Volume 17 • 2003 • Number 3

OXFORD

ISSN 0258-6770

THE WORLD BANK ECONOMIC REVIEW

The World Bank Economic Review is a professional journal for the dissemination of World Bank–sponsored and outside research that may inform 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 from work conducted by World Bank staff and consultants and from papers submitted by outside researchers. Before being accepted for publication, all articles will be reviewed by two referees who are not members of the Bank's staff and one World Bank staff member. Articles must also be recommended by a member of the Editorial Board. Non-Bank contributors are requested to submit a proposal of not more than two pages in length to the Editor or a member of the Editorial Board before sending in their paper.

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, 1818 H Street, Washington, DC 20433, USA, or wber@worldbank.org.

For more information, please visit the Web sites of the *Economic Review* at www.wber.oupjournals.org, the World Bank at www.worldbank.org, and Oxford University Press at www.oupjournals.org.

THE WORLD BANK ECONOMIC REVIEW

Volume 17 • 2003 • Number 3

The World Bank Economic Review (ISSN 0258-6770) is published three times a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009 for The International Bank for Reconstruction and Development/THE WORLD BANK. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Economic Review*, 66, avenue d'Iéna, 75116 Paris, France. E-mail: wber@worldbank.org.

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

SUBSCRIPTIONS: Subscription is on a yearly basis. The annual rates are US$40 (£30 in UK and Europe) for individuals; US$101 (£71 in UK and Europe) for academic libraries; US$123 (£84 in UK and Europe) for corporations. Single issues are available for US$16 (£12 in UK and Europe) for individuals; US$43 (£30 in UK and Europe) for academic libraries; US$52 (£35 in UK and Europe) for corporations. All prices include postage. Individual rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address. *Subscriptions are provided free of charge to non-OECD countries.* All subscription requests, single issue orders, changes of address, and claims for missing issues should be sent to:

North America: Oxford University Press, Journals Customer Service, 2001 Evans Road, Cary, NC 27513-2009, USA. Toll-free in the USA and Canada: 800-852-7323, or 919-677-0977. Fax: 919-677-1714. E-mail: jnlorders@oup-usa.org.

Elsewhere: Oxford University Press, Journals Subscriptions Department, Great Clarendon Street, Oxford OX2 6DP, UK. Tel: +44 1865 353907. Fax: +44 1865 353485. E-mail: jnls.cust.serv@oup.co.uk.

ADVERTISING: Helen Pearson, Oxford Journals Advertising, P.O. Box 347, Abingdon SO, OX14 1GJ, UK. Tel/Fax: +44 1235 201904. E-mail: helen@oxfordads.com.

BACK ISSUES: The current plus two back volumes are available from Oxford University Press. Previous volumes (from 1997) can be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: 518-537-4700.

INDEXING AND ABSTRACTING: *The World Bank Economic Review* is indexed and/or abstracted by *CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/EconLit, PAIS International*, RePEc (*Research in Economic Papers*), and *Social Sciences Citation Index*. The microform edition is available through ProQuest (formerly UMI), 300 North Zeeb Road, Ann Arbor, MI 48106, USA.

PAPER USED: *The World Bank Economic Review* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

POSTAL INFORMATION: *The World Bank Economic Review* (ISSN 0258-6770) is published three times a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Send address changes to *The World Bank Economic Review*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009.

the century. And there have been serious health shocks, such as AIDS and

institutional setting. Section IV exploits the survey data to greater depth by taking a micro perspective, assessing how poor households have been affected by the events of the 1990s and distinguishing between the effects of policies and of shocks.

I. Context and Motivation

It is now widely accepted that economic growth is at least a necessary condition for sustainable poverty reduction (Kanbur 2001). In Africa growth rates have historically been low. Reviewing the literature of cross-country growth studies, Collier and Gunning (1999) concluded that the explanations are to be found in macroeconomic policies (notably economic policy volatility and a lack of openness to international trade), in microeconomic and sectoral policies (in particular the urban bias of fiscal and agricultural policies, which eroded social capital and undermined the provision of public services while keeping rural producers largely subsistence producers), and in geography (the land-locked, tropical character of many countries). Despite these important insights from cross-country analysis, the reasons for "Africa's growth and poverty paradox" (Easterly and Levine 1997) remain open to debate. Does the Dollar and Kraay (2000, p. 27) view that "anyone who cares about the poor should favor the growth-enhancing policies of good rule of law, fiscal discipline, and openness to international trade" apply to Africa in the 1990s? Or was the growth path the reforms induced characterized by increasing inequality, denying benefits to the poorest (Stewart 1995; Mkandawire and Soludo 1999; Forsyth 2000)?

Collier and Gunning (1999) find that the messages from household-level analysis are different from those from the cross-country literature. Among the factors explaining poverty at the household level, "disease and climate feature most prominently, and these are largely omitted in the aggregate analysis" (p. 83). They hint that these growth-retarding risks might explain the results for the Africa dummy variable in growth regressions.[2] Ravallion (2001) also calls for a more microeconomic approach to the analysis of policies, growth, and poverty. Using household survey data for a sample of 50 developing economies and 120 spells of poverty change, he estimates that on average the growth elasticity of headcount poverty is –2.5.[3] But this average masks a great deal of variation across countries—variations related to income inequality levels and trends. Going beyond the averages highlights "the importance of more micro, country-specific research on the factors determining why some poor people are able to take up the opportunities afforded by an expanding economy—and so add to

2. In support of this hypothesis Guillaumont and others (1999) find that economic, political, and natural volatility are important factors explaining the poor growth performance of African economies.

3. This elasticity is based on the growth in mean household income or consumption. If growth is taken to be per capita private consumption from the national accounts, the elasticity is approximately –2.

its expansion—while others are not" (Ravallion 2001, p. 1813) This article applies Ravallion's advice to Africa in the 1990s.

II. Living Standards During the 1990s

A broader view of well-being looks not only at the evolution of private consumption but also at primary school enrollment, child malnutrition, and child mortality (table 1). This sets the scene for the analysis of the selected countries.

The first and obvious point is that living standards are very low in these countries. By the close of the decade no country enjoyed annual per capita consumption higher than US$500, and in Ethiopia it was just $86. All countries fell far short of universal primary enrollment, and in some (for example, Ethiopia) primary enrollment was unacceptably low. Malnutrition was a serious problem, especially in Ethiopia and Madagascar. In Ethiopia about two-thirds of children exhibited signs of stunting or long-term malnutrition. Even in Ghana, Mauritania, and Zimbabwe, where malnutrition indicators were better, there is evidence of stunting in about a quarter of the population under five years of age. Perhaps the most poignant indicator of the very low welfare levels is the incidence of child deaths. Under-age-five mortality exceeded 100 (per 1,000) in all countries. In Zambia about one in five children failed to survive to the fifth birthday. Too many African children are dying needlessly.

Second, there are differences in how these indicators changed over time. In Ethiopia, Ghana, Mauritania, and Uganda economic living standards appear to have improved. But in Madagascar average real consumption remained more or less unchanged, and in Nigeria, Zambia, and Zimbabwe it fell sharply. Similarly, improvements in primary school enrollment in Ethiopia, Ghana, Mauritania, and Uganda contrast with unsatisfactory outcomes in Zambia. Mauritania experienced sharp reductions in long-term malnutrition but limited progress elsewhere. In several countries the long-term downward trend in child mortality appears to have continued through the decade, but not in Nigeria, Zambia, and Zimbabwe, a result probably related to the AIDS epidemic in these countries (among other factors). Also, the latest Uganda Demographic and Health Survey (Uganda Bureau of Statistics 2001) suggests that child mortality has been unchanged (and has possibly even increased) since 1995 despite the economic gains.

Third, the trends in the indicators are generally consistent with each other, though there are important exceptions. In the four countries experiencing economic growth (Ethiopia, Ghana, Mauritania, and Uganda), the trends in human development indicators match the improvement in economic well-being, albeit to different degrees. But in the four countries experiencing stagnation and decline (Madagascar, Nigeria, Zambia, and Zimbabwe), the signals are noisier. In some the education indicator improved despite the stagnation or decline in economic living standards (Madagascar, Nigeria, and Zimbabwe). Child malnutrition improved in Zimbabwe during an episode of deteriorating economic

TABLE 1. Evolving Living Standards in Eight African Countries during the 1990s (%, unless otherwise specified)

Growth outcome/ Country and years	Real private consumption per capita (constant 1995 US$)			Net primary school enrollment rates[b]			Under-five child malnutrition[c]			Under-five child mortality[d]		
	Year 1	Year 2	AGR[a]	Year 1	Year 2	Change (%)	Year 1	Year 2	Change (%)	Year 1	Year 2	Change
Positive growth												
Ethiopia												
1994 and 1997	80	86	2.6	19	25	+6	66	55	−11	190	175	−15
Ghana												
1992 and 1999	275	304	2.0	70	82	+12	26	26	0	119	104	−15
Mauritania												
1987 and 1995	297	361	3.6	28	41	+13	48	23	−25	—	149	—
Uganda												
1992 and 1997	211	258	4.7	68	86	+18	43	39	−4	165	162	−3
Stagnation or decline												
Madagascar												
1993 and 1999	223	222	0.0	48	64	+16	50	49	−1	170	149	−21
Nigeria												
1992 and 1996	206	173	−3.4	94	98	+4	38	—	—	136	147	11
Zambia												
1991 and 1998	362	266	−6.6	73	666	−7	40	42	+2	192	202	10
Zimbabwe												
1991 and 1996	595	439	−5.2	83	886	+3	30	23	−7	80	90	10

Note: Many of these estimates are subject to margins of error, given the difficulties of obtaining reliable and accurate estimates (especially of under-five mortality and malnutrition). —, not available.

[a]AGR = Annual growth rate. Calculated based on least squares method, which is less sensitive to the choice of base and terminal period.

[b]Percentage of school-age children enrolled in primary school as a fraction of the total number of children in that age group. First year figure for Ethiopia refers to 1996. Figures for Nigeria reflect gross enrollment rates in 1994 and 1996 and are obtained from *World Development Indicators* (World Bank various years).

[c]Percentage of children stunted (z-score of height for age of less than −2); reference periods are close to those in the country column.

[d]Reference periods are close to those in the country column; values presented, per 1000 live births.

Source: World Bank data and country studies for Ethiopia, Ghana, Madagascar, Mauritania, Nigeria, Uganda, Zambia, and Zimbabwe (see text for specific studies cited). Private consumption growth is from national accounts data.

circumstance. Such outcomes (and Uganda's experience after 1995) serve as a reminder that focusing on only one dimension of well-being can be misleading when tracking poverty dynamics (World Bank 2000). It is quite possible through public intervention to raise enrollment rates, for example, even if consumption per capita is not rising.

Trends in Income Inequality

Were episodes of growth in Africa in the 1990s associated with changes in income inequality?[4] Gini coefficients, a popular measure of inequality, are used to describe how income inequality evolved in the sample of countries (table 2). All underlying "welfare measures" are based on total household real expenditures.[5] The studies were designed to enable comparisons over time within a country, though differences in survey designs require caution in making comparisons across countries. Nonetheless, the differences in the degree of income inequality in the sample of countries are striking. At one extreme Zimbabwe has a highly unequal distribution (a Gini ratio of over 0.6),[6] reflecting unequal land distribution, a result in part of its colonial history. Highly unequal income distributions are also observed in Nigeria and Zambia and are likely to be related to the importance of mineral exports. Income distributions in Ghana and Uganda are far more egalitarian.

Over the periods considered, there was very little change in overall income inequality in these countries. Reforms and growth have clearly not led to a significant deterioration in consumption inequality, as some would hold (Forsyth 2000)—although Ethiopia is an important exception.[7] Nevertheless, these aggregate measures of inequality can be misleading. They may mask a great deal of distributional change, an issue reviewed in section IV.

Trends in Income Poverty

If growth episodes were not associated with significant changes in inequality, were they associated with poverty reduction? As with the inequality measures, real household consumption per adult equivalent (or in some cases per capita) is

4. All the empirical measures of income are based on expenditures, so the terms *income* and *consumption* are used interchangeably here.

5. For most countries researchers normalized expenditures on the number of "equivalent adults" in the household. In Ethiopia, Nigeria, and Madagascar the welfare measure is real household expenditure per capita. Details of the data and methods used by the researchers in computing the welfare measure (including the specifics of the adult equivalent scales) are provided in the appendix.

6. Intuitively, the Gini index of a population represents the expected income difference between two randomly selected individuals or households. Table 1 gives Zimbabwe's real average per capita consumption in 1996 as $439. The corresponding Gini index is 0.64 (table 2). Thus in 1996 the per capita consumption of any two randomly selected Zimbabweans differed on average by $281(= 0.64 * $439)—a clear indication of high inequality given that average per capita consumption is only $439.

7. A similar picture emerges when using the Theil inequality measures. Yet it must be noted that stability in the Gini ratio over relatively short periods is not uncommon (Deininger and Squire 1996).

TABLE 2. Consumption Inequality during the 1990s in Eight African Countries, Gini Coefficients

Country, years, and area	Year 1	Year 2	Change
Ethiopia[a]			
1994 and 1997			
Rural	0.39	0.43	0.04
Urban	0.40	0.45	0.05
National	0.39	0.44	0.05
Ghana			
1992 and 1999			
Rural	0.33	0.33	0.00
Urban	0.34	0.31	−0.03
National	0.37	0.37	0.00
Madagascar			
1993 and 1999			
Rural	0.42	0.36	−0.06
Urban	0.41	0.38	−0.03
National	0.43	0.38	−0.05
Mauritania			
1987 and 1995			
Rural	0.43	0.37	−0.06
Urban	0.40	0.36	−0.04
National	0.43	0.39	−0.04
Nigeria			
1992 and 1996			
Rural	0.51	0.44	−0.07
Urban	0.51	0.51	0.00
National	0.51	0.47	−0.04
Uganda			
1992 and 2000			
Rural	0.33	0.32	−0.01
Urban	0.39	0.40	0.01
National	0.36	0.38	0.02
Zambia			
1991 and 1996			
Rural	0.62	0.48	−0.13
Urban	0.47	0.44	−0.03
National	0.59	0.50	−0.09
Zimbabwe			
1991 and 1996			
Rural	0.58	0.57	−0.01
Urban	0.60	0.59	−0.01
National	0.68	0.64	−0.04

Note: Real per capita expenditures for Ethiopia, Nigeria, and Madagascar. Otherwise, real expenditures per adult equivalent.

[a]Purposely sampled villages and urban centers; not nationally representative.

Source: World Bank data and country studies for Ethiopia, Ghana, Madagascar, Mauritania, Nigeria, Uganda, Zambia, and Zimbabwe (see text for specific studies cited).

TABLE 3. Consumption Poverty in Eight African Countries during the 1990s (%)

	Poverty headcount			Poverty severity index		
Country and years	Year 1	Year 2	Change (%)[b]	Year 1	Year 2	Change (%)[b]
Ethiopia[a]						
1994 and 1997	41	35	−14*	8	6	−27*
Ghana						
1992 and 1999	51	39	−24	9	7	−22
Madagascar						
1993 and 1997	70	73	5	17	19	12
1997 and 1999	73	71	−3	19	19	0
Mauritania						
1987 and 1995	58	35	−40*	17	6	−65*
Nigeria						
1985 and 1992	46	43	−7**	8	9	13**
1992 and 1996	43	66	53*	9	17	89*
Uganda						
1992 and 1997	56	44	−21*	10	6	−40*
1997 and 2000	44	35	−20*	6	5	−16*
Zambia						
1991 and 1996	70	80	14**	31	31	1**
1996 and 1998	80	76	−5**	31	26	−16**
Zimbabwe						
1991 and 1996	26	35	35	4	5	25

*Statistically significant at the 5% level.

**Not statistically significant at the 5% level.

[a]Purposely sampled villages and urban centers; not nationally representative.

[b]For changes without a symbol, no statistical assessment is reported by the country study authors.

Source: World Bank data and country studies for Ethiopia, Ghana, Madagascar, Mauritania, Nigeria, Uganda, Zambia, and Zimbabwe (see text for specific studies cited).

taken as the central economic welfare measure (table 3). Poverty lines in all but one case (Mauritania) are derived from a food consumption basket estimated to yield a minimum caloric intake with adjustments made for essential nonfood consumption. These poverty lines are typically much higher than the purchasing power parity (PPP) $1 a day poverty line. The average poverty incidence in 24 spells of poverty change in African countries analyzed by Ravallion (2001) was 31 percent (based on the PPP $1 a day line). This compares with the (unweighted) average poverty headcount ratio of 58 percent in this sample, indicating higher poverty lines. Because of differences in survey design and in the specifics of how the welfare measure and poverty lines are derived, the data in table 3 are not comparable across countries. But the research has been designed to ensure comparable estimates over time. Where available, table 3 reports whether the observed changes are statistically significant.

The poverty measures reported here are derived from the familiar class of poverty indices after Foster and others (1984). They include the poverty headcount and the severity of poverty index, which is sensitive to the distribution of income among the poor, particularly to changes in the living standards of the poorest of the poor. The data suggest the following.

- Most countries have to deal with "mass" poverty. Over 70 percent of the population was estimated to be poor in Madagascar and Zambia toward the end of the 1990s and 66 percent in Nigeria in 1996.
- There is no uniform trend. Although the incidence of consumption poverty declined substantially in several countries (Ethiopia, Ghana, Mauritania, and Uganda), it rose sharply in Nigeria and Zimbabwe. Poverty has fluctuated in Madagascar and Zambia, remaining more or less unchanged in Madagascar and increasing marginally Zambia (but not significantly).
- Where the incidence of poverty has declined, the data suggest that the poorest sections of the population have also benefited. This is evidenced by the downward trend in the severity of poverty index, which in several instances fell more than the poverty headcount.

Poverty, Inequality, and Economic Growth

In some cases these changes in poverty occurred in a context of economic decline (Nigeria and Zimbabwe, and Madagascar and Zambia during the earlier periods). In others they accompanied overall economic progress (Ethiopia, Ghana, Mauritania, and Uganda). To shed more light on the relation between poverty, inequality, and growth, the poverty incidence was decomposed into changes arising from changes in mean consumption, with the distribution of consumption kept constant, and changes arising from changing consumption distribution, with mean consumption kept constant (table 4). The poverty measure decomposed is the elasticity of headcount poverty with respect to changes in mean household expenditure.[8]

In most countries changes in poverty incidence were due predominantly to changes in mean expenditure (table 4). But the results of this exercise serve as a caution against overgeneralizing for Africa. Growth experience in Uganda (in which reduction in inequality bolstered the effects of rising mean consumption) contrasts with that in Ethiopia, where inequality increased and dampened the poverty reducing impact of growth. Where there has been recession, mean and redistribution effects typically have opposite signs, and the redistribution effect substantially mitigates the poverty increasing impact of lower mean incomes (in

8. This is defined as the proportionate change in headcount poverty divided by the proportionate change in mean per capita household expenditure. For details about the method used, see Kakwani and Subbarao (1993) and Kakwani and Pernia (2000).

TABLE 4. Relative Importance of Mean and Redistribution in the Evolution of Poverty Incidence in Eight African Countries during the 1990s

Country and period	Change in mean per capita expenditure (%)	Change in poverty headcount (%)	Poverty elasticity relative to change in mean expenditure	Poverty elasticity explained by changes in	
				Mean	Distribution
Ethiopia 1994–97[a]	24.8	−13.8	−0.56	−1.09	0.53
Ghana 1992–99	24.9	−23.6	−0.95	−0.99	0.04
Madagascar 1993–97	−17.5	4.7	−0.27	−0.77	0.50
Madagascar 1997–99	0.6	−2.7	−4.50	−0.78	−3.72
Madagascar 1993–99	−17.0	1.9	−0.11	−0.73	0.62
Mauritania 1987–95	49.5	−40.4	−0.82	−0.75	−0.07
Nigeria 1992–96	−41.1	53.6	−1.30	−1.3	0.02
Uganda 1992–97	17.1	−20.7	−1.21	−1.06	−0.15
Zambia 1991–96	−25.7	14.9	−0.58	−0.58	0.00
Zambia 1996–98	13.2	−4.9	−0.37	−0.44	0.07

Note: Decompositions are based on Kakwani and Subbarao (1993). This method is an exact decomposition with no residual or interactive term.

[a]Purposely sampled villages and urban centers; not nationally representative.

Source: Authors' computations based on country studies for Ethiopia, Ghana, Madagascar, Mauritania, Nigeria, Uganda, Zambia, and Zimbabwe (see text for specific studies cited).

Madagascar, Nigeria, and Zimbabwe). Better-off groups clearly bear a heavier burden of income losses during periods of economic decline in Africa.[9]

To assess further the extent to which these episodes of growth and recession are pro-poor and following the work of Kakwani and Pernia (2000), an index of pro-poor growth (ϕ) is defined:

$$\phi = \eta/\eta_g$$

where η is the observed elasticity of headcount poverty with respect to changes in mean expenditure and η_g is the elasticity of headcount poverty assuming that the distribution of income did not change during the period. Growth can be considered pro-poor if $\phi > 1$. When mean household expenditures are declining, $\phi = \eta_g/\eta$, so that a recession would also be considered pro-poor if $\phi > 1$. Note that ϕ is defined for a specific poverty line, and its value may be sensitive to the poverty line selected.[10] Comparisons of estimates of ϕ for the eight African countries with estimates for four Asian countries suggest that growth and recession episodes have tended to be more pro-poor in Africa than in Asia (table 5).

Taking the 11 spells of poverty change in the sample of African countries reported in table 4 yields a growth elasticity of poverty incidence of just -0.89.[11] This relatively low elasticity cannot be due to increasing inequality—the Gini ratios were stable. It clearly reflects the depth of poverty—large numbers of people are subsisting well below the poverty line (and poverty lines are set well above modal consumption). The growth elasticity of the poverty severity index, at -1.28 (with a standard error of 0.21), is higher, indicating that growth has improved the economic well-being of the poorest, though not enough to take many of them out of poverty. (It should be pointed out that this assessment takes the poverty lines as given.)

III. Growth and Systematic Changes in Income Distribution: A Macro Perspective

These changes in inequality and poverty have occurred during an era of economic policy reform, institutional change, and profound internal and external

9. The tendency for income inequality to narrow as higher-income groups bear the brunt of economic recession was also noted by Grootaert (1996) in analyzing poverty changes in Côte d'Ivoire in the 1980s.

10. For that reason, recent work has analyzed the impact of growth on the whole distribution (Ravallion and Chen 2003). Here, however, the Kakwani-Pernia method is retained to compare the experience of the sample of African countries with their findings for East Asia.

11. This is simply the slope coefficient in the regression of the proportionate change in headcount poverty on the proportionate change in the survey mean. The standard error on the slope coefficient is 0.11. The regression line runs almost through the origin, a reflection of the fact that income inequality has been stable over this period. The historical elasticities observed for this sample of African countries are significantly lower than that estimated by Ravallion (2001) as typical of low-income countries (-2.5). Given the different poverty lines used (Ravallion uses the much lower benchmark of PPP $1 a day) and the different method of computation, his estimates are not comparable with those in this study.

TABLE 5. Pro-Poor Growth Indices (ϕ) in Selected African and Asian Countries during the 1990s

Africa	Pro-poor growth index	Asia	Pro-poor growth index
Growth episodes			
Ethiopia 1994–97	0.51	Thailand, 1992–96	0.61
Ghana 1992–99	0.96	Lao PDR, 1993–98	0.21
Mauritania 1987–95	1.10	Korea, 1990–96	1.03
Uganda 1992–96	1.14		
Zambia 1996–98	0.87		
Recession and stagnation episodes			
Madagascar, 1993–97	2.85	Thailand, 1996–98	0.73
Nigeria, 1992–96	1.02	Korea, 1997–98	0.84
Zambia, 1991–96	0.97		
Zimbabwe, 1991–96	1.81		

Note: For details on the method see text. Asian country estimates are simple means across years within the subperiods shown.

Source: Table 4; Kakwani and Pernia (2000).

shocks, including droughts, disease, and fluctuating commodity prices. These events have effects at all levels—they influence the growth rate of the economy at large, they affect the functioning of markets and of government, they change village and community life, and they impinge directly on households and individuals. Understanding how the changes have influenced poverty outcomes thus calls for knowledge at both the macro (economy-wide) and micro (household and individual) levels. An assessment of how macro changes (in economic and institutional environments) have affected poverty outcomes provides the context for reviewing the microeconomic evidence (from recent studies) linking poverty outcomes to policies and shocks.

Macroeconomic Reforms and Poverty Trends

This review of the relationship between macroeconomic policy reforms and income poverty elaborates on and updates the analysis of Demery and Squire (1996), who examined the empirical association between improvements in macroeconomic balances and poverty reduction using data for the late 1980s and the early 1990s. The better comparable household data now available (including panel data) together with another decade of economic reform in many countries make this a good time to revisit this issue.[12]

Following Bouton and others (1994), a macroeconomic policy index or score is calculated based on changes in three key elements of sound macroeconomic policy: fiscal, monetary, and exchange rate policies. The overall macro-policy

12. The data used in many previous assessments were often of doubtful quality, and given the lags involved in implementing the reforms, the 1990s might be a more appropriate decade to examine the growth path induced by economic policy reforms in Africa (Collier and Gunning 1999, p. 101).

