusion Africaines (CEDA)
04 B.P. 541
Abidjan 04 Plateau

**CYPRUS**
Center of Applied Research
Cyprus College
6, Diogenes Street, Engomi
P.O. Box 2006
Nicosia

**CZECH REPUBLIC**
National Information Center
PO Box 668
CS-113 57 Prague 1

**DENMARK**
SamfundsLitteratur
Rosenoerns Allé 11
DK-1970 Frederiksberg C

**EGYPT, ARAB REPUBLIC OF**
Al Ahram
Al Galaa Street
Cairo

The Middle East Observer
41, Sherif Street
Cairo

**FINLAND**
Akateeminen Kirjakauppa
P.O. Box 23
FIN-00371 Helsinki

**FRANCE**
World Bank Publications
66, avenue d'Iéna
75116 Paris

**GERMANY**
UNO-Verlag
Poppelsdorfer Allee 55
53115 Bonn

**GREECE**
Papasotiriou S.A.
35, Stournara Str.
106 82 Athens

**HONG KONG, MACAO**
Asia 2000 Ltd.
46-48 Wyndham Street
Winning Centre
7th Floor
Central Hong Kong

**HUNGARY**
Foundation for Market Economy
Dombovari Ut 17-19
H-1117 Budapest

**INDIA**
Allied Publishers Private Ltd.
751 Mount Road
Madras - 600 002

**INDONESIA**
Pt. Indira Limited
Jalan Borobudur 20
P.O. Box 1181
Jakarta 10320

**IRAN**
Kowkab Publishers
P.O. Box 19575-511
Tehran

**IRELAND**
Government Supplies Agency
4-5 Harcourt Road
Dublin 2

**ISRAEL**
Yozmot Literature Ltd.
P.O. Box 56055
Tel Aviv 61560

R.O.Y. International
P.O. Box 56055
Tel Aviv 61560

**ITALY**
Licosa Commissionaria Sansoni SPA
Via Duca Di Calabria, 1/1
Casella Postale 552
50125 Firenze

**JAMAICA**
Ian Randle Publishers Ltd.
206 Old Hope Road
Kingston 6

**JAPAN**
Eastern Book Service
Hongo 3-Chome, Bunkyo-ku 113
Tokyo

**KENYA**
Africa Book Service (E.A.) Ltd.
Quaran House, Mfangano St.
P.O. Box 45245
Nairobi

**KOREA, REPUBLIC OF**
Pan Korea Book Corporation
P.O. Box 101, Kwangwhamun
Seoul

Korean Stock Book Centre
P.O. Box 34
Yeoeida
Seoul

**MALAYSIA**
University of Malaya Cooperative
Bookshop, Limited
P.O. Box 1127, Jalan PantaiBaru
59700 Kuala Lumpur

**MEXICO**
INFOTEC
Apartado Postal 22-860
14060 Tlalpan, Mexico D.F.

**NETHERLANDS**
De Lindeboom/InOr-Publikaties
P.O. Box 202
7480 AE Haaksbergen

**NEW ZEALAND**
EBSCO NZ Ltd.
Private Mail Bag 99914
New Market
Auckland

**NIGERIA**
University Press Limited
Three Crowns Building Jericho
Private Mail Bag 5095
Ibadan

**NORWAY**
Narvesen Information Center
Book Department
P.O. Box 6125 Etterstad
N-0602 Oslo 6

**PAKISTAN**
Mirza Book Agency
65, Shahrah-e-Quaid-e-Azam
P.O. Box No. 729
Lahore 54000

Oxford University Press
5 Bngalore Town
Sharae Faisal
P.O. Box 13033
Karachi-75350

**PERU**
Editorial Desarrollo SA
Apartado 3824
Lima 1

**PHILIPPINES**
International Book Center
Suite 1703, Cityland 10
Condominium Tower 1
Ayala Avenue, H.V. dela
Costa Extension
Makati, Metro Manila

**POLAND**
International Publishing Service
Ul. Piekna 31/37
00-577 Warszawa

**PORTUGAL**
Livraria Portugal
Rua Do Carmo 70-74
1200 Lisbon

**SAUDI ARABIA, QATAR**
Jarir Book Store
P.O. Box 3196
Riyadh 11471

**SINGAPORE & TAIWAN**
Gower Asia Pacific Pte Ltd.
Golden Wheel Building
41, Kallang Pudding, #04-03
Singapore 1334

**SLOVAK REPUBLIC**
Slovart G.T.G. Ltd.
Krupinska 4
P.O. Box 152
852 99 Bratislava 5

**SOUTH AFRICA, BOTSWANA**
*For single titles:*
Oxford University Press
Southern Africa
P.O. Box 1141
Cape Town 8000

*For subscription orders:*
International Subscription Service
P.O. Box 41095
Craighall
Johannesburg 2024

**SPAIN**
Mundi-Prensa Libros, S.A.
Castello 37
28001 Madrid

Librería Internacional AEDOS
Consell de Cent, 391
08009 Barcelona

**SRI LANKA & THE MALDIVES**
Lake House Bookshop
P.O. Box 244
100, Sir Chittampalam A.
Gardiner Mawatha
Colombo 2

**SWEDEN**
FritzesCustomer Service
S-106 47 Stockholm

**SWITZERLAND**
Librairie Payot
Case postale 3212
CH 1002 Lausanne

Van Dierman Editions Techniques
P.O. Box 465
CH 1211 Geneva 16

**TANZANIA**
Oxford University Press
Maktaba Street
P.O. Box 5299
Dar es Salaam

**THAILAND**
Central Department Store
306 Silom Road
Bangkok

**TRINIDAD & TOBAGO,**
Systematics Studies Unit
#9 Watts Street
Curepe
Trinidad, West Indies

**UGANDA**
Gustro Ltd.
1st Floor, Room 4, Geogiadis
Chambers
P.O. Box 9997
Plot (69) Kampala Road
Kampala

**UNITED KINGDOM**
Microinfo Ltd.
P.O. Box 3
Alton, Hampshire GU34 2PG
England

**ZAMBIA**
University Bookshop
Great East Road Campus
P.O. Box 32379
Lusaka

**ZIMBABWE**
Longman Zimbabwe (Pte.) Ltd.
P.O. Box ST 125
Southerton
Harare

*Coming in the next issue of*

# THE WORLD BANK ECONOMIC REVIEW

**January 1996**
**Volume 10, Number 1**

Articles including . . .

- **Is the Debt Crisis History? Recent Private Capital Inflows to Developing Countries**
  *by Michael Dooley, Eduardo Fernandez-Arias, and Kenneth Kletzer*

- **How Important to India's Poor Is the Sectoral Composition of Economic Growth?**
  *by Martin Ravallion and Gaurav Datt*

And a symposium on fertility in Sub-Saharan Africa, including . . .

- **An Introduction**
  *by Martha Ainsworth*

- **Fertility and Child Mortality in Côte d'Ivoire and Ghana**
  *by Kofi Benefo and T. Paul Schultz*

- **Contraceptive Use and the Quality, Price, and Availability of Family Planning in Nigeria**
  *by Bamikale J. Feyisetan and Martha Ainsworth*

- **Fertility, Contraceptive Choice, and Public Policy in Zimbabwe**
  *by Duncan Thomas and John Maluccio*