TABLE 6. Changes in Macroeconomic Policy Scores, Selected Countries and Periods

Country	Period of change	Fiscal policy	Monetary policy	Exchange rate policy	Average score	
					Unweighted	Weighted
Côte d'Ivoire						
	1985–88	−3	1	−1	−1.0	−1.5
Ethiopia						
	1989–95	−1	0.5	2.5	0.7	1.0
	1994–97	2	1.5	2.5	2.0	2.2
Ghana						
	1988–92	−1	1.5	2	0.8	0.8
	1992–99	0	−0.5	0.5	0	0.2
Madagascar						
	1993–97	0	−0.5	0	−0.2	−0.1
	1997–99	1	1.0	0	0.7	0.5
Mauritania						
	1987–95	3	0.5	2.5	2.0	2.4
Nigeria						
	1985–92	1	−0.5	3	1.2	1.9
	1992–96	1	−1	−2.5	−0.8	−1.0
Uganda						
	1992–97	2	1.5	−0.5	1.0	0.7
	1997–2000	0	0.5	0.5	0.3	0.3
Zambia						
	1991–96	1	2	2	1.7	1.6
	1996–98	1	1	−1	0.3	0.0
Zimbabwe						
	1991–96	−1	−0.5	1.5	0	0.3

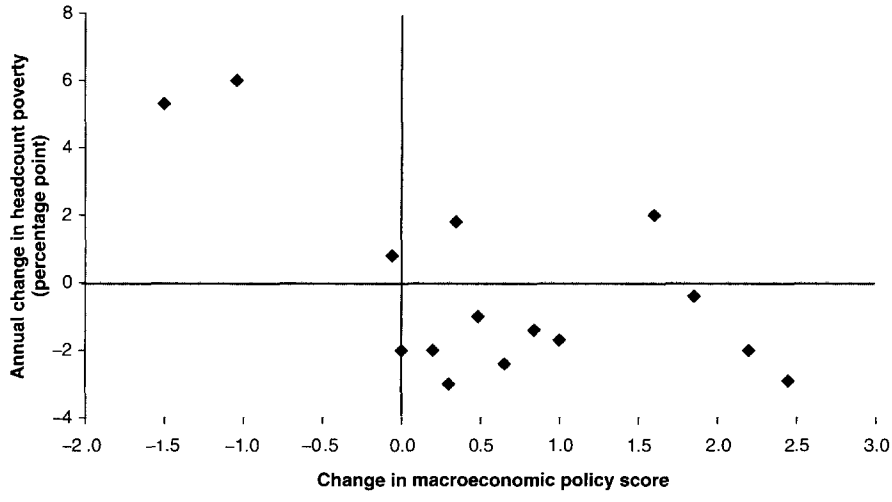
Source: Demery and Squire (1996); authors' computations based on World Bank data (see appendix table 1).

score is a weighted average of these components, the weights being derived from international cross-section growth regressions. These scores are computed for the three-year period prior to each survey, and changes in the index are then compared. The index is computed so that increases in the score (either lower negative values or higher positive values) indicate an improvement in economic policy (table 6). (Details of the changes in the different policy instrument indicators and of the computations of the macro-policy score are in appendix table 1.)

Given weaknesses in the underlying survey data, two countries included in the original Demery and Squire analysis (Tanzania and Kenya) were dropped from this analysis.[13] The data for Ethiopia, Ghana, and Nigeria, which were

13. Comparative household survey data were not available for these countries. Although Côte d'Ivoire is not among the sample of countries in this study (the survey data are not comparable for assessing trends in the 1990s), it is retained for this exercise because it was included in the Demery and Squire assessment.

FIGURE 1. Macroeconomic Policy Reform and Poverty Changes



included in the Demery and Squire analysis, were updated by adding the latest available survey. Finally Madagascar, Mauritania, Uganda, Zambia, and Zimbabwe were added for a total coverage of 15 episodes of change in 9 countries.

Most countries experienced improvements in their macroeconomic policy indicators—those for the second period (the three years prior to the second survey) generally being better than those for the earlier period (the three years prior to the first survey). But there were only marginal improvements in Ghana (1992–99) and Zimbabwe (1991–96) and no change in Zambia (1996–98). Macroeconomic destabilization is observed in two countries—Côte d'Ivoire during the 1980s and Nigeria in the 1990s.

Setting these macroeconomic trends against the trends in poverty reduction suggests that countries achieving improvements in their macroeconomic balances typically have not experienced (in the aggregate at least) increases in consumption poverty—rather the reverse (figure 1).[14] Nine of the 15 episodes of change for which there are data indicate both macroeconomic policy improvement and subsequent poverty reduction. In the two cases in which macroeconomic balances substantially deteriorated, poverty increased sharply. Only 2 of the 15 observations (Zimbabwe 1991–96 and Zambia 1991–96) are in the "wrong" quadrant in figure 1 (improved macroeconomic policy and increased poverty).[15]

14. Ali (1998) gets quite different results, with reforms being associated with increasing poverty. This is probably due to the different poverty data sets he uses (derived from International Fund for Agricultural Development data) and possibly the different time period covered (1985–95).

15. The Pearson correlation coefficient is −0.62 and statistically significant at the 5 percent level.

The association between the macro-policy stance and poverty reduction does not necessarily imply any causative or direct behavioral link.[16] Rather, it highlights the close interactions between macroeconomic policies and economic well-being at the household level. An important feature missing from this analysis is any measure of policy persistence and consistency.[17] Collier and Gunning (1999) argue that the slow investment response to the reforms derives in part from a fear of policy reversals. Countries with a longer history of consistent policies (Ethiopia, Ghana, Mauritania, and Uganda in the sample) are more likely to experience growth and poverty reduction dividends from the reforms.[18] The macro analysis is partial in another respect—the changes in the macroeconomic accounts took place alongside other reforms—mostly structural (trade liberalization, agricultural marketing reforms, privatization)—and within changing institutional environments. Both the sectoral reforms and the institutional environment are certain to be important, as is suggested by the quite similar poverty reductions that occurred in some of the countries despite different changes in their macroeconomic indicators (see bottom right quadrant in figure 1).

Institutional Change and Poverty Trends

There is an accumulation of convincing empirical evidence on the importance of political stability and good governance for growth and poverty reduction (Alesina and Perotti 1994; Knack and Anderson 1995; Collier 1999; Collier and Gunning 1999; World Bank 2000). Although the construction and consolidation of good indicators of political stability and good governance remain a work in progress, the composite political risk index of the International Country Risk Guide (ICRG) and its subsets have frequently been used by researchers to examine the effect of governance and institutional quality on growth and poverty. The composite index consists of 12 components covering different aspects of political stability (for example, government stability, internal conflict, external conflict), governance, and institutional quality (for example, corruption, democratic accountability, bureaucracy quality).[19] The key advantage of the ICRG index is its broad coverage across countries and over time (1985 to today).

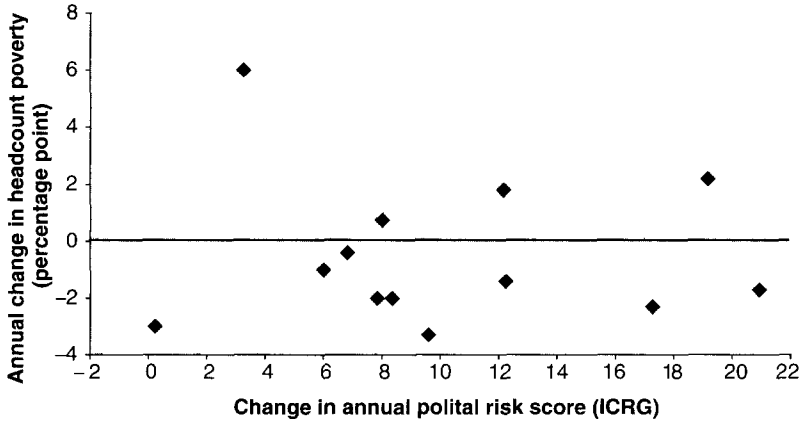
16. Both poverty changes and macro-policy scores might be favorably affected by a third factor, movements in the terms of trade, for example.

17. Although this study tracks and compares three-year averages of the macro-policy stance, it does so only for the two periods prescribed by the available household surveys.

18. These countries are described by Collier and Gunning (1999, p. 102) as "providing at least modest levels of social order, macroeconomic order and resource allocation."

19. The components of the ICRG political risk index (maximum scores in parentheses) are government stability (12), socioeconomic conditions (12), investment profile (12), internal conflict (12), external conflict (12), corruption (12), military in politics (6), religion in politics (6), law and order (6), ethnic tensions (6), democratic accountability (6), and bureaucracy quality (4). The maximum score is 100, with a score below 49.9 indicating very high political risk; 50 to 59.9 a high risk, 60 to 69.9 a moderate risk, 70 to 79.9 a low risk, and 80 or more a very low risk. The same range of scores on an individual component implies the same degree of risk for that component. Evaluations of the different aspects of the index are provided by a private consultancy. For a detailed description of the ICRG rating system, visit www.icrgonline.com/icrgMethods.asp.

FIGURE 2. Change in Political Stability and Governance and Poverty Trends



The political risk score improved during all episodes of poverty change covered by the Poverty Dynamics in Africa study. In all, 13 episodes of institutional change were examined. In Ethiopia (1989–95) the improvement followed largely from reduced risk of internal and external conflict following peace agreements with Eritrea. Better overall governance (as captured by the corruption, law and order, democratic accountability, and bureaucratic quality indices), as well as greater government stability and reduced risk of internal conflicts, induced progress in institutional quality in Ghana (1992–99) and Uganda (1992–97). Increased government stability was responsible for the change in Madagascar (1993–97). In Zimbabwe (1991–96) improvement followed from reduced risk of an external conflict, reflecting the peace process in neighboring Mozambique and the end of the cold war. Greater external security is also an important factor in explaining the large improvement in the political risk score in Zambia, in addition to substantial progress in internal political stability and security (law and order) following the peaceful handover of power by Kenneth Kaunda in 1991 after 27 years of autocratic rule.

Plotting the changes in the average annual political risk scores of the survey years of the sample countries against annual changes in the observed poverty incidence suggests that improvements in political stability and governance are generally associated with reductions in poverty, though experiences vary across countries (figure 2).[20] In 9 of the 13 episodes improvements were accompanied

20. Using two-year averages of the survey year and the year prior to the survey to account for lags in the effect of institutional change on poverty does not change the results. The findings are also robust to the use of a subset of the political risk indicator focusing on political stability (government stability, internal conflict, external conflict) and governance (corruption, law and order, democratic accountability and bureaucratic quality).

by poverty reduction. In one episode there was a modest increase in poverty (Madagascar during 1993–97), whereas in the three other cases (Nigeria, Zambia during 1991–96, and Zimbabwe) the poverty increases were more pronounced. In Nigeria the recorded improvement in the institutional environment was marginal (3.3 percentage points), likely swamped by the adverse effects of the macroeconomic deterioration.

The other exceptions, Zambia and Zimbabwe, are more of a puzzle, as the macroeconomic balances also improved during the 1991–96 episodes of poverty increase (if only moderately in Zimbabwe). So where did things go wrong? The answer cannot be provided here, but the very high level of initial inequality in both Zambia and Zimbabwe was a particularly serious challenge for growth and poverty reduction during the decade. Also, both countries experienced severe droughts in 1994, followed by mediocre rainfall in 1995, leaving households under considerable stress. This compares with above-average rainfall in 1990, the year preceding the first survey. (The Zimbabwe episode of poverty increase and the role of rainfall shocks is discussed in further detail in the section on shocks.)

Although these measures of political stability and the quality of governance are admittedly crude, the findings support the general observation that increased political stability and improved governance go hand in hand with poverty reduction. Nevertheless, many difficult questions remain that fall beyond the scope of this article. Which components of institutional change (for example, political, economic, civil rights, or social stability) have had the most significant impact? What is the direction of causality and the channels through which institutional improvements and poverty reduction affect each other (Aron 2000)?

A Multivariate Approach

The treatment so far has been partial and descriptive and does not take into account how factors other than macroeconomic policy and institutional quality might have affected economic growth and poverty during the periods covered by the household survey. A multivariate approach would control for these other factors, but the sample of countries is too small for that. However, the cross-country regression results of an African Economics Research Consortium study by O'Connell and Ndulu (2000) can be used to provide at least an international cross-section interpretation of the growth turnarounds that the study countries experienced.[21] Their model is used to speculate about the factors that have influenced the changes in economic growth (and therefore in consumption

21. Most cross-country models are ill equipped to deal with such questions because of their very long-run concern (Collier and Gunning 1999). In contrast, the O'Connell and Ndulu specification has a medium-run (half-decade) empirical orientation.

poverty) over the periods defined by the survey years.[22] This approach cannot trace precisely the situations of the first and second surveys, but it can be used to examine growth dynamics during the period broadly surrounding the surveys.

This exercise was conducted for four of the study countries—two experiencing growth surges and poverty reduction (Ghana and Uganda) and two facing declining growth and worsening poverty (Côte d'Ivoire and Nigeria).[23] The periods are selected to best describe the growth dynamics that influenced the observed poverty changes. The model correctly predicts the change in per capita gross domestic product (GDP) growth in all countries, indicating an acceleration in growth over the relevant periods in Ghana and Uganda and a deceleration in Côte d'Ivoire and Nigeria (table 7). The predictions are reasonably close for Ghana and Uganda but tend to underestimate the growth reversals in Côte d'Ivoire and Nigeria. The most striking result is the confirmation of the importance of macroeconomic policy for growth and poverty reduction. In both Ghana and Uganda changes in macroeconomic policy explain most of the growth turnaround. Of the macroeconomic variables the reduction in the black market exchange rate premium has the greatest impact. The deterioration in macroeconomic policy is important in interpreting Nigeria's growth reversal (1.2 percentage points of the decline in the growth rate is attributable to that alone).

Changes in the institutional setting (approximated in this model by the political stability variable—the mean number of revolutions, strikes, and assassinations) also play an independent role, raising the growth potential in Ghana and Uganda by 0.2 percentage points and harming growth prospects in Nigeria. Finally, luck is also crucial. In Nigeria and especially Côte d'Ivoire deterioration in the terms of trade and the slowdown in growth in trading partner countries harmed growth prospects. The decade effect (basically a time trend allowed in the O'Connell and Ndulu model that is assumed to be common to all countries) also indicates an increasingly harsh external environment during the periods covered for Côte d'Ivoire and Nigeria. In sum, interpreting the growth changes

22. Of the various regression models O'Connell and Ndulu specify to analyze the "deep determinants" of growth, this study takes the following, $G = 15.35 - 1.765\ Y_0 + 0.089\ LE - 0.052\ ADR + 0.728\ LFG + 0.004\ TT + 0.540\ TPG - 0.912 LL$ dummy $- 0.975\ POL + - 0.004\ CPI - 0.007\ BMP - 0.113\ GOV \pm$ decade dummies, where Y_0 is the log of initial GDP, LE is life expectancy at birth, ADR is the age dependency ratio, LFG is the difference between the growth of the population of working age and the total population, TT is the terms of trade, TPG is the trade-weighted GDP growth of trading partners, LL is a dummy variable for land-locked countries, POL is a measure of political stability (mean number of revolutions, strikes and assassinations), CPI is the mean annual inflation rate, BMP is the black market exchange rate premium, and GOV is government consumption (less spending on education and defense) to GDP ratio. Per capita GDP growth in Côte d'Ivoire, Ghana, Nigeria, and Uganda are obtained by applying the country regressors to this relationship.

23. The choice of countries covers cases of both economic recovery and stagnation. But it was also constrained by lack of data and by the poor performance of the model (particularly for Madagascar). In the cases, selected the model correctly predicts the growth turnarounds that took place during the years surrounding the survey observations.

TABLE 7. Decomposition of Changes in Predicted per Capita GDP Growth Based on O'Connell and Ndulu (2000), Selected Countries and Periods

Variable	Ghana 1975–84 to 1985–97	Uganda 1980–89 to 1990–97	Côte d'Ivoire 1970–79 to 1980–89	Nigeria 1970–79 to 1980–97
Initial year GDP	0.20	−0.06	0.05	−0.29
Demographic factors[a]	0.42	−0.28	0.10	0.21
Geography	0.00	0.00	0.00	0.00
Shocks	0.89	−0.26	−0.70	−0.46
Political stability	0.19	0.21	0.00	−0.12
Inflation	0.16	0.36	0.02	−0.05
Black-market exchange rate premium	3.02	1.85	0.00	−0.69
Government spending	−0.42	−1.15	−0.26	−0.42
Sum (macro policy effect)	2.76	1.06	−0.24	−1.16
Decade effect	−0.10	0.33	−0.99	−0.88
Predicted change	4.36	1.00	−1.78	−2.70
Actual change	5.17	1.91	−7.18	−5.16

Note: Decomposition values are given as percentage points.

[a]The demographic regressors are life expectancy at birth, age dependency ratio, and difference between the growth of the population of working age and the total population (see note 23).

Source: Authors' calculation based on O'Connell and Ndulu (2000) and World Bank data.

in these countries using a multivariate approach confirms the independent effects of improved macroeconomic policy and also points to the importance of institutions for growth and poverty reduction.

IV. Growth and Systematic Changes in Income Distribution: A Micro Perspective

The evidence from the African experience covered in this study indicates that growth and recession—except in Ethiopia—have been reasonably pro-poor. Yet this conclusion is true only in an aggregate sense. Further decomposition of national inequality and poverty measures—by geographic location and socio-economic group—indicates that the aggregate statistics often mask a wide variety of experiences. Some groups and regions gained disproportionately from the new opportunities following economic reforms, but others lost out or even became impoverished. Similarly, overall Gini coefficients often appear stable over time despite substantial churning within and across geographic regions. From microeconometric evidence the positive association between improved macro environments and poverty reduction seems conditioned by other factors, such as market liberalization, location and infrastructure, private endowments, and the occurrence of shocks.

Two of the case studies (Dercon 2002; Deininger and Okidi 2001) use panel data to explain the determinants of consumption and income growth and poverty dynamics. Dercon (2002) uses panel data from six rural communities in Ethiopia covering 1989–95.[24] The change in household real consumption per adult is explained through a reduced-form regression model with an Oaxaca-Blinder type decomposition. Changes in consumption and poverty are explained by changes in endowments over time and by changes in returns to those endowments. The main regressors are changes in real crop producer prices (closely related to macroeconomic and agricultural reforms), location, access to roads, private endowments (land, labor, and education), and two shock variables, rainfall and ill health. Deininger and Okidi (2001) analyze changes in consumption and income observed for a panel of about 1,200 Ugandan households during 1992–2000. They regress household-level changes in consumption and income against variables representing the change in relative producer prices of coffee, largely due to the liberalization of coffee markets (Townsend 1999), access to infrastructure, initial endowments of physical and human capital, and initial health status of households and their social capital.

These microeconometric analyses of panel data suggest that the following factors are influential in determining the relationship between economic growth and poverty reduction.

- Most poor rural households stand to benefit directly from liberalization measures, as well as from increased political stability and better governance. The gains can be substantial. Insofar as liberalization measures increase producer prices, rural producers gain, and to the extent that food marketing margins decline, rural consumers will benefit as well. Nonetheless, some gain more than others, depending on the product and consumption mix of the household.
- A household's location is key in conditioning the extent to which it benefits from liberalization measures. Whether the household had access to infrastructure and urban markets is an immensely important factor governing the growth in household income. It explains about half of household consumption growth and poverty reduction in Ethiopia during 1989–95 and is also quantitatively important for growth in Ugandan household income. So connectedness to markets, as captured by access to infrastructure (especially roads, but also electricity) and distance to urban centers, is likely to be a major factor in determining how growth in any country transmits its benefits to the population.
- The potential for economic growth and poverty reduction further depends on a household's private endowments. Households with larger private

24. Because the study is not nationally representative, the results cannot be generalized to Ethiopia as a whole. Nonetheless, the methodology and the empirical findings provide important insights to the linkages between economic policy, growth, and poverty reduction.

TABLE 8. Poverty Incidence by Rural Activity, Ghana and Uganda in the 1990s

	Uganda				Ghana			
Crop type	Population share (2000)	1992	2000	% reduction	Population share (1998)	1992	1998	% reduction
Food crop	45.9	63.3	45.7	−27.8	43.9	68.1	59.4	−12.8
Cash crop	21.3	62.7	29.7	−52.6	6.3	64.0	38.7	−39.5

Source: Appleton (2001); Coulombe and McKay (2003).

endowments—be it more and better qualified labor or land—tend to be not only less poor but also better placed to profit from new opportunities generated by liberalization and institutional change.

- It is vital to separate the effects of shocks (such as ill health or drought) from the effects of other factors when assessing poverty trends or the impact of policy changes.

Liberalization

The panel studies show that policy-induced changes in relative prices directly benefited poor households. The experience of Ghana in West Africa echoes these East African findings. Ghana experienced sharp poverty reductions among cash (export) crop producers during the 1990s, a result of more favorable world cocoa prices and an increase in cocoa production. A comparison of trends in poverty among crop producers in rural Ghana and Uganda shows that about two-fifths of the population in both countries are food-producing farmers, and about two-thirds of them were poor in the early 1990s (table 8). In both countries poverty fell among food producers, but the decline was greater among export crop producers. So, although most of the rural poor appear to have benefited from growth, those producing export crops benefited most. That a much larger share of the population grows cash crops in Uganda (21 percent) than in Ghana (6 percent) may explain the larger drop in poverty among food crop producers in Uganda. Reviewing the evidence on agricultural reforms in Sub-Saharan Africa, Kherallah and others (2002) arrive at a similar conclusion—export-crop producers seem to have benefited more than food crop producers. What needs to be better understood is the transmission mechanism that led to economic gains for households not producing for export.[25]

Location

The panel analysis of Ethiopian and Ugandan households provides strong empirical evidence that location and geography are important in determining

25. Coulombe and McKay (2003) find that remittances were an important source of income growth in rural Ghana during this time.

TABLE 9. Headcount Poverty Trends in Rural and Urban Areas of Eight African Countries during the 1990s (% unless otherwise specified)

	Rural				Urban		
Country and period	Population share, year 1	Year 1	Year 2	% point change	Year 1	Year 2	% point change
Ethiopia							
1994–97	84	42	35	−6	37	35	−2
Ghana							
1992–98	67	64	49	−15	28	19	−9
Madagascar							
1993–99	81	74	77	2	50	52	2
Mauritania							
1987–95	56	68	48	−20	45	17	−28
Nigeria							
1992–96	62	46	69	23	37	58	21
Uganda							
1992–97	88	59	48	−11	28	16	−12
Zambia							
1991–96	62	88	90	2	47	62	15
1996–98	62	90	86	−4	62	59	−3
Zimbabwe							
1991–96	63	36	48	12	3	8	5

Sources: World Bank data and country studies for Ethiopia, Ghana, Madagascar, Mauritania, Nigeria, Uganda, Zambia, and Zimbabwe (see text for specific studies cited).

how growth influences income distribution. In some countries the decline in poverty is observed in both rural and urban areas (Ghana, Mauritania, Ethiopia, Uganda; table 9). In others the change is confined mainly to urban areas (Zambia in 1991–96). It is clear from the case studies that within both rural and urban sectors, poverty changes have varied considerably by geographic location. Some geographic areas have not benefited as much as others from growth, and some have even lost ground during the period of recovery. The different experience in the evolution of poverty seems closely related to the extent to which the region or village is integrated within the overall economy.

The experiences of Ghana and Madagascar are illustrative. Although poverty fell in Ghana between 1992 and 1999, not all regions benefited. Living standards declined in urban areas other than Accra and in rural areas in the north. Recent studies by Badiane and Shively (1998) and Abdulai (2000) conclude that markets (more specifically the maize market) in the remoter northern region are not well integrated with the economy at large, likely impeding transmission of the benefits of growth to the region. Remoteness is also important in understanding geographic differences in poverty outcomes in Madagascar. Paternostro and others (2001) report an association between the degree of remoteness and the likelihood of being in poverty. They also show that although overall rural poverty remained largely unchanged during 1997 and 1999, the most

remote households experienced increased poverty—in contrast to the least remote, whose poverty indicators improved.

Private Endowments

The experiences in Ethiopia and Uganda demonstrated that better-endowed households, particularly more educated households and those with more (fertile) land, were not only less likely to be poor but also more likely to benefit from favorable changes in the macro environment. The importance of education for poverty reduction is echoed by microeconometric evidence from Ghana, Madagascar, and Zimbabwe. In Ghana and Madagascar real consumption levels increase with educational attainment. The returns to education across education levels increased between the first and second survey years (Coulombe and McKay 2003; Paternostro and others 2001). In Zimbabwe a sharper increase in poverty following the economic decline was prevented by previous investments in schooling that increased the educational attainment of the population in the 1990s (Alwang and others 2002). That incomes fell and poverty increased despite household efforts to invest in human capital, assets, and migration was due to a reduction in the rates of return to these assets.

In Madagascar consumption levels are higher for households that possess land, but only for holdings greater than 0.1 ha per capita. Returns increase with the size of the plots. Consumption levels deteriorated between 1993 and 1999 for households with less than 0.4 hectare per capita, while they improved for households with more land. As a result, poverty incidence fell by 2 percentage points among those with more land and rose by 0.82 percentage points among those with less than 0.4 ha per capita. Paternostro and others (2001) surmise that this difference follows from more extensive land use by small farmers forced by demographic pressures to expand their fields into less productive and more fragile areas.

Shocks

Poverty estimates provide a snapshot of the standard of living at a certain point in time, reflecting both policy reforms and temporary external shocks, such as droughts. In evaluating the evolution of poverty, it is thus important to control for the effect of external shocks on comparative poverty figures. Controlling for all other factors, the Ethiopian panel analysis estimated that household income growth was reduced by about a fifth because of rainfall shortage (Dercon 2002). Rainfall variations were also an important influence on household income growth in Madagascar and Zimbabwe.

Simulations of poverty change in Madagascar show that 75 percent of the predicted change in household economic well-being and poverty incidence can be traced to the effects of drought (Paternostro and others 2001). The capacity of households to insure themselves against covariate shocks is clearly extremely limited in many parts of Africa.

That poverty increased sharply in Zimbabwe during the 1990s is without question (Alwang and others 2002). Less clear is whether poverty increased because of the droughts that afflicted the country in 1991/92 and again in 1994/95 or because of the Economic Structural Adjustment Program (launched in 1991) that was being implemented at the same time. Alwang and others apply nonparametric methods to confirm that the drought did lead to an increase in poverty during the early 1990s. But they also show that the drought alone cannot explain the worsening in economic well-being. The deteriorating economic environment sharply undermined the returns to both human and physical capital, reducing incomes and increasing poverty.

V. Conclusion

In the African countries studied, episodes of growth reduced poverty, at least in the aggregate. Countries whose macroeconomic balances and institutional quality improved also saw a decline in poverty. But there are two serious qualifications. First, experiences varied enormously. Some countries enjoyed a decade of sustained growth, and others had to cope with crisis and decline. Of the eight countries covered, four experienced declines in poverty (Ethiopia, Ghana, Mauritania, and Uganda), two faced sharp increases (Nigeria and Zimbabwe), and two (Madagascar and Zambia) no exhibited discernible trend. This variety counsels caution in applying the empirical findings reported here to Africa as a whole.

The second qualification derives from the need to go beyond the averages. Although it is true that overall income distributions (evidenced by the Gini ratio) did not change during episodes of growth in these African countries (except Ethiopia) and that such growth (or recession) can be characterized as pro-poor in this aggregate sense, this conclusion can be misleading. A variety of experience lies beneath the aggregate numbers. Neglect of this reality by policymakers—and sometimes by academics—has often impeded a constructive and fruitful dialogue with civil society about appropriate poverty-reducing policies (Kanbur 2001). Focusing on income poverty, this review of the evidence finds systematic changes in income distribution and poverty in the countries covered and identifies some of the main contours of these distribution changes. These results highlight four key policy messages: the importance of economic reform and political stability for poverty reduction, the role of geographic location (especially remoteness) in conditioning how the benefits of growth are distributed, the importance of private endowments (especially education and land) for the ability of households to take advantage of new opportunities and for consequent poverty outcomes, and the need to account for shocks in understanding distributional outcomes and poverty changes over time.

The "emerging picture" described by Demery and Squire (1996) is further supported by the new, better data (reflecting also a longer time perspective than previous work). Improvements in macroeconomic balances are generally

associated with reductions in poverty in the countries examined. There is also an emerging micro-picture of the impact of market liberalization on consumption poverty. The analysis of household panel data by Dercon (2002) for Ethiopia and Deininger and Okidi (2001) for Uganda provides the most systematic and empirically convincing cases that policy-induced changes in relative prices can have poverty-reducing effects. Micro evidence from Ghana provides some corroboration from West Africa.

The second policy message is the need for a geographic perspective on poverty. Although various rounds of poverty assessments have established that the incidence of poverty varies considerably across regions of a country, this recent work on poverty dynamics has shown that some regions, by virtue of their sheer remoteness, have been left behind as growth has picked up. Households with limited access to markets and public services have not benefited from growth during the 1990s. The provision of public goods (notably infrastructure services—from the Ethiopian case especially roads, and from the Ugandan case, electricity) is crucial to help poor households benefit from growth. Third, education emerges as a key private endowment enabling households to escape poverty. Its importance for poverty reduction is brought out in the case studies—for both rural and urban areas—with the marginal returns to education typically increasing with the level of educational attainment.

Finally, the empirical evidence reviewed here underscores the importance of risk in everyday life in Africa. Two risks featured in the analysis are the impact of rainfall variations and ill health. Dercon (2002) estimates that poverty reduction in the sample of Ethiopian rural communities would have been 18 percentage points greater had households been protected from the effects of ill health and rainfall shortages. The importance of weather shocks for changes in poverty was also underscored by the findings from Madagascar and Zimbabwe. Deininger and Okidi (2001) find that in Uganda ill health in 1992 noticeably increased the probability of being in poverty eight years later. Policies to help the poor manage their risks should be an integral part of poverty reduction strategies in the region.

27. Appendix. Data Sources and Constructs

This appendix presents basic information on the survey instruments, the construction of the welfare measures used in the studies covered in this article, and the construction of the macroeconomic policy scores. Readers are referred to the individual studies for a detailed exposition of survey design and methodological choices made in building each welfare measure. Some of the salient features are reported here.

Surveys Used

For all countries investigated except Ethiopia data were obtained from household surveys collected by local statistical authorities. For Ethiopia the authors (Dercon 2002; Bigsten and others 2003) constructed a purposively selected

APPENDIX TABLE 1. Computations of Macroeconomic Policy Scores

		Fiscal policy					Monetary policy					Exchange rate policy						
		Change in overall fiscal balance excluding grants (% GDP)		Change in government revenue (% GDP)		Change in fiscal policy	Change in seigniorage		Change in inflation		Change in monetary policy	Change in real effective exchange rate		Change in black market premium		Change in exchange rate policy	Overall macroeconomic policy	
Country and Period		%	Score	%	Score	Score[a]	%	Score	%	Score	Score[b]	%	Score	%	Score	Score[c]	Average score	Weighted[d] average score
Côte d'Ivoire	1985–88	−11.6	−2	−5.2	−1	−3.0	−2.7	2	2.9	0	1.0	21.8	−2	−2.1	0	−1.0	−1.0	−1.5
Ethiopia	1989–95	0.3	0	−6.9	−1	−1.0	−0.7	1	2.9	0	0.5	−55.8	3	−56.0	2	2.5	0.7	1.0
	1994–97	2.5	1	6.1	1	2.0	−3.8	2	−6.8	1	1.5	−23.9	2	−126.6	3	2.5	2.0	2.2
Ghana	1988–92	−2.3	−1	0.1	0	−1.0	−1.2	1	−10.1	2	1.5	−23.5	2	−51.0	2	2	0.8	0.8
	1992–98	−5.0	−1	4.5	1	0.0	0.4	0	7.9	−1	−0.5	−11.9	1	−4.4	0	0.5	0.0	0.2
Madagascar	1993–97	0.8	0	−0.5	0	0.0	−1.1	1	13.7	−2	−0.5	−0.2	0	−8.0	0	0.0	−0.2	−0.1
	1997–99	1.7	1	1.6	0	1.0	−0.2	0	−16.8	2	1.0	2.3	0	1.4	0	0.0	0.7	0.5
Mauritania	1987–95	9.2	3	0.6	0	3.0	−1.3	1	−1.1	0	0.5	−35.8	3	−84.2	2	2.5	2.0	2.4
Nigeria	1985–92	0.4	0	12.3	1	1.0	1.0	−1	−1.8	0	−0.5	−518.9	3	−260.4	3	3.0	1.2	1.9
	1992–96	3.7	2	−4.6	−1	1.0	−1.2	1	31.4	−3	−1.0	53.3	−2	249.1	−3	−2.5	−0.8	−1.0
Uganda	1992–97	2.9	1	3.3	1	2.0	−1.8	1	−30.3	2	1.5	10.2	−2	−23.0	1	−0.5	1.0	0.7
	1997–00	−0.5	0	0.3	0	0.0	0.4	0	−4.5	1	0.5	−8.9	1	−5.8	0	0.5	0.3	0.3
Zambia	1991–96	1.7	1	1.0	0	1.0	−2.7	2	−63.2	2	2.0	−8.4	1	−350.7	3	2.0	1.7	1.6
	1996–98	2.2	1	−0.7	0	1.0	−0.9	1	−9.2	1	1.0	11.0	−2	1.7	0	−1	0.3	0.0
Zimbabwe	1991–96	−2.6	−1	−0.8	0	−1.0	1.6	−1	4.2	0	−0.5	−8.0	1	−40.6	2	1.5	0.0	0.3

Scoring criteria (from World Bank 1994)						
−3				≥31.0		≥51.0
−2	≤−5.0		2.0 to 3.9	10 to 30.9	≥10.0	16 to 50
−1	−4.9 to −2.0	≤−4.0	1.0 to 1.9	5.0 to 9.9	5.0 to 9.9	5.0 to 15
0	−1.9 to 0.9	−3.9 to 3.0	−0.5 to 0.9	−2.4 to 4.9	−2.0 to 4.9	−9.0 to 4.0
1	1.0 to 2.9	≥3.1	−2.0 to −0.6	−9.9 to −2.5	−2.1 to −14.9	−29 to −10
2	3.0 to 4.9		−3.0 to −2.1	−49.0 to −10.0	−15.0 to −30.9	−99 to −30
3	≥5.0				≤−31.0	≤−100

[a]Sum of scores for change in overall fiscal balance and change in revenue.

[b]Average of scores for change in seigniorage and change in inflation.

[c]Average of scores for change in the real effective exchange rate and change in the black market premium.

[d]Weights derived from cross-sectional growth regressions.

Source: Authors' computations based on World Bank (1994) and World Bank data.

panel germane to their research objectives (in a collaborative venture with the University of Addis Ababa). Details on the survey designs, time frames, coverage, and sample size can be found online at http://poverty.worldbank.org/files/14946_afr_growth.pdf. With the exception of Ethiopia, where one panel covered 6 villages and the other 15, all the surveys are nationally representative.

There are, however, a few omissions worth noting. The surveys used in Mauritania did not sample the nomadic population (about 30 percent). In Uganda the 1996–97 Monitoring Survey and the 1999–2000 Uganda National Household Survey did not cover four districts (6.9 percent of the population according to the 1991 census) for security reasons. These districts reported relatively low levels of mean consumption in the 1992–93 Household Income Survey. These omissions affect only the representativeness of the sample. The analysts adjusted the sample to ensure over-time comparability.

Welfare Measurement

The welfare indicator commonly chosen was total household expenditure (see data online at http://poverty.worldbank.org/files/14946_afr_growth.pdf). The studies used adult equivalence scales to account for household composition except in Ethiopia, Madagascar, and Nigeria, where total expenditures are computed on a per capita basis (see data online at http://poverty.worldbank.org/files/14946_afr_growth.pdf). In general, the principle guiding the selection of items included in each expenditure measure was ensuring comparability over time. Therefore, only items common across surveys and for which questions were asked in a similar fashion were retained. In some cases, like Zimbabwe, surveys maintained a common design over time, thus allowing for a wide coverage of household consumption expenditures including use of services, consumption values from assets owned and imputed values from gifts remittances, and transfers received. In some other instances, as in Madagascar, this approach led to a more restrictive coverage. The survey design changed over time so that items such as consumption of own livestock, gifts, remittances, in-kind payments, and own consumption from nonfood enterprises were omitted to ensure proper comparability.[26]

In all cases, consumption included items requiring imputations, such as imputed rent from owner-occupation and imputed income from the consumption of food produced by the household. Methods of computing these imputations differed. Although consumption of own-produced food was included in all expenditure measures, for Ghana households themselves estimated the value of such items, whereas for Uganda the imputation was obtained using median unit values from household food purchases (market prices). Such procedures were

26. Such items accounted for only 4.3 percent of total expenditure in 1993. The highest number of items omitted was in Ethiopia (1989–95 rural panel), where only food expenditures were included.

applied consistently to data sets to preserve over-time comparability. But differences among countries would counsel caution in comparing results across countries.

Adjustments were also made to account for differences in prices, across both time and space. All studies computed total expenditures in real terms, using official price series to express values in base year prices (see data online at http://poverty.worldbank.org/files/14946_afr_growth.pdf). Moreover, with the exception of Zambia and Zimbabwe, regional and rural-urban price differences are also taken into account in constructing the consumption measure.

Macroeconomic Policy Scores

A macroeconomic policy index or score is calculated for the countries covered in this article. It is based on changes in three key components: fiscal, monetary, and exchange rate policies. The overall score is a weighted average of these components, computed for the three-year period prior to each survey (the weights being derived from international cross-section growth regressions). The index is computed so that increases in the score (either lower negative values or higher positive values) indicate an improvement in economic policy. Appendix table 1 provides details of the changes in the different policy instrument indicators and of the computations of the macro-policy score.

References

Abdulai, Awudu. 2000. "Spatial Price Transmission and Asymmetry in the Ghanaian Maize Market." *Journal of Development Economics* 63:327–49.

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

Ali, A. A. G. 1998. "Structural Adjustment and Poverty in Sub-Saharan Africa: 1985–1995." In T. Mkandawire and C. C. Soludo, eds., *African Perspectives on Structural Adjustment*, vol. 2. Ottawa: International Development Research Centre.

Alwang, J., and L. Ersado. 1999. "Changes in Poverty in Zimbabwe: 1990–1996." World Bank, Africa Region, Poverty Reduction and Social Development, Washington, D.C.

Alwang, J., B. Mills, and N. Taruvinga. 2002. *Why Has Poverty Increased in Zimbabwe?* Poverty Dynamics in Africa Series. Washington D.C.: World Bank.

Appleton, S. 2001. "Poverty in Uganda, 1999/2000: Preliminary Estimates from the UNHS." University of Nottingham, Department of Economics.

Appleton, S., T. Emwanu, J. Kagugube, and J. Muwonge. 1999. "Changes in Poverty in Uganda, 1992–1997." World Bank, Africa Region, Poverty Reduction and Social Development. World Bank, Washington, D.C.

Aron, Janine. 2000. "Growth and Institutions: A Review of the Evidence." *World Bank Research Observer* 15(1):99–135.

Badiane, O., and G. Shively. 1998. "Spatial Integration, Transport Costs, and the Response of Local Prices to Policy Changes in Ghana." *Journal of Development Economics* 56:411–31.

Bigsten, Arne, Bereket Kebede, Abebe Shimeles, and Mekonnen Taddesse. 2003. "Growth and Poverty Reduction in Ethiopia: Evidence from Household Panel Surveys." *World Development* 31(1):87–106.

Bouton, Lawrence, Christine Jones, and Miguel Kiguel. 1994. "Macroeconomic Reform in Africa: 'Adjustment in Africa' Revisited." Policy Research Working Paper 1394. World Bank, Washington, D.C.

Canagarajah, S., J. Ngwafon, and F. Okunmadewa. 2000. "Nigeria's Poverty: Past, Present, and Future." World Bank, Africa Region, Nigeria Country Department, Washington D.C.

Collier, Paul. 1999. "On the Economic Consequences of Civil War." *Oxford Economic Papers* 51: 168–83.

Collier, Paul, and Jan Willem Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37(1):64–111.

Coulombe, H., and A. McKay. 2003. "Changes in Poverty and Inequality in a Slow Growth Environment: Ghana over the 1990s." World Bank, Africa Region, Poverty Reduction and Economic Management, Washington, D.C.

Deininger, K., and J. Okidi. 2001. "Growth and Poverty Reduction in Uganda, 1992–2000: Panel Data Evidence." World Bank, Washington, D.C., and Economic Research Council, Kampala.

Deininger, K., and L. Squire. 1996. "A New Data Set Measuring Income Inequality." *World Bank Economic Review* 10(3):565–91.

Demery, Lionel, and Lyn Squire. 1996. "Macroeconomic Adjustment and Poverty in Africa: An Emerging Picture." *World Bank Research Observer* 11:39–59.

Dercon, Stefan. 2000. "Changes in Poverty and Social Indicators in Ethiopia in the 1990s. (At Last) Some Good News From Ethiopia." World Bank, Africa Region, Poverty Reduction and Social Development, Washington, D.C.

———. 2002. *The Impact of Economic Reforms on Households in Rural Ethiopia: A Study from 1989 to 1995*. Poverty Dynamics in Africa Series. Washington D.C.: World Bank.

Dollar, David, and Aart Kraay. 2000. "Growth *IS* Good for the Poor." World Bank, Development Economics, Development Research Group, Washington, D.C.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4):1203–50.

Forsyth, Justin. 2000. Letter. *The Economist* June 20.

Foster, J., J. Greer, and E. Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52(3):761–66.

Grootaert, Christiaan. 1996. *Analyzing Poverty and Policy Reform*. Aldershot, U.K.: Avebury.

Guillaumont, P., S. Guillaumont, and J. F. Brun. 1999. "How Instability Lowers African Growth" *Journal of African Economies* 8(17):87–107.

Kakwani, Nanak, and Kalanidhi Subbarao. 1993. "Poverty in India: 1973–87." In Michael Lipton and J. van der Gaag, eds., *Including the Poor*. Washington, D.C.: World Bank.

Kakwani, Nanak, and Ernesto M. Pernia. 2000. "What Is Pro-poor Growth?" *Asian Development Review* 18(1).

Kanbur, Ravi. 2001. "Economic Policy, Distribution and Poverty: The Nature of Disagreements." *World Development* 29(6):1083–94.

Kherallah, Mylène, Christopher Delgado, Eleni Gabre-Madhin, Nicolas Minot, and Michael Johnson. 2002. *Reforming Agricultural Markets in Africa*. Baltimore, Md.: John Hopkins University Press.

Knack, Stephen, and Gary Anderson. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7:207–27.

McCulloch, N., B. Baulch, and M. Cherel-Robson. 2000. "Growth, Inequality and Poverty in Mauritania, 1987–1996." World Bank, Africa Region, Poverty Reduction and Social Development, Washington, D.C.

———. 2001. *Poverty, Inequality and Growth in Zambia during the 1990s*. WIDER Discussion Paper 2001/123. Helsinki: United Nations University.

Mkandawire, Thandika, and Charles C. Soludo. 1999. *Our Continent, Our Future: African Perspectives on Structural Adjustment*. Asmara, Eritrea: Africa World Press.

O'Connell, Stephen A., and Benno J. Ndulu. 2000. "Africa's Growth Experience: A Focus on Sources of Growth." African Economic Research Consortium, Nairobi.

Paternostro, S., J. Razafindravonona, and D. Stifel. 2001. "Changes in Poverty in Madagascar: 1993–1999." Africa Region Working Paper Series 19. World Bank, Washington, D.C.

Ravallion, Martin. 2001. "Growth, Inequality and Poverty: Looking Beyond Averages." *World Development* 29(11):1803–15.

Ravallion, Martin, and Shaohua Chen. 2003. "Measuring Pro-poor Growth." *Economic Letters* 78(1):93–99.

Stewart, Frances. 1995. *Adjustment and Poverty: Options and Choices*. London: Routledge.

Townsend, Robert. 1999. "Agricultural Incentives in Sub-Saharan Africa: Policy Changes." World Bank Technical Paper 444. World Bank, Washington, D.C.

Uganda Bureau of Statistics. 2001. "Uganda Demographic and Health Survey 2000 – Preliminary Report." Entebbe.

White, Howard, and Tony Killick. 2001. *African Poverty at the Millennium: Causes, Complexities and Challenges*. Strategic Partnership with Africa. Washington, D.C.: World Bank.

World Bank. 1994. *Adjustment in Africa: Reforms, Results and the Road Ahead*. New York: Oxford University Press.

———. 2000. *World Development Report 2000/01: Attacking Poverty*. New York: Oxford University Press.

———. Various years. *World Development Indicators*. Washington, D.C.

Risk Sharing in Labor Markets

Arne Bigsten, Paul Collier, Stefan Dercon, Marcel Fafchamps, Bernard Gauthier, Jan Willem Gunning, Abena Oduro, Remco Oostendorp, Cathy Pattillo, Mans Söderbom, Francis Teal, and Albert Zeufack

Empirical work in labor economics has focused on rent sharing as an explanation for the observed correlation between wages and profitability. The alternative explanation of risk sharing between workers and employers has not been tested. Using a unique panel data set for four African countries, we find strong evidence of risk sharing. Workers in effect offer insurance to employers: when firms are hit by temporary shocks, the effect on profits is cushioned by risk sharing with workers. Rent sharing is a symptom of an inefficient labor market. Risk sharing, by contrast, can be seen as an efficient response to missing markets. Our evidence suggests that risk sharing accounts for a substantial part of the observed effect of shocks on wages.

If labor markets were perfectly competitive, all firms would pay the same wages for the same type of work. But the empirical evidence is very different, showing large wage differentials across firms. Do such differentials indicate the need for policies promoting competition? In many situations, "yes" may well be the appropriate answer. But in this article we argue that policymakers may have to look elsewhere. Rather than a labor market imperfection, the problem may be the limited ability of firms to deal with risk. If this is what explains wage differentials, policymakers should aim to improve the functioning of credit and insurance markets.

Evidence for developed economies typically shows strong correlations across firms of profits and wages per worker. Labor economics has suggested various explanations for this. Theoretical articles have focused on two such explanations:

Arne Bigsten is with Göteborg University; Paul Collier, Stefan Dercon, Marcel Fafchamps, Mans Söderbom, and Francis Teal are with the University of Oxford; Bernard Gauthier is with École des Hautes Études Commerciales, Montreal; Jan Willem Gunning and Remco Oostendorp are with the Free University, Amsterdam; Abena Oduro is with the Center for Policy Analysis, Accra, Ghana; Cathy Pattillo is with the International Monetary Fund; and Albert Zeufack is with the World Bank. Correspondence should go to Jan Gunning at jgunning@feweb.vu.nl and Remco Oostendorp at roostendorp@feweb.vu.nl. The authors gratefully acknowledge the very useful comments of three referees and François Bourguignon. The authors form the Industrial Surveys in Africa (ISA) group, which uses multicountry panel data sets to analyze the microeconomics of industrial performance in Africa. This article uses data collected by the Regional Programme on Enterprise Development, organized by the World Bank and funded by the Belgian, British, Canadian, Dutch, French, and Swedish governments. The support of these governments for workshops is gratefully acknowledged.

THE WORLD BANK ECONOMIC REVIEW, VOL. 17, NO. 3 349–366

DOI: 10.1093/wber/lhg026

rent sharing and risk sharing between employers and workers (see, for example, Oswald 1995; Malcomson 1999). Under rent sharing workers demand higher wages from employers in more profitable firms, and thanks to wage bargaining workers in those firms receive higher wages. In risk-sharing models, changes in value added are shared between the firm and the workers when the firm is exposed to shocks, so that profits and wages change in the same direction.[1] In effect, workers provide insurance to their employers, and they are compensated for the resulting volatility of their incomes by a wage premium. Most risk-sharing models assume that there are no financial markets, so employers cannot respond to risk by using credit or by buying insurance; there are no alternatives to risk sharing with employees.

Both the rent- and the risk-sharing literature are concerned with profit–wage correlations (which may occur in both the short-run and the long-run equilibriums), but the focus of the rent-sharing literature is primarily on long-run or steady-state correlations. The reason is that rent sharing is inherently long run but that a long-run correlation is difficult to reconcile with nonrent-sharing explanations (Blanchflower and others 1996). Conversely, risk sharing is, by definition, primarily concerned with short-run wage–profit correlations.

Empirical work on labor markets in developed economies has usually tested the first of these models, rent sharing. Evidence of rent sharing between workers and employers has also been found for labor markets in developing economies, as by Teal (1996) for Ghana and Velenchik (1997) for Zimbabwe. Mazumdar and Mazaheri (1998) investigate rent sharing and efficiency wage explanations for interfirm wage differences for Ghana, Kenya, Zambia, and Zimbabwe. To the best of our knowledge, there has been no empirical work on risk sharing.

The aim of this article is to test for risk sharing, using panel data for manufacturing firms in Africa. That the data are for African firms is fortunate. Industrial firms in Africa are exposed to very high risks, reflecting demand shocks, price volatility, unreliable infrastructure, and poor contract enforcement (Fafchamps 1996; Collier and Gunning 1999; Bigsten and others 2000b). In most African economies, the financial markets, particularly those for insurance, are poorly developed. This conjunction of high risk and weak financial markets suggests that if risk sharing through implicit labor contracts is to be found anywhere, it is in African manufacturing.

There is a growing literature on the responses that African manufacturing firms have adopted to the risks they face. These include holding inventories and liquid assets (studied for Zimbabwe by Fafchamps and others 2000) and risk sharing between firms in networks (Bigsten and others 1999). Within networks delays in payments or delivery are usually accepted from partners, because network members can observe whether the delays reflect bad luck (making

1. Risk sharing models apply the optimal contract framework of Baily (1974) and Azariadis (1975). In those early studies, workers are risk-averse but the firm is risk-neutral. In this framework employers insulate workers from shocks: risks faced by the firm are not shared with workers. In the optimal contract the wage is independent of the state of nature. In later models, however, both agents are risk-averse, and this leads to risk sharing.

contract flexibility possibly efficient) or bad faith (opportunism). Although evidence of such risk sharing *between* firms is accumulating, there has been no applied work on risk sharing between workers and employers *within* firms. There is, however, interesting anecdotal evidence of risk sharing. For example, in Zimbabwe a Christmas bonus is common (in the form of a "13th month" payment), but it is not paid if the firm has fared poorly during the year.

It is important to be able to distinguish between these two explanations for a correlation between wages and profits because they have different policy implications. Rent sharing clearly is a symptom of an inefficient labor market. But risk sharing also implies labor market imperfections—for if workers could costlessly move from one firm to another, it would be impossible to enforce the risk sharing contract. If a firm experienced a negative shock, its workers would leave the firm rather than accept the wage cut implied by risk sharing. However, given the transaction costs that make the risk-sharing contract enforceable, risk sharing may well be second best in the sense that it makes efficient use of the labor market to compensate for a missing insurance market. In this case, policymakers should aim to improve the functioning of financial markets (credit and insurance) rather than labor markets.

In this article we investigate whether risk sharing occurs in African manufacturing firms by using a unique multicountry panel data set. We use the data to construct measures of sustained and temporary changes in firms' value added, so that we can test for both rent sharing and risk sharing: permanent changes will raise wages through rent sharing, temporary changes through risk sharing. The demarcation between permanent and temporary shocks (and thus between rent sharing and risk sharing) is, of course, debatable. Rather than use a single operational classification of shocks as temporary or permanent, we use several.

The data set is unique in containing data both on firms and on a sample of workers in these firms. This makes it possible to estimate earnings functions while controlling for characteristics of the workers and for characteristics of the firm employing them. Here we use data for four countries: Cameroon, Ghana, Kenya, and Zimbabwe. We find evidence of risk sharing under each of the several concepts of permanent shocks.

We set out the model in the next section and discuss the survey data in section II. The econometric results are in section III. In section IV we consider whether the effect of risk sharing on wages is quantitatively important. Section V discusses the policy implications of our findings.

I. Modeling Rent Sharing and Risk Sharing

We use a hybrid model that encompasses the two canonical models for rent and risk sharing (Blanchflower and others 1996).[2] Consider a bargaining problem

2. Our hybrid model is based on models 1 and 3 in the theoretical appendix to Blanchflower and others (1996).

involving two agents: workers and an employer. Both agents are risk averse. The employer maximizes the expected value of $v(\pi)$ where v is a strictly concave function of profits π. Workers have an outside wage of $\overline{w}$ when bargaining breaks down or when they are unemployed—and w if they are employed by the firm. With n workers employed and the labor force equal to 1 (by choice of units) the average utility of workers (in the absence of a breakdown of bargaining) is $nu(w) + (1-n)u(\overline{w})$ where $u(y)$ is a strictly concave function of income y (w or $\overline{w}$). Production is subject to short-run or temporary (s^s) and long-run or permanent (s^l) shocks, with both types multiplicative. The firm's value added f depends only on employment, n. So profits are given by $\pi = s^s . s^l . f(n) - wn$. The probability density function (PDF) of s^s is $g(s^s)$. The bargaining problem can now be written as a Nash problem:

$$\max \phi \log W + (1-\phi) \log E,$$

where ϕ measures the relative bargaining power of workers and

$$W = \int [nu(w) + (1-n)u(\overline{w})]g(s^s)ds^s - u(\overline{w}) = \int n[u(w) - u(\overline{w})]g(s^s)ds^s$$

$$E = \int v(\pi)g(s^s)ds^s.$$

Risk is shared between the two agents if the wage rate, employment, or both are state-contingent, depending on the realization of s^s : $n = n(s^s)$ or $w = w(s^s)$.[3] If w is state contingent, the first-order condition for w is:

$$\frac{\phi}{W}u'(w) - \frac{1-\phi}{E}v'(\pi) = 0 \quad \text{for all } s^s, s^l$$

hence

$$\frac{\phi}{1-\phi}\frac{E}{W} = \frac{v'(\pi)}{u'(w)}. \tag{1}$$

Shocks will affect both profits and wage rates so that (1) traces out the relationship that will be observed between w and π over time as the firm is exposed to a series of s^s shocks. Alternatively, it is the relationship that will be found between wages and profits in a cross-section of firms.

We consider the two types of shocks in turn. In the case of a s^s shock we assume that agents will recognize it as a draw from the given PDF $g(s^s)$ so that in (1) neither E nor W is affected. The first-order condition can thus be written as $u'(w) = \lambda v'(\pi)$, with λ a constant. It follows that

$$\frac{dw}{d\pi} = \lambda\frac{v''}{u''} = \left(\frac{w}{\pi}\right)\frac{-\pi v''/v'}{-wu''/u'}$$

3. In Oswald (1995) and Blanchflower and others (1996) both n and w are state-contingent.

so that

$$\frac{\pi}{w}\frac{dw}{d\pi} = \frac{\Omega}{r} \tag{2}$$

where Ω and r denote the relative risk aversion of the employer and of the workers, respectively. Equation (2) is a well-known result in the literature (for example, equation [25] in Blanchflower and others 1996). It indicates that the elasticity of wages to profits depends only on the risk aversion of the two agents. Clearly, if the employer is risk-neutral so that $\Omega = 0$, wages will be unaffected by shocks (as in Baily 1974 or Azariadis 1975). Because the result is derived from the first-order condition for the wage rate, it holds irrespective of whether only w or both w and n are adjusted in response to shocks.

Now consider s^l shocks. In this case the model reduces to[4]:

$$\max \phi \log n[u(w) - u(\overline{w})] + (1 - \phi) \log v(\pi),$$

with first-order conditions for the wage rate:

$$\phi \frac{u'(w)}{nu(w) - nu(\overline{w})} - (1 - \phi)\frac{v'(\pi)}{v(\pi)} = 0$$

and for employment:

$$\phi \frac{u(w) - u(\overline{w})}{nu(w) - nu(\overline{w})} + (1 - \phi)\frac{v'(\pi)(f'(n) - w)}{v(\pi)} = 0.$$

From the first of these:

$$\frac{u(w) - u(\overline{w})}{u'(w)} = \frac{\phi}{1 - \phi}\frac{v(\pi)}{nv'(\pi)}.$$

Assuming constant relative risk aversion, we can substitute $v/v' = \pi/(1 - \Omega)$. Taking a first-order approximation to $u(\overline{w})$, this gives

$$w = \overline{w} + \frac{\phi}{1 - \phi}\frac{1}{1 - \Omega}\frac{\pi}{n}.$$

Note that this implies (for a constant degree of relative risk aversion and constant ϕ) that the elasticity of wages to profits per head is equal to $(w - \overline{w})/w$. If the outside wage ($\overline{w}$) is low relative to the wage paid by the firm (w), this elasticity is likely to be larger than Ω/r, the elasticity for risk sharing.[5] To this

4. To achieve comparability with the s^s shock model, we changed the risk-sharing model of Blanchflower and others (1996) by replacing π in the maximand by $v(\pi)$. Note that the difference between average utility $nu(w) + (1 - n)u(\overline{w})$ and the utility of the outside wage $\overline{w}$ reduces to $n[u(w) - u(\overline{w})]$.

5. For example, if the degree of relative risk aversion is 1 for the employer and 2 for workers, the statement is true if and only if $\overline{w} < 2w$.

extent we expect permanent shocks to have a stronger effect on wages than temporary shocks.

The theoretical model treats labor as homogeneous and firms as identical. Our empirical strategy is to estimate earnings functions in which wages are determined by characteristics of the firm and the workers. We include among the regressors two measures of shocks: permanent and temporary. If risk sharing occurs, temporary shocks should be a significant determinant of wage levels. We investigate whether this is so and whether the evidence is sensitive to the way the demarcation line between permanent and temporary shocks is drawn. We cannot observe when implicit contracts are drawn up or renegotiated. So we implicitly assume that the same contract remains in force during the period covered by our data (three rounds of panel data, collected at one-year intervals).

II. Data

Our data are from the Regional Programme on Enterprise Development surveys in Cameroon, Ghana, Kenya, and Zimbabwe. In each country three rounds of interviews were held during 1992–95. For Ghana we also have data for 1996, from a survey Oxford University organized using a similar survey instrument. The samples were chosen from four manufacturing subsectors: food, textile and clothing, wood and furniture, and metal working and machinery. Firm size ranged from fewer than five employees to well over a thousand. The average size of firms was smallest in Ghana (36 employees), largest in Zimbabwe (303). The questionnaire covered the firm's finance, technology, contractual relations with suppliers and clients, labor force, marketing, capital, investment, and exports.[6]

Many of the firms did not have audited accounts. So data on such concepts as value added or profits were not collected directly but constructed on the basis of questions about sales and direct and indirect cost items. The labor market module of the questionnaire included questions about total wage costs, the composition of the labor force, and workers' educational attainments. A separate questionnaire was administered to a small sample of (at most) 10 workers in each firm. These individual data concerned wages, education, work experience, and job description. Workers were interviewed only once; although we have panel data for firms, we have only cross-section data for workers. In the analysis we use the linked firm–worker data to test for risk and rent sharing. We focus on total earnings and thus do not distinguish between basic earnings and allowances (including bonuses) because the information on bonuses does not always seem reliable.

All available observations were used, except suspicious data excluded using the following rules: (1) value added was negative; (2) sales, wages, or employment increased or fell by more than a factor 5 between one year and the next;

6. A very detailed description of the data, sampling, and questionnaire can be found online at www.worldbank.org/research/pics.

and (3) mean wages in the firm were less than 5 percent or more than 500 percent of the mean wage in the sample.

III. Results

Risk Sharing: Prima Facie Evidence

If the risk-sharing model is correct, firms facing the most volatility in performance should exhibit the greatest volatility in average wages per worker or in employment. As a first step we calculated the volatility of profits, value added, and sales, checking whether they are systematically related to the observed volatility in employment or average wages per worker. Volatility is measured as the standard deviation of the log of annual observations for profits, value added, sales, employment, and average wages per worker for firms with at least three annual observations. For Ghana we have observations on four rounds for some of the firms, but for the other countries, at most three. Average wages per worker are calculated by the total wage bill (including allowances) divided by the total number of workers as reported in the firms' questionnaires. Because volatility is measured as the standard deviation of logs, the volatility of profits is calculated only for firms with profits. This may bias the result, but the standard deviation of profits *per worker* is not a suitable alternative because it is directly affected by the volatility in employment.

Table 1 reports the findings, with firms ordered by quartiles of profits, value added, and sales volatility. If the risk-sharing hypothesis is correct, the standard

Table 1. Volatility of Employment and Average Wages by Volatility of Sales, Value Added, and Profits

	Median of firm-specific standard deviations of	
	Annual employment	Average wages per worker
Quartiles of $\sigma_{profits}$		
25	0.12	0.23
50	0.16	0.28
75	0.18	0.36
100	0.23	0.45
Quartiles of $\sigma_{value\ added}$		
25	0.12	0.22
50	0.14	0.27
75	0.24	0.39
100	0.22	0.44
Quartiles of σ_{sales}		
25	0.10	0.25
50	0.18	0.34
75	0.19	0.40
100	0.28	0.53

Note: Standard deviations are calculated in terms of logs. Values are in PPP US$. Only firms with at least three observations are included.

deviation (measured by the median across firms in a given quartile) of the log of employment and of average wages per worker should increase going down the columns.

This is indeed the case: both employment and wages are more volatile in firms with more volatile profits. The effect is remarkably strong: for both variables the standard deviation roughly doubles from the bottom profits quartile to the top.[7] This is a necessary condition for risk sharing: under risk sharing (temporary) shocks should affect both profits and wages or employment.

Changes in profits are, of course, endogenous to shocks. We therefore present the same information for the standard deviation of sales and value added, more direct measures of exogenous shocks. The evidence for these two variables is very similar to that for profits: greater volatility is associated with higher volatility of both wages and employment. Our focus here is on risk sharing in the form of workers accepting wage volatility, but note that there is also prima facie evidence of substantial employment volatility.

Clearly the evidence is only suggestive. The standard deviation records all changes and thus does not distinguish between permanent and temporary shocks. In addition, table 1 does not control for firm characteristics that may also affect wages (other than the shocks to which the firm is exposed).

Our next step is to construct measures of temporary and permanent shocks and include these as regressors in an earnings function regression. Bigsten and others (2000a, table 7) estimated a Mincerian earnings function using the Regional Programme on Enterprise Development data for Cameroon, Ghana, Kenya, Zimbabwe, and Zambia. The dependent variable is the log of earnings (in PPP dollars) and the regressors are human capital measures (dummy variables indicating whether a worker had completed primary, secondary, or university education), the worker's age and age squared, tenure with the firm (also squared), and firm dummies to control for observable and unobservable firm differences. Bigsten and others also include the log of firm size (number of employees) and capital intensity because workers in larger and more capital-intensive firms tend to receive higher wages. This is commonly interpreted as some type of rent sharing between the larger and more capital-intensive firms and the workers. We adopt the same specification but replace firm size and capital intensity with explicit measures for temporary and permanent shocks to test for risk and rent sharing directly.

But first we need to address reverse causality or endogeneity. In the risk- and rent-sharing literature, better firm performance leads to higher wages. By contrast, the efficiency wage literature emphasizes the opposite direction of

7. Interestingly, this is not true for wages as reported in the workers' survey. Using those data we find very little change across quartiles, the median standard deviation of wages taking the values 0.32, 0.28, 0.28, and 0.33 (top quartile). This strongly suggests that risk sharing in the form of wage volatility is dominated by bonuses.

causality: higher wages lead to better firm performance through less shirking, lower labor turnover, or higher morale (Shapiro and Stiglitz 1984; Dasgupta and Ray 1986). Such an efficiency wage mechanism introduces an endogeneity bias. Blanchflower and others (1996) suggest two possible remedies. First, if it takes time for wages to adjust to shocks the equation is recursive and wages become a function of past shocks. With this remedy, past movements in value added can be viewed as predetermined, and wages can be estimated as a function of lagged values of shocks. The second remedy is to find valid instruments to correct for potential endogeneity bias.

We adopt the first approach, for three reasons. First, there is evidence that the impact of shocks on wages is indeed recursive. Blanchflower and others (1996, table 2) show that the effect of profit shocks on wages increases over time. Second, the lag structure test for risk versus rent sharing: the effect of permanent shocks on wages should persist, whereas the effect of temporary shocks should taper off quickly. This implies that longer lags for permanent shocks should still be significant, while becoming insignificant for temporary shocks. Third, there are no obvious valid instruments.

Blanchflower and others (1996) suggest cost measures, such as the cost of energy, as instruments. Teal (1996) uses the amount of foreign borrowing per employee and the share of intermediate imported input costs in total output for a study on Ghana, where the exchange rate fell far faster than domestic prices rose during the study period. But even these cost measures are problematic if better firm performance comes with the capability of workers to adapt to international best practice and thus the greater use of energy-intensive foreign inputs. The basic problem is that clearly exogenous instruments, such as firm-specific input or output prices, are simply difficult to measure and therefore not available.

For all these reasons we use lagged values of permanent and temporary shocks to test for risk and rent sharing. Endogeneity is also mitigated because we use shocks in value added rather than profits to test for risk and rent sharing. The use of value-added data allows two period lags for each of the countries because of the availability of retrospective data.[8]

Risk Sharing Using a Production Function Specification

We estimate a Cobb-Douglas production function regressing the log of value added on the log of the firm's physical capital and the log of the number of employees. We allow the coefficients for physical capital and employment to vary

8. Retrospective data are available for sales, employment, and investment but not for the use of intermediate inputs. We therefore calculate value added for the presurvey years from the retrospective sales data by applying the ratio of value added to sales observed in the survey years for each firm. The results are virtually the same if we use the estimated value added series also for the survey years, suggesting that this procedure does not create any serious bias.

TABLE 2. Cobb-Douglas Production Function for Log Value Added

	(1) Random effects		(2) Fixed effects	
	Coefficient	*t*-score	Coefficient	*t*-score
Log capital				
Cameroon	0.21	6.38	0.05	1.09
Ghana	0.21	7.50	0.01	0.34
Kenya	0.19	6.62	0.07	1.78
Zimbabwe	0.25	7.13	0.01	0.31
Log employment				
Cameroon	1.09	18.54	0.58	6.15
Ghana	0.79	11.85	0.28	2.57
Kenya	0.93	17.64	0.44	5.21
Zimbabwe	0.88	14.14	0.34	2.81
N	2,788		2,788	
R^2	0.81		0.52	

Note: Also included in the regression are country- and year-specific sector dummies.

across countries and also include country and year-specific sector dummies.[9] We also include random or fixed effects to control for unobserved firm heterogeneity.

The coefficients for physical capital and employment are mostly significant but typically lower in the fixed effects specification (table 2). The Hausman test of random versus fixed effects is rejected in favor of fixed effects. But the coefficients for physical capital are implausibly low in the fixed-effects regression, suggesting serious measurement errors. Fortunately none of the following results is affected in any serious way if we use the random-effects specification rather than the fixed effects, which we do.[10]

We use the production function results to construct the long-run shock s^l as the value predicted by the regressors and the controls not reported in table 2, except for employment. We do not include the part predicted by employment in the construction of the long-run shock because employment may vary in the short run because of risk sharing. This implies that an increase in value added as a result of an increase in the firm's capital stock is treated as a permanent shock, possibly giving rise to rent sharing. Similarly, if there are productivity changes at the sectoral level, we treat them as permanent. We consider all other changes in value added—movements along the production function as a result of changes in employment or deviations from the regression line—as our measure of short-run shocks (s^s). In effect we treat changes in value added not due to changes in the firm's physical capital or firm sector as temporary shocks. Of course we can interpret the impact of short-run

9. This is tantamount to estimating the production function country by country. We do this because pooling is rejected.

10. In fact, the results are also almost unaffected if we use a pooled (random or fixed effects) production function for all countries.

shocks on wages as risk sharing only if these shocks have low persistence. This turns out to be the case: the first-order autocorrelation (allowing for firm fixed effects) is 0.19, suggesting that persistence is a small proportion of our short-run shock variable.[11]

We scale the long-run shock by the number of employees because the model suggests a stable relationship between wages and profits per worker for rent sharing (but between wages and total profits for risk sharing).[12]

The question we address is whether these two measures of shocks affect wages paid by the firm, controlling for the human capital variables and the other determinants of wages identified by Bigsten and others (2000a). We therefore introduce the shock variables as regressors in the earnings function (table 2). Firm fixed effects and country-specific year dummies are included to control for firm heterogeneity and macroeconomic variations. We find a strong pattern across length of lag and type of shock. The short-run shock is positive and significant only for the one-period lag, and the long-run shock is positive and significant for the two-period lag. That the coefficient for long-run shocks increases with the length of the lag has also been reported by Blanchflower and others (1996, table 2).

The risk- and rent-sharing coefficients pool across countries at the 10 percent significance level. Virtually the same estimates are found if the fixed effects Cobb-Douglas production function is used to construct short- and long-run shock variables.

Worker Heterogeneity

We controlled for such observed worker characteristics as age, tenure, and education. But we cannot control for worker heterogeneity in terms of unobservables, because in our panel data firms were followed over time but workers were interviewed only once. So we can introduce fixed effects for firms (but not for workers) to control for unobserved heterogeneity.[13] This is a serious limitation. For example, Abowd and others (1999) show (using French data) that unobserved worker heterogeneity (which we do not control for) is large relative to unobserved firm heterogeneity (which we do control for).

11. The 0.19 estimate is subject to the Nickell bias because of the short length of our panel. But in our estimates (see table 3) we observe that the short-run shocks are significant for a one-period lag but not for a two-period lag. This also suggests low persistence in the short-run shock variable.

12. This point is also noted by Blanchflower and others (1996, p. 239) who therefore view the wage correlation with profit per employee as evidence for rent sharing rather than risk sharing (which implies a wage correlation with total profits). If we also scale the short-run shock by the number of employees, the results remain generally the same. To be more specific, the effect is to leave the coefficients of the short-run shock variable exactly as reported in table 3 and to change their t-scores marginally. For the long-run coefficients, as before, the coefficient is larger and more significant for the two-period lag.

13. One could argue that worker characteristics, such as tenure, are endogenous because they may be correlated with omitted worker characteristics. But even if we exclude tenure from the specification, the size and significance of the coefficients of the shock variables are unaffected.

TABLE 3. Effects of Value-Added Shocks on Monthly Wages

	One-period lag		Two-period lag	
	Coefficient	*t*-score	Coefficient	*t*-score
Short-run shock s^s	0.03	2.13	−0.01	0.59
Long-run shock s^l	0.01	0.49	0.07	1.77
Primary education	0.04	1.79	0.03	1.12
Secondary education	0.20	8.58	0.18	6.88
University education	0.80	17.33	0.79	14.33
Age	0.08	17.11	0.09	15.75
Age^2	−0.001	14.22	−0.001	13.16
Tenure	0.01	2.70	0.01	1.82
$Tenure^2$	−0.00001	0.13	−0.0001	0.63
Firm dummies	yes		yes	
N	6,789		5,248	
R^2	0.52		0.47	

Note: Also included in the regression are country-specific year dummies and dummies for sex and occupation of worker. The number of observations is lower in the 2-period case because lags are not available for all firms. Limiting the regression to the same observations does not affect the results.

Workers may differ in how they are affected by shocks.[14] For example, risk sharing could take the form of reducing the number of hours worked but only for junior workers. Or risk sharing could impose wage volatility on junior workers but not the senior. In either case the effect of short-run shocks on our wage variable (monthly earnings) would differ by worker category. We investigate this by including an interaction term ($s^s \times$ age) in the table 3 regression. This interaction term is insignificant (*t*-values of 0.95 and 1.20 for the one- and two-lag specifications, respectively), so this form of heterogeneity does not appear to pose a problem.

Is It Really Risk Sharing?

The evidence in table 3 clearly suggests that workers in Sub-Saharan manufacturing share risks (and rents) with their firms. But is the relationship between the short-run shocks in value added and wages really a consequence of risk sharing and not something else? We investigate this by looking at three elements of risk sharing: labor market imperfections, credit market imperfections, and compensating differentials.

Risk Sharing and Labor Market Imperfections

Risk sharing implies some kind of labor market imperfection or friction. If workers could move costlessly to another firm when the firm experienced a negative shock, a risk-sharing arrangement clearly could not be enforced. Risk-sharing contracts

14. We are grateful to one of the referees for this point.

can be enforced only when firms and workers are somehow locked in. We expect production workers to be more mobile than other workers who typically have more firm-specific skills and experience. The reason: the label "production workers" covers people with few skills (unlike accountants, supervisors, managers, and technicians). This leaves "production workers" as a relatively unskilled and presumably more mobile group, an interpretation supported by data on tenure. For example, in a regression of tenure on a dummy for production workers (controlling for age, quadratically, and including firm fixed effects) the dummy is negative and significant. So production workers in our sample are indeed more mobile between firms than other workers. This makes a risk-sharing contract more difficult to implement for this category.

To test for this, we estimated the wage regression for production and other workers separately (the first two regressions of table 4). We include only one-period lags because there is no evidence of risk sharing for two-period lags. The result confirms the risk-sharing interpretation: risk sharing is observed only for nonproduction workers, with a positive and highly significant coefficient for the short-run shock variable.

Risk Sharing and Credit Market Imperfections

Risk sharing is a form of informal credit for the firm, unnecessary if credit markets function perfectly. But there is ample evidence that credit markets function less than perfectly, especially in developing economies, and risk sharing

TABLE 4. Effects of Value Added Shocks on Monthly Wages: Production versus Other Workers and Demand Volatility

	Production workers		Other workers		Demand volatility	
	Coefficient	*t*-score	Coefficient	*t*-score	Coefficient	*t*-score
Short-run shocks s^s	0.02	1.31	0.06	2.71		
×Low demand volatility					0.02	0.54
×High demand volatility					0.07	2.69
Long-run shocks s^l	−0.03	0.94	0.06	1.42	0.05	1.32
Primary education	0.03	1.35	0.05	1.26	0.05	1.26
Secondary education	0.16	5.23	0.27	6.48	0.27	6.47
University education	1.05	13.02	0.70	11.25	0.70	11.19
Age	0.09	14.85	0.05	6.06	0.05	6.03
Age^2	−0.001	12.67	−0.0005	4.30	−0.0005	4.27
Tenure	0.01	2.97	0.005	1.12	0.01	1.14
Tenure^2	−0.0001	0.91	−0.0001	0.46	−0.0001	0.51
Firm dummies	yes		yes		yes	
N	4,292		2,497		2,478	
R^2	0.36		0.38		0.23	

Note: Also included in the regression are country-specific year dummies and dummies for sex and occupation of worker. The number of observations in columns 1 and 2 sums to 6,789, as in column 1 of table 3. The number of observations in column 3 is lower than in column 2 because demand volatility has been calculated only for firms reporting at least 3 years of data.

may be an efficient response to such imperfections. We would expect firms that are more credit constrained to be more likely to use risk sharing as a credit tool. In principle, we could test whether more credit-constrained firms exhibit more risk sharing using some proxy for how credit constrained they are.

Causality can, however, also run from risk sharing to credit rationing. Presumably firms that are highly effective in sharing risks with workers will be less credit constrained than firms that have to rely entirely on outside finance. But if we assume that all firms are credit constrained to some extent, the ones exposed to the greatest outside shocks (such as those in demand and input supply) will be most likely to use some form of risk sharing to cope with liquidity crises. Outside shocks are exogenous, and therefore a valid test for the risk sharing interpretation of our result.

We tested whether firms that face greater demand shocks also share more risks. We calculated demand volatility as the standard deviation of annual log of sales for each firm and divided the firms into low and high demand volatility depending on whether the standard deviation was below or above the median. Next we tested whether wages react differently to short-run value-added shocks for low and high demand volatility firms.

The third regression in table 4 presents the results for the nonproduction workers, for whom we found evidence of risk sharing. The results confirm that these workers share risk but only in firms with high demand volatility: there is no evidence that nonproduction workers share risks in low demand volatility firms.

Risk Sharing and Wage Compensation

The results may be viewed as strong evidence of risk sharing in African manufacturing. The efficiency wage explanation for the correlation between wages and value added was excluded through the use of lagged values of the shock variables. There is also evidence of rent sharing, but this cannot explain the variation of wages with short-run variations in value added. But Blanchflower and others (1996) discuss an alternative hypothesis that could explain why wages are affected by short-run shocks in value added. If wages are set competitively but there is slow adjustment, positive value-added shocks may lead to increases in wages if firms temporarily move up the labor supply curve in booming industries. The rent-sharing hypothesis can withstand this criticism because it has been shown that firm and industry differentials in wages tend to persist over very long periods. The competitive-slow-adjustment model cannot explain such long-term variations from the competitive model. But it can explain the short-run variations in wages following short-run shocks in value added. Risk sharing also implies that workers are compensated if they enter the implicit contract ex ante or punished if they are forced to accept the contract ex post. If they enter the contract voluntarily, they will demand (and receive) a risk premium for incurring some of the risks the firm is facing. If they are forced to accept risk sharing because of unexpected volatility occurring during the contract period and because of the costs of job mobility, they will pay a risk

TABLE 5. Testing for Wage Compensation: Effects of Intrafirm Wage Volatility

	Coefficient	*t*-score
Short-run shock s^s	0.03	2.50
Long-run shock s^l	0.00	0.09
Primary education	0.03	1.35
Secondary education	0.18	6.94
University education	0.77	15.14
Age	0.08	15.89
Age^2	−0.001	13.14
Tenure	0.01	3.06
Tenure^2	−0.0001	0.85
Intrafirm wage volatility	0.12	2.18
Firm dummies	yes	
N	5,748	
R^2	0.50	

Note: Also included in the regression are country-specific dummies for year and for the sex amd occupation of the worker. The number of observations is lower than in table 3 (6,789) because intrafirm wage volatility is not available for all firms.

premium. Either way, risk sharing implies a correlation between wages and short-run firm-specific wage volatility that cannot be explained by the competitive-slow-adjustment model.

We measure the short-run firm-specific wage volatility as the standard deviation of the log average wage across time within a firm. The average wage is measured as the ratio of the total wage bill (including allowances) and the total number of employees. To create a firm-specific measure of compensation that also varies by type of worker, the standard deviation of the log wage (within a firm) is calculated for production and nonproduction workers. We need intrafirm variation in the wage volatility measure to include a volatility measure in a wage regression together with firm fixed effects.

Firms with greater intrafirm wage volatility pay higher wages to the workers, even after controlling for firm and worker characteristics (table 5). This confirms the risk-sharing hypothesis as opposed to the competitive-slow-adjustment hypothesis. With intrafirm wage volatility increasing from 0.06 to 0.63 between the 10th and 90th percentile, workers receive an additional 7 percent wage as compensation if they were to move from the low-wage-volatility firm to the high.

IV. How Important is Risk Sharing?

How important are the risk-sharing effects identified in the article? How important are they relative to the other determinants of wages? To capture the importance of each of these determinants, we consider how much wages change if we move from

TABLE 6. The Importance of Risk Sharing in Determining Wage Differentials

	90th Percentile	Median	10th Percentile	% Increase
Log monthly wages	4.46	5.26	6.05	392
Transient shock	1.74	3.68	5.49	15
Permanent shock	3.81	5.27	6.82	28
Education (dummy)	secondary completed	primary completed	primary not completed	20

Note: The percentage increase in the rows for transient shocks, permanent shocks, and education indicate how much wages would rise in response to a change in those variables from the 10th to the 90th percentile.

the 10th to the 90th percentile of the distribution for each of these determinants. Table 6 shows how much of the rise in earnings between the 10th and 90th percentile of the distribution can be attributed to each of the explanatory variables. Average wages increase 392 percent between the 10th and 90th percentile. Wages would increase by 15 percent if the short-run shock increases from the 10th to the 90th percentile of its distribution. The wage increase for a similar move for the long-run value-added shock is 28 percent. This suggests that one-third of the combined impact of risk and rent sharing can be traced to risk rather than rent sharing. This is important, because it qualifies the common perception that African labor markets are inefficient in light of widespread rent sharing. In fact, African labor markets appear to be highly efficient. They compensate for poorly developed financial markets by allowing firms to share negative shocks with their workers. These risk and rent sharing effects are comparable to the effect of education (20 percent).

V. CONCLUSION

The developed economy literature has focused on rent sharing as an explanation for the correlation across firms between wages and profits typically observed in cross-section data. The alternative explanation of risk sharing has been explored in theoretical work but not tested empirically. Under risk sharing, only part of the effect of shocks is on profits. Workers provide insurance to employers by accepting changes in employment or wages.

We investigated this hypothesis using a four-country panel data set for African manufacturing firms. The descriptive analysis suggests a prima facie case for risk sharing: the more volatile a firm's value added, the greater the volatility of the average wages it pays—and of the size of its labor force. Risk sharing in our sample thus involves changes both in employment and in average wages.

In the econometric analysis, we focused on changes in wages. We estimated an individual earnings function including a firm-specific measure of shocks among the regressors and controlling for worker and firm characteristics. We separated measures of permanent and temporary shocks by estimating a

production function (with time-variant stocks of physical capital and employment). Value-added changes resulting from investment or time trends (sector and country specific) are treated as long-run shocks. Deviations from this permanent component are treated as temporary shocks. We found evidence of risk sharing, with temporary shocks having a positive and significant effect on wages.

Because we used lagged values for the shock variables, this effect cannot be easily explained by an efficiency wage hypothesis. The observed lag structure for short-run versus long-run shocks also suggests that it is risk sharing rather than rent sharing. Short-run shocks affect wages but only temporarily, whereas long-run shocks have a longer effect on wages. We also find that nonproduction workers are more likely to share risks than production workers. This is consistent with the risk-sharing hypothesis because the skills of nonproduction workers are more firm-specific than those of production workers. So a risk-sharing contract can be enforced more easily for nonproduction workers. Finally, the positive effect of short-run shocks on wages cannot be explained by a short-run, upward-sloping labor supply curve because we also find that workers in firms with high intrafirm wage volatility receive a risk premium. The risk-sharing hypothesis can explain this premium, but the competitive-slow-adjustment model cannot.

When the correlation between wages and profits reflects risk sharing rather than rent sharing, it need not be interpreted as a symptom of a labor market inefficiency. Instead, risk sharing can be seen as an efficient (second-best) use of the labor market to substitute for a missing insurance market. Our evidence suggests that risk sharing is quantitatively important: one-third of the combined impact of risk and rent sharing can be traced to risk sharing rather than to rent sharing, the literature's focus.

What does risk sharing mean for policy? Consider insurance for firms. Because firms would substitute formal insurance for risk sharing, wage volatility would be reduced. But to the extent that workers are compensated for the wage volatility they experience under risk sharing, they will not benefit from the insurance. Firms, however, would clearly be better off because the possibilities for risk sharing with workers are obviously rather limited. Formal insurance would give the firm access to a bigger risk pool. If wages are rigid, there would be even less scope for risk sharing. Wage rigidity is therefore damaging not just by distorting the labor market but by blocking the possibility of risk sharing as a substitute for insurance.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67(2):251–333.

Azariadis, C. 1975. "Implicit Contracts and Underemployment Equilibria." *Journal of Political Economy* 83:1183–202.

Baily, M. N. 1974. "Wages and Employment under Uncertain Demand." *Review of Economic Studies* 41:37–50.

Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Isaksson, A. Oduro, R. Oostendorp, C. Pattillo, M. Söderbom, F. Teal, and A. Zeufack. 1999. "Investment in Africa's Manufacturing Sector: A Four Country Panel Data Analysis." *Oxford Bulletin of Economics and Statistics* 61:489–512.

Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Söderbom, F. Teal, A. Zeufack, and S. Appleton. 2000a. "Rates of Return on Physical and Human Capital in Africa's Manufacturing Sector." *Economic Development and Cultural Change* 48:801–27.

Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Söderbom, F. Teal, and A. Zeufack. 2000b. "Contract Flexibility and Dispute Resolution in African Manufacturing." *Journal of Development Studies* 36:1–37.

Blanchflower, D. G., A. J. Oswald, and P. Sanfey. 1996. "Wages, Profits and Rent-Sharing." *Quarterly Journal of Economics* 111:227–51.

Collier, P., and J. W. Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37:64–111.

Dasgupta, P., and D. Ray. 1986. "Inequality as Determinant of Malnutrition and Unemployment." *Economic Journal* 96:1011–34.

Fafchamps, M. 1996. "The Enforcement of Commercial Contracts in Ghana." *World Development* 24:427–48.

Fafchamps, M., J. W. Gunning, and R. Oostendorp. 2000. "Inventories and Risk in African Manufacturing." *Economic Journal* 110:861–93.

Malcomson, J. M. 1999. "Individual Employment Contracts." In O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*. Amsterdam: North-Holland.

Mazumdar, D., and A. Mazaheri. 1998. "The Structure of Labor Market and Wages in African Manufacturing." Mimeo. University of Toronto.

Oswald, A. J. 1995. "Rent Sharing in the Labor Market." Mimeo. London School of Economics.

Shapiro, C., and J. E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74:433–44.

Teal, F. 1996. "The Size and Sources of Economic Rents in a Developing Country Manufacturing Labor Market." *Economic Journal* 106:963–76.

Velenchik, A. 1997. "Government Intervention, Efficiency Wages and the Employer Size Wage Effect in Zimbabwe." *Journal of Development Economics* 53:305–38.

Trade Facilitation and Economic Development: A New Approach to Quantifying the Impact

John S. Wilson, Catherine L. Mann, and Tsunehiro Otsuki

This article analyzes the relationship between trade facilitation and trade flows in the Asia-Pacific region. Country-specific data for port efficiency, customs environment, regulatory environment, and e-business usage are used to construct indicators for measuring trade facilitation. The relationship between these indicators and trade flows is estimated using a gravity model that includes tariffs and other standard variables. Enhanced port efficiency has a large and positive effect on trade flows. Regulatory barriers deter trade. Improvements in customs and greater e-business use significantly expand trade but to a lesser degree than improvements in ports or regulations. The benefits of specific trade facilitation efforts are estimated by quantifying differential improvements in these four areas among members of the Asia Pacific Economic Cooperation (APEC). A scenario in which APEC members with below-average indicators improve capacity halfway to the average for all members shows that intra-APEC trade could increase by $254 billion, or 21 percent of intra-APEC trade flows. About half the increase is derived from improved port efficiency.

Economic theory suggests a relatively direct and simple chain of causality: human development is enhanced through income growth; income growth is greater with more cross-border trade; trade is increased through trade facilitation efforts. Recent empirical work has focused on quantifying each of these links. The human development index is positively related to gross domestic product (GDP) per capita, and countries with a growing income have a higher GDP per capita. Though the positive relationship between trade and growth has come under scrutiny recently, there is no evidence that increased cross-border trade reduces income growth. The focus of this article is on the last (or perhaps first) link in the chain—the empirical relationship between trade facilitation and trade flows.

Trade facilitation most often implies improving efficiency in administration and procedures, along with improving logistics at ports and customs. A broader

John S. Wilson is Lead Economist in Transport and Urban Development, Infrastructure Vice Presidency at the World Bank. Catherine L. Mann is Senior Fellow at the Institute for International Economics. Tsunehiro Otsuki is Research Analyst in the Development Research Group at the World Bank. Correspondence should be directed to Tsunehiro Otsuki at [illegible] This article is part of a series of research efforts exploring the link between trade facilitation and development at the World Bank. The authors thank Baishali Majumdar of the World Bank for assistance in producing the manuscript. Comments by Caroline Freund, Carsten Fink, and Bernard Hoekman on the work are also appreciated.

THE WORLD BANK ECONOMIC REVIEW, VOL. 17, NO. 3 367–389

DOI: 10.1[illegible]

definition includes streamlining regulatory environments, deepening harmonization of standards, and conforming to international regulations (Woo and Wilson 2000). Emphasizing broader concepts of trade facilitation is particularly important given the increasing volume of global trade, the time sensitivity of intermediate goods trade (Hummels 2001), reductions in protective tariff rates, and increased availability of modern technology that can improve the management of cross-border trade. This article examines the empirical relationship between relatively broad concepts of trade facilitation and trade flows. It also compares trade facilitation initiatives with reductions in traditional trade barriers (such as tariffs and quotas) for their impact on trade flows and considers complementarities between them.

As part of individual development strategies, some countries, with or without donor assistance or private funding and partnerships, are considering whether to engage in unilateral trade facilitation efforts. They also confront the challenge of priorities within a broad range of trade facilitation measures. At the 1996 Singapore Ministerial Conference of the World Trade Organization (WTO) trade facilitation was added to the new basket of trade issues. Discussions continue on priorities in trade facilitation in the WTO Doha Development Agenda. Decisions on modalities for negotiations on trade facilitation, including customs procedures, were on the agenda for the WTO Ministerial Conference in Mexico in September 2003. Debate also continues on such issues as whether to extend the Information Technology Agreement to nontariff measures and standards, as well as whether international standards should be mandated in national regulations. The lack of empirical measures of trade facilitation and of their impact on international commerce limits informed debate.

At least three challenges are apparent in empirical research on trade facilitation: defining and measuring trade facilitation, choosing a modeling methodology to estimate the importance of trade facilitation for trade flows, and designing a scenario to estimate the effect of improved trade facilitation on trade flows. The research approach taken here contributes to a deeper understanding in each of these areas. The article explores these topics for trade among members of the Asia Pacific Economic Cooperation (APEC), which accounts for about 57 percent of world GDP and about 47 percent of global trade.

First, trade facilitation is defined and measured using four indicators (port efficiency, customs environment, regulatory environment, and e-business usage) rather than a single parameter to proxy trade facilitation, such as import prices, international transport costs, or productivity of the transport sector. Second, a gravity model of bilateral trade flows rather than a computable general equilibrium (CGE) approach is used to model cross-border trade and to estimate the effect of trade facilitation on trade. Third, the scenarios explored to determine the benefits of trade facilitation do not assume that all countries improve capacity by the same amount to support trade flows. The simulations acknowledge that some countries have farther to go to reach best practice in regulatory reform or port efficiency, for example, than do others.

Section I reviews definitions of trade facilitation and previous efforts to measure the impact on trade. Sections II and III discuss the data, methodology,

and results from the empirical model used to estimate the relationship between bilateral trade flows and country-specific trade facilitation measures. Section IV presents simulation exercises exploring the consequences of improving trade facilitation measures for APEC as a whole, for individual countries as exporters to APEC, and for individual countries as importers from APEC.

I. Overview of Previous Work on Measuring the Impact of Trade Facilitation

There is no standard definition of trade facilitation. In a narrow sense, trade facilitation simply addresses the logistics of moving goods through ports or customs at the border. A broader definition includes the environment in which trade transactions take place, including the transparency of regulatory environments, harmonization of standards, and conformance to international or regional regulations. The focus has now moved behind the border to domestic policies and institutional structures, where capacity building plays an important role. In addition, the rapid integration of networked information technology into trade means that service sector infrastructure that supports technology use is also relevant. The definition of trade facilitation used here incorporates such border elements as port efficiency and customs administration and such behind-the-border elements as the domestic regulatory environment and the infrastructure to enable e-business.

The empirical literature on trade facilitation is limited. Maskus and others (2001) address some of the more important empirical methods and challenges in quantifying the gains of trade facilitation in the area of harmonized regulations. The Asia Pacific Foundation of Canada (1999) outlines the relative importance of three kinds of trade facilitation measures (customs, standards and regulatory conformance, and business mobility) for APEC business but does not assess the impact on APEC trade of trade facilitation improvements. The Australian Department of Foreign Affairs and Trade and Chinese Ministry of Foreign Trade and Economic Cooperation (2001) suggest that moving to electronic documentation for APEC trade would yield a cost savings of some 1.5 to 15 percent of the landed cost of an imported item. Applying a simple average of a 3 percent reduction in landed costs from electronic documentation for intra-APEC merchandise trade yields gross savings of US$60 billion. The Organisation for Economic Co-operation and Development (OECD 2001) summarizes other studies, but most use a limited definition of trade facilitation or old data.

Several recent studies use CGE models to quantify the benefits of improved trade facilitation. In CGE models an improvement in trade facilitation can be modeled equivalently as a reduction in the costs of international trade or as an improvement in the productivity of the international transport sector. Because this sector is already included in the CGE model, the effect of improved trade facilitation comes from "shocking" the sector by an appropriate amount. The United Nations Conference on Trade and Development (UNCTAD 2001) uses CGE analysis to consider trade facilitation in the broader context of creating an environment conducive to developing e-commerce. The objective of the CGE

analysis is to consider the relationship between a given size shock to productivity growth, applied equally to all members of the group, on the GDP of regional groups of countries. The results show that a 1 percent reduction in the cost of maritime and air transport could increase Asian GDP some $3.3 billion. If trade facilitation is considered in a broader sense to include an improvement in wholesale and retail trade services, a 1 percent improvement in the productivity of that sector could increase GDP an additional $3.6 billion.

APEC (1999) also uses CGE analysis. The shock reduction in trade costs from trade facilitation efforts differs among members of the group, with a 1 percent reduction in import prices for the industrial countries and the newly industrializing economies (the Republic of Korea, Singapore, and Taiwan, China) and 2 percent for the other developing economies. APEC (1999) estimates that APEC merchandise exports would increase by 3.3 percent from the trade facilitation effort to reduce costs. In comparison, the model estimates a long-run increase of 7.9 percent in APEC merchandise exports from completing Uruguay Round commitments.

Hertel and others (2001) use CGE analysis to quantify the impact on trade of harmonizing standards for e-business and automating customs procedures between Japan and Singapore. They find that these reforms will increase trade flows between these countries as well as with the rest of the world.

Other research addresses specific aspects of the trade facilitation agenda and uses gravity model analysis. Freund and Weinhold (2000) find that a 10 percent increase in the relative number of Web hosts in one country would have increased trade flows by 1 percent in 1998 and 1999. Fink and others (2002b) find that a 10 percent decrease in communication costs is associated with an 8 percent increase in bilateral trade. Moenius (2000) finds that bilaterally shared standards can promote trade. Otsuki and others (2001b), looking at food safety standards, find that a 10 percent tighter EU standard on aflatoxin contamination levels would reduce African exports by 4.3 percent for cereals and 11 percent for nuts and dried fruit. The study reported here uses a gravity model of bilateral trade in the region and incorporates a richer set of indicators of trade facilitation in the analysis. The model also includes tariffs to determine which of these factors might have a greater effect on trade flows within APEC.

II. Data for Measuring Trade Facilitation and Trade Flows

The greatest challenge to new research on trade facilitation is to find conceptually distinct measures of trade facilitation to meet policymakers' needs for specificity. Should policymakers focus scarce resources on port modernization, customs reform, regulatory harmonization, or e-commerce infrastructure? There are clearly synergies among these reforms. Limited resources, however, mean that not all reforms can be addressed simultaneously. Previous efforts to proxy trade facilitation using import prices or transportation costs are not adequate for informing policy priorities.

This analysis includes four indicators that measure four different categories of trade facilitation effort:

- *Port efficiency*, designed to measure the quality of infrastructure of ports and airports.
- *Customs environment*, designed to measure direct customs costs as well as administrative transparency of customs and border crossings.
- *Regulatory environment*, designed to measure the economy's approach to regulations.
- *E-business usage*, designed to measure the extent to which an economy has the necessary domestic infrastructure (telecommunications, financial intermediaries, logistics firms) and is using networked information to improve efficiency and transform activities to enhance economic activity.[1]

Each indicator is generated from data specific to each APEC economy. The indicators alone can help policymakers judge how their economy rates relative to APEC's best practice in each of these areas. Self-assessments against best practice and estimation results on the effect of indicators on trade flows provide useful information to policymakers about what might be the most fruitful direction for reform, capacity building, and negotiation.

Survey data were used to generate the four indicators because no other empirical data are available on a consistent basis for all the APEC members. Although some APEC members have done empirical studies of, for example, improvements in customs costs or release times from customs warehouses, the gains obtained by a country (such as Singapore) cannot be assumed to apply equally to other countries. The objective of the research for this study is to distinguish one country from another in the need for capacity building or pilot projects in the various trade facilitation areas.

In addition, the data available on the conceptual basis relevant for the trade facilitation analysis are also limited. Consistent and country-specific assessments are needed for port efficiency, customs environment, regulatory environment, and e-business usage. Survey data are used in the analysis because they are available for the range of trade facilitation indictors to be examined. Although the data must be used with caution and checked across alternative sources for similar proxies, they offer the potential for cross-country qualitative and quantitative analysis to inform policy discussion and debate.

Generating Trade Facilitation Indicators

Each of the four trade facilitation indicators is constructed with multiple data inputs (oversampling) to reduce dependence on any one survey response. The

1. For further discussion of the relationship between domestic infrastructure and e-commerce, see Mann and and others (2000).

inputs can be analyzed to gain even greater information about trade facilitation measures for individual economies and across APEC. Because some of the data are actual values and some come from surveys with different response ranges (1 to 7, 1 to 10, and so on), the raw data need to be put on a comparable basis. Each APEC-specific observation of a raw series is indexed to the average of all the APEC members' value for the raw series, yielding an indexed input.[2]

Next, the indexed inputs into the four specific trade facilitation indicators are averaged. For greater transparency—and because there is no specific argument (theoretical or statistical) for choosing a different aggregation method—a simple average is used.[3] Details of the sources and survey questions underpinning each of the indexed inputs are in the appendix.

Examining the indexed inputs that are averaged to generate the indicators is informative for several reasons. First, summary statistics on the indexed inputs and the aggregated indicators identify where countries fall in the range from best practice to worst practice (table 1). Knowing the range and where countries are in the range is important for building the scenarios on the benefits of trade facilitation and for considering which areas of trade facilitation might be most fruitful for a country or for APEC as a whole. Second, correlation matrixes of the indexed inputs into the averages help determine how well the oversampling of surveys works to reduce dependency on a single raw data input while still measuring the relevant trade facilitation concept. Within each trade facilitation indicator the correlation of the indexed inputs is high—above 0.85—suggesting robustness of the trade facilitation indicator with respect to the source of the data. But the fact that the correlations are not 1 indicates that the use of multiple inputs for each trade facilitation indicator is valid.

Trade Flows and Other Variables

Trade data are bilateral trade flows of manufactured goods among APEC member nations from 1989 to 2000. The data come from the Commodity and Trade Database (COMTRADE) of the United Nations Statistics Division. Manufactured goods are defined as commodities in categories 5 to 8 at the one-digit level of the Standard International Trade Classification (Revision 1) except those in category 68 (nonferrous metals), which are at the two-digit level. Trade flow data

2. So an indexed input for APEC member $J(J = 1, 2, 19)^2$ is constructed as: $\overline{II}_J = II_J/(\sum_{J=1}^{19} II_J/19)$ where II_J denotes the raw data for APEC member J, where II_J denotes the raw data for APEC member J.

3. The statistical properties of the trade facilitation indicators may require further consideration. The raw data come from different metrics (percent, survey ranges from 1 to 7 or 1 to 10, numbers of users). So the standard deviations around the mean of each of these indicators will differ from the standard deviation of the indexed inputs that they become. When averaged into the trade facilitation indicator, the standard deviation of the final product and its relationship to the standard deviation of the original data are unclear. What implication this has for using the trade facilitation indicators for estimation in the gravity model is also unclear.

TABLE 1. Summary Statistics for Values of Trade Facilitation Indicators

Indicator and indexed inputs	Source	SD	Min	Economy	Max	Economy
Port efficiency						
Port Efficiency Index (higher is better)	Clark and others 2002	0.284	0.612	Philippines	1.482	Singapore
Ports (higher is better)	World Econ. Forum 2000	0.264	0.617	Philippines, Vietnam	1.447	Singapore
Air transport (higher is better)	World Econ. Forum 2000	0.216	0.688	Peru, Vietnam	1.319	Singapore
Aggregate index		0.248	0.658	Philippines	1.416	Singapore
Customs environment						
Irregular payments (higher is fewer)	World Econ. Forum 2000	0.324	0.464	Russia	1.372	New Zealand
Import fees (higher is fewer fees)	World Econ. Forum 2000	0.359	0.569	Russia	1.821	Singapore
Hidden import barriers (higher is fewer barriers)	World Econ. Forum 2000	0.267	0.461	Indonesia	1.384	Hong Kong, China
Improper practices (higher is better administration)	IMD 2000	0.566	0.142	Russia	1.779	Singapore
Corruption Perceptions Index (higher is less corruption)	Transparency Int'l 2001	0.467	0.343	Indonesia	1.694	New Zealand
Aggregate index		0.375	0.456	Russia	1.590	Singapore
Regulatory environment						
Effectiveness of regulations	World Econ. Forum 2000	0.190	0.748	Vietnam	1.402	Singapore
Regulatory standards	World Econ. Forum 2000	0.235	0.628	Vietnam	1.342	United States
Compliance with agreements	World Econ. Forum 2000	0.183	0.683	Peru	1.256	Singapore
Enforcement of regulations	World Econ. Forum 2000	0.250	0.638	Philippines	1.448	Singapore
Aggregate index		0.207	0.735	Philippines	1.335	Singapore
E-business						
E-commerce (% business use)	World Econ. Forum 2000	0.305	0.461	Russia	1.683	United States
Aggregate index		0.306	0.460	Russia	1.680	United States

Note: Mean equals 1.0. SD, standard deviation; Min, minimum; Max, maximum.

Source: Authors' computations based on data from indicated sources.

over a decade are used to ensure that data points exist for all the study countries because trade flow data can be sparse for some years for some countries.

The data for gross national product (GNP) and per capita GNP come from the World Bank's *World Development Indicators* (World Bank various years). Tariff data were derived from the Trade Analysis and Information System (TRAINS) of UNCTAD. Average applied tariff rates are used, weighted by the values of bilateral trade. The applied tariff rates are the most favored nation (MFN) tariff rates when MFN status applies or preferential rates, if available, when there are preferential trading agreements between two countries. For years when no data are available on preferential rates, the MFN rate is applied. Applied tariff records are sparse. To avoid a significant loss of observations, the applied rates are linearly interpolated or extrapolated over the period 1989–2000 for a given pair of importing and exporting countries when records for at least two years are available.

III. The Econometric Model and Results

Developed by Tinbergen (1962) and Pöyhönen (1963) to explain bilateral trade flows by trading partners' GNP and geographic distance between countries, the gravity model is a common approach to modeling bilateral trade flows. Recent theoretical and empirical work supporting this modeling approach includes Evenett and Keller (1998), Feenstra and others (1998), and Frankel (1997). Besides GDP and distance, other factors relevant for bilateral trade may include population, GDP per capita (to account for intraindustry trade effects that may be associated with countries of similar incomes but varied tastes), regional trade arrangements, and language or ethnic similarities.

Some studies add additional structural elements to the gravity model to better reflect real-world observations. These concern mainly the heterogeneity of traded goods in quality and price by origin and price differentials associated with border and transportation costs. Anderson (1979) develops a gravity model in line with a general equilibrium framework, incorporating consumer preferences for goods differentiated by region of origin, assuming a constant elasticity of substitution structure on consumer preferences. Anderson and von Wincoop (2001) additionally introduce border costs as premiums on the export prices. Balistreri and Hillberry (2001) further extend those results to estimate transport and border costs separately by distinguishing consumer and producer price indices. Using a standard specification of the gravity model, Otsuki and others (2001a, 2001b) control for differences in the prices and unobservable factors that are specific to exporting countries by allowing fixed effects for exporting countries. Though somewhat crude, such a model is less data demanding and more applicable for developing economies whose price data are less reliable and complete.

The model reported here uses the key economic variables of the gravity model, such as GNP and the geographic distance between corresponding pairs of importing and exporting countries, and augments the standard gravity model

specification with the various indicators of trade facilitation. In the general specification of the gravity model, the log of bilateral trade flows in real value is regressed on logs of GNP of exporters and importers, geographic distance between each pair of importers and exporters, and other variables that can account for the rest of the variation (Maskus and others 2001). The model used here employs the specification of the exporter-specific fixed effects developed in Otsuki and others (2001a).

The Gravity Model Analysis

The basic structure of the specific gravity equation is:

$$
\begin{aligned}
(1)\quad \ln(V_{IJ}^t) = {} & b_1\ln(100 + TARIFF_{IJ}^t) + b_2\ln PE_I + b_3\ln CE_I + b_4\ln RE_I \\
& + b_5\ln EB_I + b_6\ln(GNP_I^t) + b_7\ln(GNP_J^t) + b_8\ln(GNPPC_I^t) \\
& + b_9\ln(GNPPC_J^t) + b_{10}\ln(DIST_{IJ}) + b_{11}D_{NAFTA} + b_{12}D_{ASEAN} \\
& + b_{13}D_{LAIA} + b_{14}D_{ENG} + b_{15}D_{CHN} + b_{16}D_{SPN} + b_{17}D_{ADJ} + e_{JI}^t
\end{aligned}
$$

where the b terms are coefficients, I is the importer and J the exporter, and t denotes trading years ($t = 1989, \ldots 2000$). The value of manufactures exports from country J to country I is denoted as V_{IJ}. The term $TARIFF_{IJ}^t$ denotes the applied ad valorem tariff specific to trading partners I and J in year t. The inclusion of the tariff variable is useful for reducing omitted variable biases. It is particularly important for APEC because, unlike the EU's harmonized tariffs, applied tariff rates generally vary across member countries and possibly across their exporting partners.

The terms PE_I, CE_I, RE_I, and EB_I denote importing country I's indicators of port efficiency, customs environment, regulatory environment, and e-business usage. These indicators were constructed from data sources with base year either 1999 or 2000. A time series for these indicators is not available. Positive signs are expected for *PE*, *CE*, and *EB*, but the sign for *RE* is ambiguous because of counteracting effects—the increased transparency of importer's regulations tend to encourage trade, but more stringent regulations might discourage trade.

The terms *GNP* and per capita GNP, *GNPPC*, are both expressed in 1995 U.S. dollars. Geographic distance between capital cities I and J is denoted as DIST_{IJ}. Dummy variables capture the effect of preferential trade arrangements, language similarity, and adjacency. Dummy variables are included for three trade arrangements: North American Free Trade Agreement (D_{NAFTA}), Association of Southeast Asian Nations (D_{ASEAN}), and Latin American Integration Association (D_{LAIA}). The language dummy variables include English (D_{ENG}), Chinese (D_{CHN}), and Spanish (D_{SPN}). The dummy variables for trade arrangements and language assume a value of 1 if both countries are part of the same agreement or both speak the same language. The adjacency dummy variable, D_{ADJ}, takes the value of 1 if country I shares a land border with country J and 0 otherwise.

The error term e_{JI}^t is defined as

$$e_{JI}^t = \alpha_J + \gamma^t + \varepsilon_{JI}^t \tag{2}$$

a composite of exporting country fixed effects, α_j, such as variations in trade flows due to the unobserved difference in quality of goods, domestic policies, and trade facilitation measures in exporting countries; time-specific fixed effects, γ^t; and the random error term, $\varepsilon_{JI}{}^t$, which is assumed to be normally distributed with mean 0. A complete specification would include fixed effects for both exporters and importers (see Moenius 2000, for example). However, incorporation of fixed effects causes a technical problem when at least one of the explanatory variables is invariant within groups for which a cross-section panel is formed (Wooldridge 2002). Importer-specific fixed effects are not included in the model because for a given importing country the trade facilitation measures are invariant over exporting partners and years. The problems associated with the time invariance of these measures is will be discussed shortly.

Regression Results

The regression results indicate that the approach used here—generating a set of distinct trade facilitation indicators and deploying them in a gravity model of trade—is generally successful (table 2). In the first specification (Model I) the coefficients for the four trade facilitation measures are generally significant, and all are of the expected sign. The remaining results reported in table 2 explore the sensitivity of the estimated coefficients to the definition of the trade facilitation indicators and to the functional form of the gravity equation.

Alternative Specifications

Model II relaxes the normalization of raw inputs to mean of 1, because this transformation results in different ranges across the indexed inputs (table 2). The ranges of the values of raw inputs are identically set, and the values of the raw inputs are adjusted for the new range and then added without normalization. This revised approach affects the customs indicator the most, because the raw inputs come from three sources. The signs and significance of the coefficients are generally maintained, but notable changes are observed on the coefficient for customs environment. This implies that the inputs to the customs environment indicator are sensitive to the relative ranges of their values.

Models III and IV examine alternative functional forms. The double-log specification is consistent with the typical specification of gravity models, but this transformation can stretch the scale of a raw input at the bottom and compress it at the top. Model III estimates a linear relationship between trade flows and the explanatory variables and yields the elasticities of trade flow with respect to the explanatory variables (except for the dummy variables). The signs and significance

TABLE 2. Gravity Model Results for Regression of Trade Flows on Trade Facilitation Indicators and Other Standard Variables

Variable	Model I Coefficient	Model II Coefficient	Model III Elasticity	Model IV Elasticity
Constant	−81.790***	−96.401***		
	(8.465)	(7.898)		
Tariff	−0.749**	−0.028**	−1.975***	−0.037***
	(0.375)	(0.014)	(0.713)	(0.008)
Port efficiency	4.200***	4.924***	1.344***	2.613***
	(0.219)	(0.254)	(0.465)	(0.128)
Customs environment	0.422**	−0.222	1.946***	0.388***
	(0.169)	(0.319)	(0.458)	(0.098)
Regulatory environment	−1.562***	−0.888***	−5.913***	−1.601***
	(0.308)	(0.317)	(0.601)	(0.182)
E-business	0.631***	1.363***	1.225***	0.515***
	(0.094)	(0.174)	(0.225)	(0.054)
GNP of importing country	0.846***	0.801***	19.589***	9.021***
	(0.021)	(0.022)	(1.509)	(0.192)
GNP of exporting country	3.870***	3.950***	−7.550	33.362***
	(0.521)	(0.517)	(28.374)	(4.688)
Per capita GNP of importing country	−0.376***	−0.346***	0.395	−0.519***
	(0.041)	(0.039)	(0.762)	(0.057)
Per capita GNP of exporting country	−1.906***	−1.962***	13.574	−2.422**
	(0.679)	(0.678)	(12.186)	(1.001)
Geographic distance	−0.687***	−0.671***	−2.375***	−1.051***
	(0.027)	(0.027)	(0.477)	(0.039)
NAFTA membership dummy variable	0.794***	0.784***	7.657***	1.104***
	(0.164)	(0.165)	(0.377)	(0.112)
ASEAN membership dummy variable	0.712***	0.695***	−0.292	0.414***
	(0.096)	(0.096)	(0.217)	(0.066)
LAIA membership dummy variable	1.624***	1.726***	−0.333	0.957***
	(0.279)	(0.279)	(0.639)	(0.191)
English language dummy variable	0.290***	0.244***	0.168	0.163***
	(0.075)	(0.076)	(0.174)	(0.052)
Chinese language dummy variable	1.138***	1.105***	0.589	0.945***
	(0.189)	(0.190)	(0.427)	(0.130)
Spanish language dummy variable	2.284***	2.339***	0.689*	1.506***
	(0.168)	(0.169)	(0.390)	(0.115)
Adjacency dummy variable	0.162	0.186	3.343***	0.353***
	(0.128)	(0.128)	(0.293)	(0.088)
Box-Cox parameter				0.089
Number of observations	3,304	3,304	3,304	3,304
Adjusted R^2	0.865	0.866	0.536	0.888

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Note: A fixed-effects model with respect to exporting countries and years is used for all models. Numbers in parentheses are standard errors. Model I is log-log with normalized indicators; model II is log-log with unnormalized indicators; model III is nonlog with normalized indicators; and model IV is Box-Cox transformation with normalized indicators.

Source: Authors' computations based on survey data for trade facilitation indicators, COMTRADE for trade flows, UNCTAD TRAINS for tariffs, and World Bank *World Development Indicators* for GNP.

are consistent with those from the double-log specification for the trade facilitation indicators, but the magnitudes of coefficients are very different.

Model IV explores functional form further, using the more flexible Box-Cox transformation, which allows a nonlinear intermediate specification between linearity and logarithm. The Box-Cox estimated relationship is found to be sufficiently close to the double-log form as the Box-Cox parameter is estimated to be 0.089.[4] The estimated elasticities are also found to be similar to those for the double-log. Thus, although these results imply that functional form matters, the estimates associated with the specification in equation 1 are acceptable compared with the alternative specification.

In model I, the preferred specification, the estimated coefficients differ for the different trade facilitation indicators. From a policy perspective, these differences in estimated elasticities of trade flows with respect to the trade facilitation indicators implies that different approaches to trade facilitation will differentially affect exports of individual countries and of the APEC region as a whole.

Overall, the analysis reveals that trade facilitation involves more than reducing the cost of transportation, although this factor is important. The results indicate that other empirical research on quantifying the benefits of trade facilitation that used transport costs as a proxy for trade facilitation likely underestimated the elasticity of trade with respect to broad trade facilitation efforts. This is an important first consideration for policymakers as they identify trade and development priorities in the future. As expected, tariffs have a significant and negative effect on intra-APEC manufactures, as does distance. The coefficients on these two variables are both about 0.7 (slightly higher for customs and slightly lower for distance).

Port efficiency has the largest elasticity among the trade facilitation indicators, at approximately 4.2, suggesting that the greatest gains to intra-APEC manufactures trade would come from improvements in this area. Such a high elasticity of trade with respect to port functions is supported by internal analyses reported by Hong Kong (China) and Japan at a Trade Facilitation Seminar in Bangkok, Thailand (APEC 2002). Fink and others (2002a) also support this finding in the context of maritime-based trade. In sum, the fact that trade is more elastic with respect to direct border costs than to indirect costs appears reasonable.

Customs environment is positively associated with intra-APEC manufactures trade. The coefficient is not large (0.42) relative to the port efficiency parameter. Equal-sized improvements in the customs environment will complement port improvement, but the additional effect would be relatively small overall. On the other hand, improvements in customs can make up for less improvement in tariff barriers. Moreover, the range of potential for country performance in the area of customs is large (for example, see Russia and Indonesia in table 1), suggesting

4. A common Box-Cox parameter is used for the dependent and trade facilitation variables. See Greene (1993:39) for the specification and properties of the Box-Cox transformation.

opportunities in some countries for great improvements in this area compared with improvements in the ports indicator. This potential should raise the profile of this trade facilitation indicator in policy discussion in those countries.

Regulatory environment has a negative and significant effect on intra-APEC manufactures trade with a coefficient of −1.56. To the extent that regulations are used as border barriers, reducing these regulations will be positively associated with increased trade flows. The relatively large coefficient suggests the costly consequences for trade of nonmarket barriers to trade. The large absolute value of the coefficient points out that tightening regulations can offset improvements in other trade facilitation measures.

E-business usage has a positive and significant effect on intra-APEC manufactures trade, with a coefficient of 0.63. This result is consistent with the findings in Fink and others (2002b) and in Freund and Weinhold (2000) that good telecommunications and greater access to the Internet could increase bilateral trade flows. The range of performance among APEC members on this measure of trade facilitation is the largest among the trade facilitation indicators (table 1). So the opportunities for increased trade from improvements in this indicator are large. These results would tend to support efforts within APEC to enhance e-commerce usage through the e-APEC Strategy and Paperless Trading initiatives.

Caveats and Robustness of the Specifications

Certain caveats should be noted, however. The analytic approach applied here is designed to overcome limited data availability. For trade flows, tariffs, and GDP 12 years of data were used to overcome the sparse availability of trade flow and tariff data for the studied countries, whereas the trade facilitation indicators are available only for a single year. This mismatch creates measurement errors, because the values of these indicators have probably changed over time. Accordingly, the estimated coefficients are less likely to represent true elasticities of trade flow for the earlier period. The use of interpolated tariff rates for the years in which data are not available may also result in measurement errors if tariff changes were not linear. The estimated coefficients also may be subject to specification bias because observations with zero trade are omitted from the sample.

The robustness of the estimation results is examined by comparing the regression results with alternative specifications (table 3). First, all the time-variant variables (value of trade, tariff, GNP, and per capita GNP of the importing and exporting country) are averaged for 1989–2000 to match these variables to the trade facilitation variables. The signs and significance of the coefficients for the tariff and some of the trade facilitation indicators are different from those in the regressions with unaveraged data. The tariff coefficient is positive and significant, and the coefficients for customs environment and regulatory environment are insignificant (table 3, second column).

Compared with the trade facilitation measures, the 12-year-averaged time-variant variables are still subject to measurement error. To match the time period

TABLE 3. Robustness of the Regression Specification for Trade Flows and Trade Facilitation Indicators

Variable	With averaged data	With data only for 2000	With averaged tariff	With uninterpolated tariff
Constant	−56.196***	−41.473***	−93.291***	−35.334***
	(8.935)	(10.952)	(8.033)	(3.626)
Tariff	5.344***	1.971	0.636	−0.929**
	(1.833)	(2.233)	(0.500)	(0.472)
Port efficiency	2.174**	1.826*	4.178***	3.179***
	(0.889)	(0.994)	(0.220)	(0.273)
Customs environment	0.826	2.629***	0.141	0.581***
	(0.684)	(0.722)	(0.178)	(0.219)
Regulatory environment	−1.295	−3.967***	−1.147***	−1.875***
	(1.236)	(1.322)	(0.325)	(0.413)
E-business	0.761*	0.391	0.719***	0.484***
	(0.053)	(0.428)	(0.097)	(0.117)
GNP of importing country	0.808***	1.026***	0.819***	0.868***
	(0.089)	(0.101)	(0.022)	(0.026)
GNP of exporting country	1.082***	1.120***	4.131***	1.359***
	(0.061)	(0.069)	(0.512)	(0.115)
Per capita GNP of importing country	−0.053	−0.221	−0.307***	−0.312***
	(0.172)	(0.195)	(0.042)	(0.053)
Per capita GNP of exporting country	0.171***	0.033	−2.167***	0.250**
	(0.065)	(0.073)	(0.673)	(0.129)
Geographic distance	−0.790***	−1.119***	−0.631***	−0.650***
	(0.103)	(0.121)	(0.026)	(0.035)
NAFTA membership dummy variable	0.044	0.968	0.786***	1.177***
	(0.661)	(0.805)	(0.163)	(0.205)
ASEAN membership dummy variable	1.702***	1.337***	0.964***	0.246**
	(0.372)	(0.447)	(0.094)	(0.123)
LAIA membership dummy variable	0.386	0.142	1.710***	1.466***
	(1.143)	(1.228)	(0.282)	(0.363)
English language dummy variable	0.503*	0.697**	0.293***	0.312***
	(0.278)	(0.334)	(0.074)	(0.092)
Chinese language dummy variable	1.438***	1.161***	1.063***	1.034***
	(0.317)	(0.369)	(0.167)	(0.244)
Spanish language dummy variable	0.689	0.336	2.395***	2.172***
	(0.674)	(0.728)	(0.170)	(0.222)
Adjacency dummy variable	0.540	0.662	0.147	0.112
	(0.509)	(0.733)	(0.123)	(0.168)
Number of observations	303	279	3413	1949
Adjusted R^2	0.746	0.725	0.864	0.874

Note: A fixed-effects model with respect to exporting countries and years is used for all models. Numbers in parentheses are standard errors.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Source: Authors' computations based on survey data for trade facilitation indicators, COMTRADE for trade flows, UNCTAD TRAINS for tariffs, and World Bank *World Development Indicators* for GNP.

exactly, the specification was run using data only for 2000. Observations are lost for some country pairs, but the measurement error with respect to trade facilitation measures is avoided. These results suggest a greater importance for customs and regulatory environment than in the initial specification reported in table 2. This result appears reasonable, given that many of these countries have undertaken improvements in their port facilities over the 12-year period (third column).

A second robustness check examined measurement error due to interpolation of tariff rates, using two alternative specifications of tariff rates: one with averaged tariff rates instead of interpolated tariff rates over the period 1989–2000 (fourth column) and one with uninterpolated tariff rates (fifth column). With averaged tariff rates the coefficient for tariff rates is positive but insignificant, and coefficients on the trade facilitation variables are of about the same magnitude as in the initial specification in table 2, although the customs variable is not significant. With uninterpolated tariff rates the tariff variable is negative and significant, although the trade facilitation variables remain robust and their coefficients are of approximately the same size as those in the table 2 specification. It is not clear from these results whether the interpolated or averaged tariff rates have introduced measurement error or simply missing data. The results suggest overall, however, that the trade facilitation variables are robust to this change in the tariff specification.

Finally, the estimated coefficients may be biased due to the sample selection bias that results from omitting observations with zero trade (see, for example, Wall 2000). Downward bias is likely for the coefficients on the trade facilitation measures because observations with zero trade caused by poor conditions of trade facilitation, other things being equal, are ignored. The implications of this selection bias could not be examined because the source data do not distinguish zero trade from missing records.

Cross-section regression analysis inevitably involves ambiguous causal relationships. In addition, use of a set of trade facilitation measures for a single year limits the interpretation of the coefficients as elasticities. The possibility cannot be excluded that greater bilateral trade will lead to higher values of trade facilitation measures rather than to the reverse relationship that was postulated in the estimation. Port efficiency, customs environment, and e-business usage may improve with a country's import flows, and if this endogeneity is present the estimated coefficients for these variables would be biased upward.

A logical approach to the endogeneity problem when time-series data for the explanatory variable of interest are unavailable is to employ instrumental variables for the trade facilitation variables so that the error term does not correlate with the trade facilitation measures. This approach requires instrumental variables that are exogenous to the trade facilitation measures and trade flow. But such instruments are difficult to find, and their power is uncertain. Moreover, the endogeneity problem remains if instruments that best account for the state of trade facilitation are likely to be dependent on trade flows. The use of instruments is consequently not an effective solution to endogeneity. Time series

for the raw inputs to trade facilitation indicators may become available in the future, which should mitigate the endogeneity problem.

IV. Potential Benefits from Trade Facilitation: Simulation Results

This section examines various scenarios of improved trade facilitation and tariff reduction, focusing on improved port efficiency, customs environment, e-business usage, and regulatory harmonization. The objective is to inform policymakers about specific trade facilitation initiatives with the greatest potential to increase trade and economic well-being.

Simulation Design

The simulations using the gravity model and trade facilitation indicators provide three perspectives on trade facilitation in APEC. First, the simulations present the implications of different trade facilitation initiatives for intra-APEC trade as a whole. Second, the simulations permit examination of an individual APEC member's exports to other APEC members (bilateral and total). Finally, the simulations can provide a proxy for the costs to businesses and consumers in an individual APEC country when its trade facilitation indicators are lower than APEC best practice.

The simulation methodology of applying a common percentage improvement to each trade facilitation indicator implies that even an economy that is already using best practice will also have to improve. Instead, an approach is used that acknowledges the differential potential for improvement revealed by table 1, to better inform policy decisions on the kind of trade facilitation initiative that might yield the greatest improvements. To that end the goal is to identify improvements that bring below-average members halfway to the APEC average.[5] The focus is on below-average performers on the grounds that donor attention and capacity-building efforts should be extended to members of this group, which because of their lower scores will have to make greater efforts. The goal of halfway to average recognizes that it is not realistic to presume a scenario in which all APEC members achieve best practice as measured by the APEC member with the highest score on a particular measure of trade facilitation. Not only does the best-practice economy differ by trade facilitation measures, but the range between lowest value and highest value varies significantly, being greatest for e-business usage and customs environment.

Thus the countries for which an improvement in trade facilitation is simulated will differ by trade facilitation indicator. However, because trade facilita-

5. The simulation is a one-time shot. Thus the problem arises that improvements in the below-average APEC members will change the target average. A simulation exercise that considered sequential rounds of reforms using this strategy would have to take into account the endogeneity of the target levels of trade facilitation reforms.

tion links exporters and importers, all economies enjoy an increase in intra-APEC trade even when only some have an improvement in their trade facilitation indicator. Consider this example for Chile and New Zealand. Chile is below average in port efficiency, and the scenario for improvement in that indicator will bring Chile halfway to the APEC average. But Chile is above average for customs environment, so no improvement is postulated for Chile in the scenario for that indicator. New Zealand has above-average trade facilitation indicators for all except e-business usage. Thus only in the scenario of improved e-business will the trade facilitation indicator for New Zealand be improved. However, because Chile and New Zealand trade with each other in APEC, when Chile improves its ports, New Zealand gains. When New Zealand improves its e-business usage, Chile gains.

Simulation Results

For port logistics, customs environment, and e-business usage, the simulation is designed to bring below-APEC-average members halfway to the initial APEC average. For regulatory environment, for which research suggests that standards harmonization increases trade (Moenius 2000; Hertel and others 2001), the simulation brings the above-average members halfway down to the APEC average, as a proxy for how relaxing regulatory barriers increases trade, and the below-average members halfway up to the APEC average, as a proxy for how standards harmonization increases trade.

Together, the simulations yield an increase in intra-APEC trade of approximately $250 billion, or 21 percent of total intra-APEC manufactures trade (table 4). Some $117 billion of the total gain (a 10 percent increase in trade) comes from improvements in port efficiency and $22 billion from improvements in the customs environment. Another $116 billion might come from improvements behind the border in regulatory harmonization and e-business usage.

The large increase in intra-APEC trade derived from improved port efficiency is due partly to the large coefficient on the relationship between trade and port logistics (4.2; see table 2 model I) and partly to the broad room for improvement in countries such as China and Mexico, which are very large intra-APEC traders. Many APEC countries would exhibit double-digit increases in dollar value exports to the APEC region from the measures, with the greatest gains going to large APEC exporters, such as the United States, Japan, and the Republic of Korea (Wilson and others 2003). These scenarios suggest that the attention devoted by policymakers to improvements in port efficiency is warranted. However, the range from best practice to worst practice is smaller for port logistics than for other trade facilitation indicators, suggesting that there are countries in which port efficiency is not the principal bottleneck to trade.

Each APEC member has a unique trade pattern with other members. From an exporting country's perspective, export gains will depend on which APEC countries it trades with and how much improvement its trading partners achieve under a particular trade facilitation scenario. From an importing country's

TABLE 4. Overview of Simulation: Halfway to APEC Average

Trade facilitation measure	Goal	Change in trade flow: Amount ($ billion)	Change in trade flow: Share of total trade (%)
Border measures			
Port efficiency	Bring below-average members up to the APEC average	116.89	9.7
Customs environment	Bring below-average members up to the APEC average	21.63	1.8
Behind-the-border measures			
E-business	Bring below-average members up to the APEC average	27.69	2.3
Regulatory environment	Regulatory harmonization: Bring above average members down to the APEC average, and below-average members up to the APEC average	88.15	7.3
Total		254.36	21.0

Source: Authors' computations based on survey data for trade facilitation indicators, COMTRADE for trade flows, UNCTAD TRAINS for tariffs, and World Bank *World Development Indicators* for GNP.

perspective efficiency gains (measured as increased imports) depend only on the country's own trade facilitation efforts. Wilson and others (2003) simulate trade gains for individual APEC members from these two perspectives.

Because of countries' dependencies on their own trade facilitation indicators and their own trading patterns, trade facilitation measures that generate the greatest gains for an individual APEC member might not be the same as trade facilitation measures that generate the greatest gains for APEC as a whole. For example, Thailand's port efficiency indicator is near the APEC average. A small improvement (which would still cost resources) in the APEC average would increase Thailand's imports by some $4.4 billion. But Thailand's customs environment and e-business usage are much further away from the APEC average. An improvement halfway to the APEC average in customs environment would increase Thailand's imports by $2.4 billion. If the cost of improving customs is much less than the cost of improving port efficiency, then the net gain might be greater from focusing policy efforts on customs than on port efficiency. Even greater gains would result from an improvement halfway to the APEC average in e-business usage—Thailand's imports would increase $7.9 billion, nearly 50 percent more than they would from the two border measures taken together. Therefore, a Thai policymaker might want to consider reforms that enable greater e-business usage. This example shows that selecting the best target for policy effort requires careful attention not only to the estimated coefficient of trade with respect to individual trade facilitation indicators but also to where an economy ranks in the range of APEC economies.

Tariff Reductions or Trade Facilitation?

The regression results also enable comparison of the potential trade gains from improvements in trade facilitation and from tariff reductions. The estimated coefficients point to tradeoffs between the two. Reducing tariffs to 0 is used as the benchmark for evaluating what equiproportionate improvement in each trade facilitation indicator would generate the same amount of gain in total intra-APEC exports. When all members make equiproportionate improvements in trade facilitation indicators, the total gain is 0.55 percent for port efficiency and 5.46 percent for customs environment. When only below-APEC-average members make equiproportionate improvements in the trade facilitation indicator,[6] the gain is 0.83 percent for port efficiency and 27.7 percent for customs environment.

An average reduction of 6.5 percent in the applied ad valorem tariff is needed to reduce to 0 all tariffs on manufactures in all APEC members. The gain in trade flows from this reform would be $27.8 billion. To achieve the same increase through trade facilitation measures the port efficiency indicator would have to be improved by 0.55 percent if all members are included or by 0.83 percent if only below-APEC-average members are included, the customs environment indicator would have to be improved by 5.5 percent for all APEC members or by 27.7 percent for below-average members, and the e-business indicator by 3.7 percent for all members or by 13.2 percent for below-average members. Thus the required improvements in trade facilitation indicators, especially for port efficiency, are relatively small compared with the required tariff reductions to generate similar gains in trade. This implies that trade facilitation can be a good policy alternative if tariff reduction is not feasible.[7]

Scope of the Analysis and Extension

The scope of the analysis was limited to elicit information useful to the policy target. The analysis has focused on improvement in the trade facilitation measures of importing countries. But the estimates from importers' improvement in trade facilitation account for only partial gains in trade flows. The simulation analysis also needs to be performed for exporters to obtain the total gains in trade flows.

The simulations reported herein focus on the relationship between trade facilitation and trade flows. In a public policy dimension, however, information on the cost of investment in trade facilitation is also indispensable for a country

6. A condition is imposed in the second analysis that the value of a trade facilitation indicator of below-average members will not exceed the APEC average after the improvement.

7. This simulation returns to the tradition of an equiproportionate change for each APEC member. As noted, this equiproportionate increase masks significant differences among the member economies in their trade facilitation indicators and also does not exploit these differences to generate greater gains from one indicator over the others.

government to make socially desirable decisions. Although the analysis indicates that improvement in port efficiency has the greatest positive impact on trade over all the APEC members, this may not be the most cost-effective option for an individual economy. The direct costs of improving port infrastructure may be greater than the direct costs of improving the customs environment, for example (although indirect costs of organization change could be higher in customs). In any case, cross-country data and analysis of the cost of investment and reform in each area of trade facilitation would make possible research that complements the findings here to better inform policymaking.[8]

V. Conclusion

The key innovation in the research approach reported here centers on considering a variety of indicators of trade facilitation and of pragmatic simulations suited to policymaking. Collectively, the country-specific trade facilitation indicators embrace the multiple approaches to trade facilitation reflected in modern international commerce. The simulation analysis also considers the importance of focusing on best practices and achieving benchmarks tied to what is known from experience in best practices in trade facilitation. Considered completely separately from any model estimation of their effect on trade, this set of indicators helps policymakers judge where their economy stands relative to their peers on each of these measures. In the context of quantifying the benefits of trade facilitation efforts, this multiple-indicator approach and realistic simulation design, along with decomposition of the impact of the various indicators on trade, may enable more targeted decisionmaking by policymakers.

The simulation approach offers several perspectives on the potential benefits of improvements in trade facilitation. It permits analysis of the implications for intra-APEC trade as a whole. It also allows examination of an individual member's exports to other APEC members and use of the results as a proxy for the costs to an APEC member whose trade facilitation indicators are below best practice. This three-sided analysis can be a particularly valuable input when considering alternative pilot projects for individual APEC members. Of course, the resource costs of alternative policy reforms must be considered to gauge the net gain.

In sum, using this set of indicators and modeling approach offers policymakers more information about what type of trade facilitation efforts might provide the largest gains in terms of increasing trade flows. Whereas it remains true that a comprehensive effort yields the greatest increase in trade, examination of different kinds of trade facilitation and of disaggregated trade flows could be useful for targeting policy efforts and launching pilot projects in capacity building.

8. For a start on the analysis of costs versus benefit, see the case studies in Wilson and others (2002).

Appendix. Data Sources

- World Economic Forum (2000), *Global Competitiveness Report 2000*. All survey data are from the World Economic Forum's Executive Opinion Survey of senior business leaders in 4,022 firms in different countries.
- IMD (2000) Lausanne, *World Competitiveness Yearbook 2000*. The yearbook uses a 115-question survey sent to executives in top and middle management of firms in 49 countries. The sample size of each country is proportional to its GDP, and firms "normally have an international dimension." The firms are selected to be a cross-section of manufacturing, service, and primary industries. There were 3,532 responses to the survey in 2000.
- Transparency International (2001), *The Global Corruption Report*. Transparency International is the only international nongovernmental organization devoted to studying and fighting corruption. Through 80 independent chapters around the world it monitors government compliance, corruption levels, and the transparency of regulations.
- Clark and others (2002) provides data on port efficiency for maritime transport.

The various raw data series were chosen for their relevance to the four concepts of trade facilitation.

- Port efficiency for each APEC member J is the average of three indexed inputs:
  - Port Efficiency Index (1 = worst and 7 = best; Clark and others 2002).
  - "Port facilities and inland waterways are extensive and efficient" (1 = strongly disagree and 7 = strongly agree; World Economic Forum 2000).
  - "Air transport is extensive and efficient" (1 = strongly disagree and 7 = strongly agree; World Economic Forum 2000).
- Customs environment for each APEC member J is the average of five indexed inputs:
  - "Irregular, additional payments connected with import and export permits, business licenses, exchange controls, tax assessments, police protection, or loan applications are very rare" (1 = strongly disagree and 7 = strongly agree; World Economic Forum 2000).
  - "Import fees are high" (1 = strongly disagree and 7 = strongly agree; World Economic Forum 2000).
  - "Hidden import barriers other than published tariffs and quotas are: 1 = an important problem and 7 = not an important problem" (World Economic Forum 2000).
  - "Bribery and corruption exist in the economy" (1 = agree and 10 = disagree, IMD Lausanne 2000).
  - Corruption Perceptions Index (Transparency International 2001).
- Regulatory environment for each APEC member J is constructed as the average of four indexed inputs (World Economic Forum 2000):
  - "Environmental regulations in your country are 1 = confusing and frequently changing and 7 = transparent and stable."

- "Regulatory standards (product, energy, safety, environmental standards) are among the world's most stringent" (1 = strongly disagree and 7 = strongly agree).
- "Compliance with international environmental agreements is a high priority in your country's government" (1 = strongly agree and 7 = strongly disagree).[9]
- "Environmental regulation in your country is: 1 = not enforced or enforced erratically and 7 = enforced consistently and fairly."

• E-business for each APEC member *J* (World Economic Forum 2000):
- "Percentage of companies that use the Internet for e-commerce."

REFERENCES

Anderson, James E. 1979. "A Theoretical Foundation for the Gravity Equation." *American Economic Review* 69:106–16.

Anderson, James E., and Eric van Wincoop. 2001. "Gravity with Gravitas: A Solution to the Border Puzzle." NBER Working Paper 8079. National Bureau of Economic Research, Cambridge, Mass.

APEC (Asia Pacific Economic Cooperation). 2002. "Implementing the APEC Trade Facilitation Principles." Room Documents. Bangkok.

———. 1999. *Assessing APEC Trade Liberalization and Facilitation: 1999 Update, Economic Committee.* Singapore.

Asia Pacific Foundation of Canada. 1999. *Survey on Customs, Standards and Business Mobility in the APEC Region.* Vancouver.

Australian Department of Foreign Affairs and Trade and Chinese Ministry of Foreign Trade and Economic Cooperation. 2001. *Paperless Trading: Benefits to APEC. Commonwealth of Australia.* Online document available at www.dfat.gov.au/publications/paperless/paperless_trading.pdf.

Balistreri, Edward J., and Russell H. Hillberry. 2001. "Trade Friction and Welfare in the Gravity Model: How Much of the Iceberg Melts?" U.S. International Trade Commission, Washington, D.C.

Clark, Ximena, David Dollar, and Alejandro Micco. 2002. "Maritime Transport Costs and Port Efficiency." World Bank Working Paper 2781. World Bank, Washington, D.C.

Evenett, Simon J., and Wolfgang Keller. 1998. "On Theories Explaining the Success of the Gravity Equation." NBER Working Paper 6529. National Bureau of Economic Research, Cambridge, Mass.

Feenstra, Robert C., James R. Markusen, and Andrew K. Rose. 1998. "Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods." NBER Working Paper 6804. Cambridge, Mass.: National Bureau of Economic Research.

Fink, Carsten, Aaditya Matoo, and Cristina Ileana Neagu. 2002a. "Trade in International Maritime Services: How Much Does Policy Matter?" *World Bank Economic Review* 16(1):81–108.

———. 2002b. "Assessing the Role of Communication Costs in International Trade." World Bank Working Paper 2929. World Bank, Washington, D.C.

Frankel, Jeffrey A. 1997. *Regional Trading Blocs in the World Economic System.* Washington, D.C.: Institute for International Economics.

Freund, Caroline, and Diana Weinhold. 2000. "On the Effect of the Internet on International Trade." International Finance Discussion Papers 693. Board of Governors of the Federal Reserve System, Washington, D.C.

Greene, William H. 1993. *Econometric Analysis*, 2nd ed. New York: Macmillan.

9. For indexing, this index value is reversed to make it consistent with the other indexes.

Hertel, Thomas W., Terrie Walmsley, and Ken Itakura. 2001. "Dynamic Effect of the 'New Age' Free Trade Agreement between Japan and Singapore." *Journal of Economic Integration* 16(4):446–84.

Hummels, David. 2001. "Time as a Trade Barrier." GTAP Working Paper. Purdue University, Department of Agricultural Economics, Center for Global Trade Analysis, West Lafayette.

IMD. 2000. *World Competitiveness Yearbook 2000*. Lausanne.

Mann, Catherine L., Sue E. Eckert, and Sarah Cleeland Knight. 2000. *Global Electronic Commerce: A Policy Primer*. Washington, D.C.: Institute for International Economics.

Maskus, Keith E., John S. Wilson, and Tsunehiro Otsuki. 2001. "An Empirical Framework for Analyzing Technical Regulations and Trade." In Keith Maskus and John S. Wilson, eds. *Quantifying the Impact of Technical Barriers to Trade: Can It Be Done?* Ann Arbor: University of Michigan Press.

Moenius, Johannes. 2000. "Three Essays on Trade Barriers and Trade Volumes." Ph.d. diss., University of California, San Diego.

OECD (Organisation for Economic Co-operation and Development). 2001. *Business Benefits of Trade Facilitation*. TD/TC/WP(2001)21 Final. Paris.

Otsuki, Tsunehiro, John S. Wilson, and Mirvat Sewadeh. 2001a. "What Price Precaution? European Harmonisation of Aflatoxin Regulations and African Groundnut Exports." *European Review of Agricultural Economics* 28(3):263–84.

———. 2001b. "Saving Two in a Billion: Quantifying the Trade Effect of European Food Safety Standards on African Exports." *Food Policy* 26(5):495–514.

Pöyhönen, P. 1963. "A Tentative Model for the Volume of Trade between Countries." *Welwirtschaftliches Archiv* 90(1):93–99.

Tinbergen, J. 1962. *Sharing the World Economy: Suggestions for an International Economic Policy*. New York: Twentieth Century Fund.

Transparency International. 2001. *The Global Corruption Report*. Berlin. Online document available at www.globalcorruptionreport.org/gcr2001.shtml.

UNCTAD (United Nations Conference on Trade and Development). 2001. *E-Commerce and Development Report*. Geneva.

Wall, Howard J. 2000. "Gravity Model Specification and the Effects of the Canada–U.S. Border." Federal Reserve Bank of St. Louis Working Paper Series 2000-024A. St. Louis.

Wilson, John S., Catherine Mann, Yuen Pau Woo, Nizar Assanie, and Inbom Choi. 2002. *Trade Facilitation: A Development Perspective in the Asia-Pacific Region*. Singapore: Asia Pacific Economic Cooperation.

Wilson, John S., Catherine Mann, and Tsunehiro Otsuki. 2003. "Trade Facilitation and Economic Development: Measuring the Impact." World Bank Working Paper 2988. World Bank, Washington, D.C.

Woo, Yuen Pau, and John S. Wilson. 2000. *Cutting through Red Tape: New Directions for APEC's Trade Facilitation Agenda*. Vancouver: Asia Pacific Foundation of Canada.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass., and London: MIT Press.

World Bank. Various years. *World Development Indicators*. Washington, D.C.

World Economic Forum. 2000. *Global Competitiveness Report 2000*. Geneva.

Water Subsidy Policies: A Comparison of the Chilean and Colombian Schemes

Andrés Gómez-Lobo and Dante Contreras

Analysis of two water subsidy schemes—a means-tested subsidy in Chile and a geographically targeted subsidy in Colombia—shows that the means-tested system is better able to identify poor households than the geographically targeted scheme. However, the overall distributive impact of both schemes is quite similar, at least for the three lowest income deciles, because the amount of benefits per household in the geographically targeted Colombian scheme are differentiated by the socioeconomic classification of household. Despite the relative merits of the Chilean means-tested scheme, targeting errors are still quite large. More than 60 percent of subsidies accrue to households that are above the third decile of the income distribution. If the policy objective in Chile is to benefit a significant proportion of households in the lowest income deciles, then either the targeting mechanism must be improved or the number of subsidies has to increase to take into account these targeting imperfections. In Colombia almost all households receive some kind of benefit, implying an unnecessarily high fiscal cost. An improvement in the targeting mechanism could lower this cost without jeopardizing benefits to lower-income households. Some suggestions for additional research and for improving both schemes are discussed.

In recent years many countries have undertaken important reforms of their electricity, gas, water, and telecommunications industries to increase productive and allocative efficiency and promote higher levels of investment. However, the reforms have sometimes created new social and distributive concerns. For example, the need to guarantee the financial viability of firms usually entails an increase in household utility bills. Reforms usually aim to improve the commercial efficiency of providers, which, if successful, reduces the level of nonpayment by certain groups of customers. Reforms normally include an end

Andrés Gómez-Lobo is Assistant Professor, Department of Economics, at the University of Chile; his e-mail address is agomezlo@econ.uchile.cl. Dante Contreras is Associate Professor, Department of Economics, at the University of Chile; his e-mail address is dcontrer@econ.uchile.cl. The authors thank Jose Luis Lima and Victor Macias for their very able research assistance. They are also indebted to Francisco Socias from Chile's Ministry of Social Planning and Elena Peralta from Chile's Ministry of the Interior for information provided during several interviews. The authors also thank Antonio Estache and Vivien Foster, three anonymous referees, and especially François Bourguignon for comments on an earlier draft of this article. Financial support for this study from the World Bank Institute is gratefully acknowledged. This project forms part of the efforts of the World Bank Institute to research the relation between poverty and utility services in Latin America.

THE WORLD BANK ECONOMIC REVIEW, VOL. 17, NO. 3 391–407

DOI: 10.1093/wber/lhg028

to cross-subsidies. All of these policies, although recommended from an economic point of view, may increase the financial burden on poorer households.[1]

Conventional wisdom indicates that such social and distributive effects should be tackled through the general tax and benefits system and not through sectoral reform programs. However, there are several reasons why this may not always be a wise strategy in developing economies. First, policymakers may be concerned that all households consume a minimum level of utility services as well as with the income effects of utility bills. This is especially so in the case of water and sanitation services, which involve public health issues. Second, many countries have no established welfare system capable of providing an income supplement to poorer households to compensate for rising utility prices. Finally, ignoring social issues in utility industries may introduce distortions in regulatory decisionmaking that, among other effects, may increase the risks faced by potential private operators in these industries.[2]

Thus, in developing economies general welfare payments that increase total disposable household income may not always be a good or viable policy substitute for programs designed to allow poor households to access and consume specific utility services. In these cases, welfare policies specifically designed for these industries may be warranted. However, there are several risks associated with such a sectoral approach. Perhaps the most important is that such subsidies may be poorly targeted.[3] Informal subsidies benefiting mainly middle-class customers abound in utility industries throughout the developing world. Therefore, if a utility service is to be subsidized, the distributive effects of such a scheme need careful attention.

This article analyzes the distributive effects of two subsidy programs for water utility services applied in developing economies: a program introduced in Chile in the early 1990s and one adopted in Colombia after 1994. Both provide explicit subsidies targeted to vulnerable households, and both were introduced as part of wider sectoral reforms.

There are several reasons why these two programs are of interest to policymakers. First, they constitute two of the few examples in the developing world of formal subsidies for utility services. Both programs are national in coverage, benefit customers from different regional operators, and are grounded in an explicit legal and regulatory framework. Moreover, both programs tried to target benefits to poor people. As such, these schemes offer an interesting contrast to the informal and universal subsidies more commonly encountered

1. Estache, Foster, and Wodon (2002) present a comprehensive analysis of the potential links between infrastructure reforms and poverty.

2. In extreme cases, neglect of social considerations may lead to civil unrest—as in Cochabamba's (Bolivia) water concession process a few years ago—or the termination of existing concession contracts, as was the case of the water concession in Tucuman, Argentina, in the 1990s.

3. This will be the case, for example, when a significant proportion of poor households is not connected to a utility service whose consumption is being subsidized.

in many utility industries.[4] The Chilean and Colombian experiences are often touted as examples of best practice in the design of subsidy schemes in infrastructure services.[5] However, so far little empirical analysis has been undertaken to evaluate the results of these programs. This article is a first step in that direction.

Second, the Chilean and Colombian schemes use different targeting mechanisms. Chile uses individual means testing of households to determine eligibility, whereas Colombia uses a geographical targeting system. Because both targeting mechanisms are valid design options, policymakers may be interested in knowing how each type of scheme performs in particular applications.

Relative concentration curves typical of distributional analysis are used to show that these two very different targeting schemes have comparable results in terms of reaching poor households, at least for monetary transfers. However, if the number of beneficiaries rather than monetary transfers are used in the concentration curve analysis, the Chilean means-tested scheme performs better than the Colombian geographically targeted scheme. Therefore, there is some evidence that a means-tested subsidy, as applied in Chile, is better able to identify poorer households than a geographically targeted subsidy. However, because the Colombian system is more generous to comparably poor households than the Chilean system, the overall effect on poverty incidence is similar in both schemes.

An interesting contribution of this article is the use of an absolute beneficiary concentration curve to further analyze the targeting properties of these subsidies. With this curve it is straightforward to evaluate both the errors of exclusion (eligible households that do not receive a benefit) and the errors of inclusion (noneligible households that receive a benefit) associated with each program. The Colombian scheme provides benefits to almost all poor households but at a high fiscal cost because the program gives some benefits to almost all households in the country. The Chilean program, which is much smaller and targeted to a narrower group of households, misses many deserving households because of targeting errors.

The next two sections provide a brief overview of the main features of each subsidy scheme. This is followed by a comparison of the targeting properties of each subsidy using relative and absolute concentration curve analysis of the total monetary transfers and the total number of beneficiaries. Some policy implications are also drawn from the analysis.

4. These subsidies usually take the form of direct transfers from local or national governments to loss-making public utilities. According to the World Bank (1994), the average ratio of revenues to costs for public utilities across the world was 0.8 for gas, 0.6 for electricity, 0.3 for water, and 1.6 for telecommunications. Thus, except for the telecommunications sector, informal universal subsidies of questionable distributional impact were very common, at least at the beginning of the past decade.

5. The Chilean water subsidy was positively reviewed in a recent book on the design of subsidies for public services (Brook and Smith 2001). Both experiences are presented as best practice in a popular training course for regulators at the University of Florida and sponsored by the World Bank. The Latin American version of this course, undertaken by the Centre for the Economic Study of Regulation in Buenos Aires, also presents these experiences as best practice in this field.

I. The Chilean Water Subsidy Scheme

The water consumption subsidy in Chile is one of the few means-tested subsidies (households are individually screened for eligibility based on their socioeconomic circumstances) applied in the utility industry of a developing country. The subsidy scheme, which became operational in 1990, was designed to counter the adverse social impacts of rising water charges.[6]

The subsidy program is administered by the Ministry of Social Planning along with municipal governments. The ministry determines the number of subsidies that will be offered to each region the following fiscal year and the value of each subsidy by region. Municipalities are responsible for administering the subsidies.

The subsidy program gives eligible households the right to consume water at a lower price (expressed as a percentage of the full tariff) up to a certain limit.[7] Beyond this limit additional consumption is charged at the full tariff. Thus for eligible households the subsidy operates like a rising block tariff structure that subsidizes the first block of consumption. The Ministry of Social Planning determines the consumption ceiling that the subsidy will cover (currently 15 m^3 per month in all regions) and the percentage discount for this first block (which varies by region).

The number and value of each subsidy by region are determined yearly, and the aggregate projected expenditure of the program is included in the national budget each fiscal year. The subsidy is funded entirely from general tax revenues, and the water regulator—responsible for setting tariffs—is not involved in determining subsidy levels or in the operational aspects of the scheme. Thus there is complete separation between the welfare policies applied in the water sector and the economic regulation of the industry.[8]

To be eligible for a subsidy, a household must apply for the benefit at its municipality.[9] Eligibility is then determined by socioeconomic need, based on the Communal Social Assistance Committees (CAS) point score. The CAS score is the main instrument used in the Chilean welfare system to gauge the socioeconomic condition of households. A household's CAS score is based on information gathered during a personal interview conducted at the household's dwelling. The interview consists of 50 questions divided in nine sections: general information, environmental conditions, overcrowding conditions, health conditions, identification of family members, occupation and earning, monetary subsidies, education, and wealth. Personal interviews are conducted at the household's dwelling so that the interviewer can verify some of the answers,

6. Between 1989 and 1998 water charges more than doubled for each of the 13 regional water companies—and more than tripled in some regions, such as the water-scarce northern regions.

7. By law the percentage of the bill that the subsidy will cover can vary between 20 percent and 85 percent.

8. However, there is some dependence in the other direction. A tariff revision process will have a direct impact on the design and projected budget of the subsidy program in the following years.

9. To qualify for a subsidy, a client cannot have arrears with the service provider.

such as the size and materials of the dwelling and durable goods ownership, key variables in the CAS score calculation.

A household's CAS score is valid for two years. Although this lowers the administrative costs of the welfare system in general, it also reduces the efficiency of the CAS score as a targeting instrument because a household's socioeconomic condition may change in the course of two years.

The municipalities must award water subsidies (as well as other public subsidies) based on the ranking of CAS scores of all new applicants and current beneficiaries. The total number of subsidies awarded is capped by the total number assigned annually to the region by the Ministry of Social Planning and later distributed to each municipality by the regional governor. If all subsidies have already been assigned and a new applicant has a lower CAS score than the best-off current beneficiary, the mayor must redistribute benefits in favor of the new applicant.

Once a household is awarded a subsidy, the service provider is notified and the client's subsequent monthly water bills will be net of the subsidy. Once a month water companies bill municipalities in their service area to recover the subsidies given to beneficiary households in the last billing period.

Data on the number of assigned subsidies by region, the percentage of connected households covered, and the average value of subsidies in 1998 show that nearly a third of households are subsidized in some regions (table 1). The absolute monthly value of the subsidy is highest in the northern regions (I to III), where

TABLE 1. Number and Value of Assigned Water Subsidies by Region in 1998, Chile

Region	Number of assigned subsidies	Share of regional households covered[a] (%)	Average value per subsidy[b] (US$/month)
I	18,836	20.7	15.74
II	24,576	22.9	17.94
III	16,869	28.3	13.41
IV	21,793	17.9	9.14
V	63,034	17.3	6.36
VI	17,108	10.7	5.95
VII	28,380	16.3	7.45
VIII	68,314	17.9	7.10
IX	36,637	25.1	9.20
X	31,277	17.7	8.11
XI	6,002	29.9	11.71
XII	6,013	14.9	8.24
Metropolitan	105,114	7.0	4.28
National	443,953	13.3	9.59

[a]Calculated as the ratio of the number in the first column to the total number of households with shared or own water connections in each region estimated using the expansion factors of the CASEN survey data.

[b]Calculated using data from Mideplan (1999) deflated by the average exchange rate from the Central Bank of Chile for that year.

Source: Mideplan 1999 and CASEN 1998.

production costs are high. The national average benefit is close to US$10 per beneficiary household per month.

The aggregate budget for the subsidy reached $42.5 million in 2000. As Serra (2000) points out, the cost of the subsidy is well below the cost of the universal subsidy given to loss-making providers before the reform. In 1998 water and sewerage companies had net profits of $107 million, more than twice the cost of the subsidy scheme (not considering administrative costs).

One important point to make is that subsidies are normally awarded to a family for a three-year period (although benefits accrue to the household monthly). This means that a large fraction of the budget for a given year is already determined by the number of ongoing subsidies that were distributed in previous years.

II. The Colombian Water Subsidy Scheme

Unlike Chile, Colombia has opted for a scheme based on cross-subsidies between different clients, although in practice benefits are mostly funded from general tax revenues.[10] Colombia is not unique in the use of cross-subsidies. These subsidies are ubiquitous in water tariffs across the world. What makes the Colombian system interesting is that the cross-subsidies are explicitly laid out in the Public Residential Services Law of 1994, and a geographic targeting system is used to determine whether a client pays a surcharge or receives a subsidy from the tariff structure. The objective of the 1994 reform was to unify criteria for the application of cross-subsidies across utility industries and to guarantee consistent application of the scheme across the country. Another motivation for the reform was to reduce the magnitude of cross-subsidies, which had grown to extremely inefficient levels in some cases.

In each municipality dwellings are classified into six socioeconomic categories. Households are eligible to receive a subsidy of up to 50 percent of the average service cost if they live in dwellings classified as level 1 (low-low) and of up to 40 percent if they live in dwellings classified as 2 (low). Households living in level 3 dwellings (medium-low) may also receive a subsidy of up to 15 percent of the average service cost. The decision on whether to grant a subsidy to this middle group is up to the regulatory commission in each case.

Subsidies are funded through a variety of sources. First, a surcharge can be applied to clients in dwellings classified as 5 or 6 and to industrial and commercial customers.[11] These surcharges are capped at a maximum of 20 percent of the water and sewage bill. When these surcharges are not sufficient to fund the subsidies to the first three socioeconomic groups, the difference can be financed by transfers from the national and provincial budgets.

10. In cross-subsidy schemes, some clients pay a tariff greater than the cost they impose on the service provider to finance lower tariffs for other clients.

11. Hospitals, schools, and other nonprofit organizations are exempt from paying surcharges or receiving subsidies. These clients must pay the full tariff.

TABLE 2. Tariffs Paid as a Proportion of the Cost-Recovery Tariff for 15 Companies, by Socioeconomic Classification of Dwellings, December 1997, Colombia (percent)

	Socioeconomic classification					
Range of tariffs paid	1 (low-low)	2 (low)	3 (medium-low)	4 (medium)	5 (medium-high)	6 (high)
Minimum	13	27	41	44	60	75
Average	35	48	57	74	87	114
Maximum	92	94	91	109	117	150

Note: Data were available for companies in Bogotá, Cali, Medellin, Barranquilla, Cartagena, Bucaramanga, Pereira, Manizales, Pasto, Armenia, Valledupar, Popayan, Girardot, Magangue, and Santa Rosa.

Source: Sánchez Torres and Núñez Méndez 1999, based on data from Unidad de Política Regional y Urbana, Departamento Nacional de Planeación.

Data on the tariffs for each group as a proportion of the cost-recovery tariff for the 15 water companies for which information was available at the end of 1997 indicate that despite the residential services law there was still an element of universal subsidy in water charges (table 2). Even consumers living in dwellings classified in groups 5 and 6 received a subsidy in some regions, although on average across regions consumers in the highest socioeconomic classification paid a surcharge. It is also apparent that for most of the water companies the subsidies accruing to the first three groups are higher than the 50 percent, 40 percent, and 15 percent reduction allowed in the law. Thus actual tariffs did not conform to the surcharge and subsidy limits imposed by the residential services law. This is explained by the fact that there was an initial five-year transition period for implementing the new limits, and in 1997 this transition was still incomplete.[12]

The municipalities are responsible for classifying dwellings according to the six socioeconomic categories, following the guidelines and methodology provided by the National Planning Department. The basic classification unit is a geographic area with homogeneous characteristics according to criteria set by the National Planning Department. All dwellings within this unit are classified in one socioeconomic group, although particular dwellings that have different characteristics within the unit can be individually classified in another group. A dwelling's category is the same for all utility services. Individuals can request a revision of their dwelling's classification.

A municipality can hire public or private institutions to undertake the classification study on its behalf. It can also count on the technical assistance of national and regional entities. Municipalities within a larger metropolitan area can also undertake the classification jointly. Finally, the same classification used by

12. To date, tariffs have not been fully adjusted to meet the surcharge and subsidy limits, and the transition period for meeting these limits has been extended for another five years.

municipalities to differentiate property taxes can be used for differentiating utility charges.

III. Comparison of the Distributional Impacts of the Chilean and Colombian Water Subsidies

This section evaluates the targeting results of the two subsidy programs using the traditional tools of distributional analysis (see the appendix for some data issues and sources). However, besides the distributional impact, there are other complementary issues that should be considered when making an overall evaluation of a particular subsidy scheme. The administrative costs of a program, the relative efficiency of the funding mechanism used, and the possible distortions generated in other markets are also relevant. In this article, however, the focus is exclusively on the distributional impact of these subsidies.[13]

Estimating the Distributional Impact of the Chilean Program

There are two difficulties in estimating the distributional impacts of the Chilean subsidy. First, determining the target population is difficult because the aim of the program is to benefit poor households that without the subsidy would spend more than 5 percent of their income on water and sewage services. Therefore, the target population will be a function of water tariffs as well as household income and will be different for each region. In some regions close to 30 percent of households are eligible for the subsidy, whereas in other regions the share is much lower (see table 1). Because the total number of subsidies for each region is further differentiated among municipalities, a poor household in a municipality with low tariff rates might not receive a subsidy, and a relatively better-off household in another municipality with higher water charges does receive a subsidy. This should not constitute a targeting error in light of the declared objectives of the subsidy. However, extensive work for this study reveals that analysis at the regional or local level does not improve the targeting results compared to a national income stratification (results available on request).

The second difficulty relates to the fact that the survey data used, the 1998 Chilean National Characterization Survey (CASEN), is able to identify only 221,821 households that receive the subsidy, whereas Ministry of Social Planning records show 443,953 beneficiary households (see appendix for more information on the survey). If the raw survey data are used, there may be important biases in the estimated distributional impacts. To control for this, the missing subsidies were distributed across deciles according to the proportions recorded in the survey. This is the only reasonable option because it assumes that the underreporting of water subsidies was random. Conversations

13. A wider assessment of the two subsidy programs discussed here is provided in Gómez-Lobo and Contreras (2000).

TABLE 3. Share of Households Receiving the Water Subsidy by Deciles of per Capita Household Income, Chile (percent)

Decile	Raw data	Adjusted data[a]
1	14	27
2	12	23
3	10	21
4	9	19
5	7	14
6	7	14
7	4	7
8	3	5
9	1	2
10	0	0

[a]The difference in subsidies between CASEN and Ministry of Social Planning records was distributed across deciles in the same proportions as recorded in CASEN; see text and appendix for details. Includes households with shared or own water connections only.

Source: CASEN 1998.

with professionals linked to the design and application of CASEN uncovered no reason to believe that underreporting was biased among income groups.

According to the raw survey data, only 14 percent of households in the lowest decile of the income distribution receive the water subsidy (table 3).[14] In the second lowest income decile only 12 percent receive the subsidy. That the majority of households in the first quintile do not receive a subsidy implies a very high error of exclusion—households that in principle should be eligible for the subsidy do not receive it. The results improve once the data are corrected for underreporting (all further results are based on the corrected data). About a quarter of households in the first quintile receive the benefit. Many nondeserving households also receive the subsidy, although most subsidies are concentrated in the first two income quintiles.

Estimating the Distributional Impact of the Colombian Program

In Colombia the vast majority of households that have water connections are in socioeconomic groups 2 and 3. Nearly 83 percent of households are classified in the first three groups eligible to receive a subsidy. This means that the tax base (households that are liable to pay a surcharge) on which to finance the subsidy is quite small, suggesting that there are not enough richer households to fund transfers to poorer ones.[15] In fact, according to the 1997 data only 1 percent of

14. Deciles were based on per capita household income. Using an equivalent income measure that takes into account the demographic composition of households does not change the results.

15. This problem is compounded by the fact that the cross-subsidies are internal to service suppliers. Thus, in a poor region, where more funds are required given the number of dwellings in the lowest socioeconomic categories, there are fewer rich dwellings to tax.

TABLE 4. Distribution of Socioeconomic Classification by Deciles of Household per Capita Income, Colombia (percent)

	Socioeconomic classification						
Decile	1 (low-low)	2 (low)	3 (medium-low)	4 (medium)	5 (medium-high)	6 (high)	Total
1	18.8	47.3	31.4	1.6	0.7	0.3	100.0
2	15.3	45.6	33.8	2.5	2.0	0.8	100.0
3	16.2	38.7	39.2	2.8	2.0	1.1	100.0
4	14.5	40.4	38.4	4.5	1.7	0.5	100.0
5	8.8	43.4	40.4	5.3	1.7	0.4	100.0
6	7.9	32.5	49.6	6.5	2.7	0.8	100.0
7	5.7	30.9	44.5	13.1	4.7	1.1	100.0
8	1.4	27.0	44.4	20.1	6.5	0.6	100.0
9	0.5	20.5	46.5	16.4	11.8	4.4	100.0
10	1.2	8.6	31.7	26.6	23.1	8.7	100.0

Source: ENCV 1997. Observations with zero income were dropped.

distributed benefits were funded through negative transfers from people living in dwellings classified in groups 5 and 6.

The classification system shows some correlation with poverty (table 4), but the correlation is quite imperfect, a result consistent with earlier work on the distributional effects of residential utility service subsidies in Colombia (see Vélez Echavarría 1993; Sánchez Torres and Núñez Méndez 1999). Most households in the poorest five income deciles live in dwellings classified in socioeconomic groups 1 to 3.[16] However, errors of inclusion are large. More than 75 percent of households in income deciles 6 to 9 fall into one of the three lowest socioeconomic groups. Only for the richest decile are most households classified in the highest three socioeconomic groups, but even in this case less than 32 percent of households are classified in the top two categories, making them liable to pay surcharges.

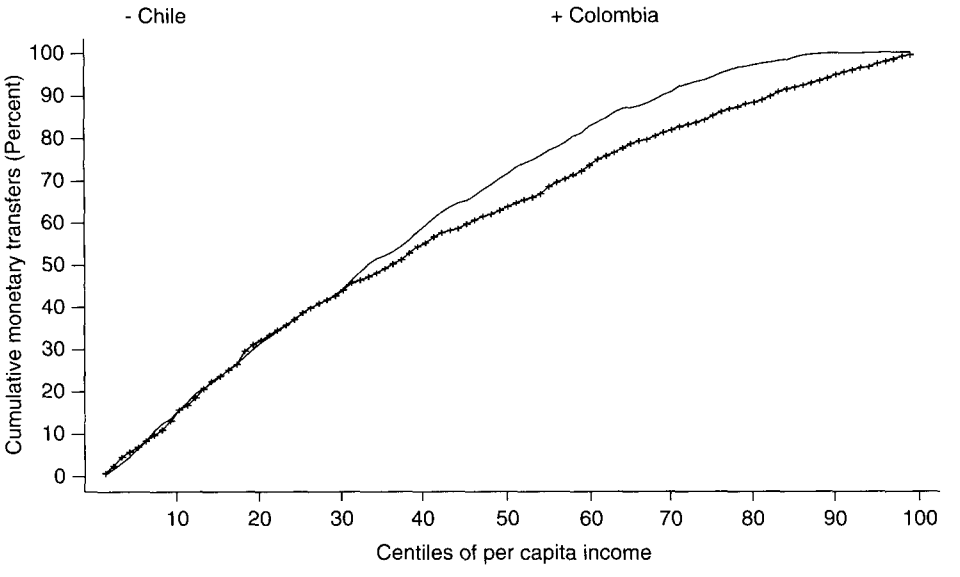
Relative Concentration Curves

Examination of the relative concentration curves for monetary transfers and total number of beneficiaries permits a more detailed comparison of the distributional impacts of the Chilean and Colombian water subsidy schemes.

The relative monetary transfer distribution curves for each country show the percentage of total monetary transfers that accrue to households that are at or below a certain range of the income distribution (figure 1). The horizontal axis measures centiles of income distribution, from poorest to richest, and the

16. Income was defined as primary income plus rental values of any dwelling owned by the household. Secondary income, such as lottery money, loans, property sales, and other infrequent revenues, were not considered.

FIGURE 1. Cumulative Monetary Transfer Curve per Centile of per Capita Income



Source: Authors' computations based on CASEN 1998 and ENCV 1997.

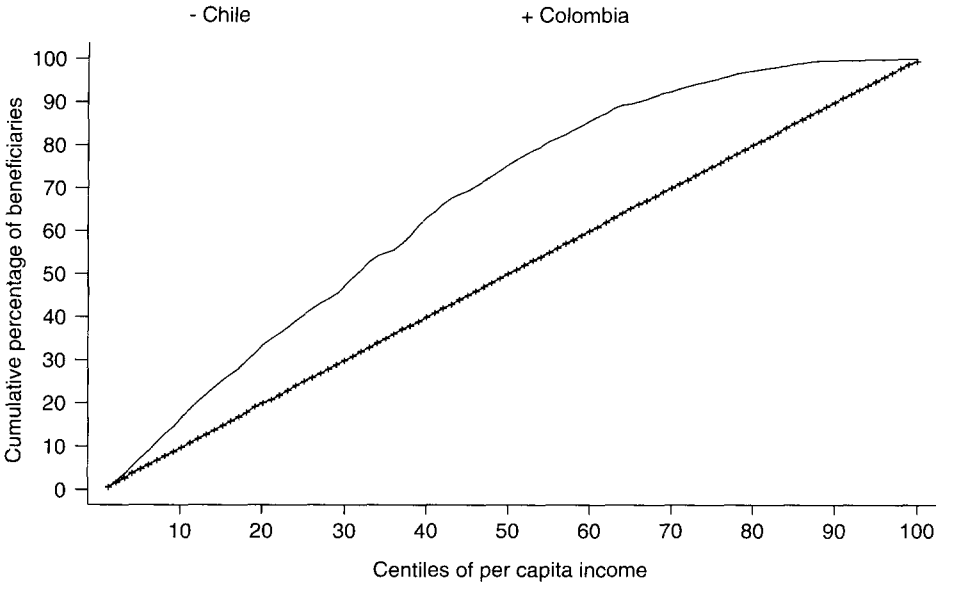
vertical axis measures the accumulated percentage of total transfers.[17] The higher and more concave the curve, the better the targeting property of the subsidy. Up to the 30th income centile (the third income decile) the effects of monetary transfers are almost identical in both countries. Roughly the same percentage of transfers, close to 45 percent, accrue to the poorest 30 percent of households. Above the third income decile, however, the Chilean scheme performs better, transferring more income to middle-income groups rather than to higher income groups.

The relative concentration curves with respect to the number of beneficiaries is another source of information on the targeting performance of each subsidy scheme (figure 2). Again, the horizontal axis measures the centiles of the income distribution and the vertical axis is the accumulated percentage of the total number of beneficiaries. For a given point in the income distribution the graph shows the percentage of beneficiaries that have incomes at or below that point. Once again, the higher and more concave the curve, the better the targeting property of the subsidy. Although in Chile close to 45 percent of beneficiaries have incomes in the first three deciles of the distribution, in Colombia only 30 percent do.

The beneficiary concentration curve suggests that the Chilean means-tested subsidy scheme is more efficient in identifying poor households. Indeed, the

17. The surcharges (negative transfers) for higher income categories are omitted as they are considered part of the funding mechanism of the subsidy. This issue is discussed later.

FIGURE 2. Cumulative Number of Beneficiaries per Centile of per Capita Income



Source: Authors' computations based on CASEN 1998 and ENCV 1997.

Chilean concentration curve lies above the Colombian curve at each centile of per capita income. This result suggests that the means-tested subsidy performs better than the geographic system in identifying needy households, confirming expectations about the relative efficiency of means-tested benefits. However, because the Colombian scheme is more progressive, providing larger transfers to the poorest households, the overall poverty impact of both programs is very similar, at least for the poorest 30 percent of households. Thus in the overall poverty impact of the scheme, the differentiation of benefits among different socioeconomic categories in Colombia compensates somewhat for the failure of the geographical targeting system in identifying the poor.

Thus a program that combines the targeting properties of a means-tested subsidy, as in Chile, with the progressiveness of monetary transfers per household, as in the Colombian scheme, would have a greater impact on poverty than either program as they are currently designed. In Chile this change could easily be accomplished by calibrating the amount of the water subsidy to each household's CAS score.

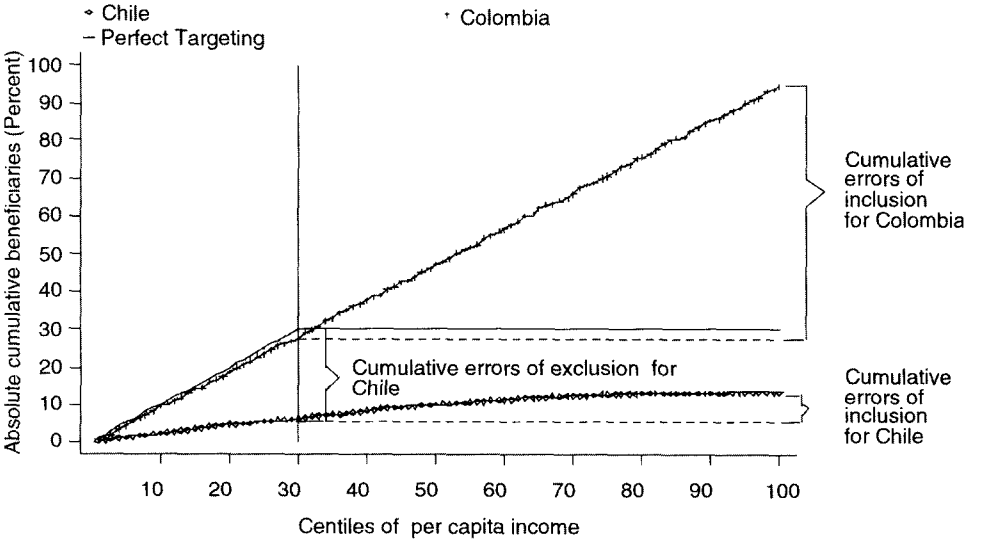
The Colombian scheme also has the potential to be made more progressive. Tightening the classification of dwellings to make the subsidy more targeted (by transferring more dwellings to the two highest classifications) would increase the tiny fraction of benefits (1 percent) funded by households in these two groups, improving not only the distributional effects of the scheme by focusing resources on more needy households but also the progressivity of the scheme by taxing higher income households.

The extent of the errors of inclusion that occur in both countries is striking when the target population is assumed to be the first few deciles of the income distribution (see figure 2). Although the Colombian scheme has no explicitly declared targeting objective, the Chilean subsidy law defines beneficiary households as households in the first two income deciles. Yet close to 70 percent of beneficiaries are not in this group.

What these relative concentration figures do not capture is the extent of errors of exclusion (already discussed). The relative concentration curves (figures 1 and 2) show the relative targeting properties of both schemes, whereas errors of exclusion depend on the scale of each program. In other words, given the total number of subsidies offered in Chile and Colombia, the Chilean system does a better job of giving the subsidies to poorer households (see figure 2). However, the Colombian scheme covers a much larger share of the population, covering close to 95 percent of households compared with Chile's 15 percent (see table 1). Thus in Colombia most poor households will receive some benefit, whereas in Chile the errors of exclusion are much higher.

This difference can be seen by analyzing the absolute concentration curve of beneficiaries—the analogue of Shorrocks (1983) absolute Lorenz curves—which is constructed by multiplying the relative concentration curves by the percentage of the population that receives a subsidy (figure 3). The solid line in figure 3 represents perfect targeting under the assumption that the target population is the first three income deciles. It is a 45-degree line up to the 30th percentile of the income distribution and then flat afterward. It reflects the fact that perfect

FIGURE 3. *Absolute Beneficiary Concentration Curve*



Source: Authors' computations based on CASEN 1998 and ENCV 1997.

targeting in this case implies that only 30 percent of the population are beneficiaries (as reflected in the vertical axis) and that all 30 percent are concentrated in households in the first three deciles. Households above this limit are not intended beneficiaries and so the curve is flat afterward.

The empirical concentration curve will reflect targeting errors. The distance between the perfect targeting curve and the empirical curve at the kink point reflects the errors of exclusion. For example, in Chile the cumulative absolute beneficiary curve, which should include 30 percent of households at the perfect targeting level, includes only 5 percent, implying that about 25 percent of the population that should receive the subsidy under the targeting goal does not. In Colombia the errors of exclusion are close to zero, because all of the poorer households receive a transfer—because of the almost universal subsidy characteristic of the Colombian program rather than because of the efficiency of the targeting mechanism.

The distance between the empirical concentration curve at the 100th centile and the point where this curve crosses the 30th centile measures the errors of inclusion of each program—the percentage of the population that receives the subsidy but is not in the target population. In Colombia, nearly universal coverage reflects large errors of inclusion. That makes the program extremely expensive from a fiscal perspective. A better targeting mechanism would lower these costs without necessarily endangering the protection of lower-income households.

Thus although errors of inclusion are much higher in Colombia, errors of exclusion are much higher in Chile, even if the target population is limited to the first decile. These errors of exclusion may be worrisome from a policy perspective. If the objective of the water subsidy is to allow the whole target population to consume the service at lower rates, then clearly the chosen policy instrument is failing. Although the targeting mechanism is better in relative terms than the one used in Colombia, it still leaves many poor households without benefits. Unless the targeting mechanism is improved, the only way to benefit more of the target population would be to increase the number of subsidies.[18]

IV. Conclusion

This study analyzed and compared two water subsidy schemes. Chile's individual means-tested subsidy was expected to be more efficient in identifying needy beneficiaries, although it can be expensive to administer and requires sophisticated institutional capacity at the local government level. Colombia's scheme, based on a geographical classification of dwellings and cross-subsidies, was

18. In the summer of 2002 (January–February) the government decided to redistribute a substantial number of water subsidies based on an evaluation showing that the subsidy was not reaching the target population precisely enough. With data from the next CASEN (2003) it will be possible to judge whether this change improved the targeting results of the subsidy program.

expected to be less efficient in identifying needy beneficiaries but also less expensive to administer than a means-tested scheme.

This study bore out these expectations about the targeting ability of each scheme. Chile's means-tested system was better able to identify poor households than Colombia's geographically targeted scheme case, but because the Colombian scheme differentiates the amount of benefits per household according to the socioeconomic classification of its dwelling, the overall impacts of the two schemes are quite similar, at least for the first three income deciles.

However, both subsidy programs could be improved. In Chile the scheme could have a greater impact on poverty if the monetary transfers were differentiated by socioeconomic levels, as in Colombia. This could be done by conditioning the amount of transfers to eligible household on their relative CAS scores. The poverty impact could also be increased if targeting errors were reduced. More than 60 percent of subsidies go to households that are above the third decile of the income distribution. Because the number of subsidies available each year is determined without taking imperfections in the targeting mechanism into account, many poor deserving households are not receiving any benefits. If the policy objective in Chile is to reach a significant proportion of lower-income households, then either the targeting mechanism must be further improved or the number of subsidies given in a year has to increase to compensate for targeting imperfections.

In Colombia almost all poor households receive some benefit. However, this is achieved through an overly generous program that gives some benefits to almost all households. The high fiscal costs of this system could be reduced without jeopardizing the benefits given to poorer households by improving the targeting scheme by reclassifying more dwellings of higher income households into the top two socioeconomic categories. This would improve not only the distributional effects of the scheme, by focusing resources on more needy households, but also the progressivity of the scheme on the financing side, by taxing more higher income households.

Finally, the empirically imperfect targeting results in both countries found by this study should not be interpreted as a case against targeted welfare programs. The results are very likely better than they would have been had no special effort been made to reach vulnerable households. In addition, both governments have introduced changes to improve targeting performance. In 2002 the Chilean government decided to redistribute a substantial number of water subsidies based on an evaluation showing that the subsidy was not reaching the target population precisely enough. The changes, which generated some political opposition, should improve the targeting performance of the scheme. In Colombia the government is planning a new household classification effort that should also improve the targeting properties of the subsidies.

Further research could help improve the schemes in both countries. In Chile it would be important to investigate why there are still large errors of inclusion and exclusion despite the relatively sophisticated targeting mechanism. Perhaps

targeting such a narrow segment of the population is difficult. Given the skewed nature of the income distribution, the average household in decile three is probably very similar in wealth, durable goods ownership, and family income compared with the average household in decile four. Any targeting mechanism will have difficulties distinguishing between them. Alternatively, it may be that the targeting mechanism works well at the time a subsidy is originally awarded. However, because subsidies are generally given for a three-year period, the socioeconomic conditions of households can subsequently improve, thus lowering ex post targeting effectiveness. These and other hypotheses should be explored in future research.

In Colombia there is an important aspect of the welfare system that merits further analysis. Water, gas, and electricity subsidies and property taxes are all linked to the socioeconomic classification of a dwelling. Because that classification is easily verifiable, it is reasonable to conjecture that some of the benefits of the scheme are capitalized into housing and rental prices. If this is the case, then some part of the benefit of the subsidies would go to the owner of a dwelling at the time the scheme was introduced rather than to the current occupier. Thus the distributive impact of the Colombian scheme measured in this article could well be biased.

Appendix. Data Issues

The data for Chile came from the 1998 CASEN, a cross-sectional survey designed to collect information on the socioeconomic situation of Chilean families, including housing, education, subsidies, and labor characteristics. CASEN began to incorporate information on water subsidy benefits in 1996. The 1998 survey includes information on 48,107 households as well as on the individuals within each household.

CASEN accounts for only about half of the water subsidies distributed in the country. The Ministry of Social Planning is aware of this problem but assumes that the underreporting is random, so that the distribution of benefits from the sample can be extrapolated to the whole water subsidy program. Some evidence that this may be correct is provided by a comparison of the average subsidy per beneficiary, Ch$3,280 (US$7.1) per month according to program records, with the average subsidy per beneficiary from the 1998 survey, Ch$3,137 (US$6.8) per month.

The data used for Colombia came from the 1997 National Quality of Life Survey (ENCV), a household survey undertaken by the National Planning Department. The survey includes a wide range of socioeconomic variables, including the amount paid by a household for water and sewage services. It does not, however, register the amount of surcharge or subsidy for these services, only the actual amount spent. Therefore, subsidies must be estimated indirectly. Following Sánchez Torres and Núñez Méndez (1999), indirect information on the portion of costs covered by each tariff for different companies was used to infer the amount of the subsidy (details can be found in Gómez-Lobo and Contreras 2000).

This approach implies that only observations from the departments where information was available could be used. These included the Departments of Cauca, Nariño, Valle, Caldas, Risaralda, Quindío, Bogotá, Cundinamarca, Santander, Antioquía, Cesar, Atlantico, and Bolivar. In addition, all households without connections to piped water were dropped, because they would not receive services from the water companies. Finally, households without a socioeconomic classification were dropped.[19] Of the original 9,121 households in the sample, 4,094 remained in the analysis after these criteria were applied. Some observations do not register the water bill. Therefore, concentration curves were constructed using only observations with positive water charges.

References

Brook, P. J., and S. M. Smith. 2001. *Contracting for Public Services: Output-Based Aid and Its Applications*. Washington, D.C.: World Bank.

Estache, A., V. Foster, and Q. Wodon. 2002. *Accounting for Poverty in Infrastructure Reform: Learning from Latin America's Experience*. WBI Development Studies. Washington D.C.: World Bank Institute.

Estache, A., D. Leipziger, and A. Gómez-Lobo. 2001. "Utilities 'Privatization' and the Poor's Needs in Latin America: Have We Learned Enough to Get It Right?" *World Development* 29(7):1179–98.

Gómez-Lobo, A., and D. Contreras. 2000. "Subsidy Policies for the Utility Industries: A Comparison of the Chilean and Colombian Water Subsidy Schemes." World Bank Institute, Washington, D.C.

Mideplan. 1999. "Antecedentes Administrativos, Operativos y Legales del Subsidio al Pago del Consumo de Agua Potable y Servicios de Alcantarillado de Aguas Servidas." Documentos Mideplan, Ministerio de Planificación y Cooperación, División Social, Departamento de Información Social, Santiago.

Sánchez Torres, F., and J. Núñez Méndez. 1999. "Decentralización, Pobreza y Acceso a los Servicios Sociales: ¿Quién se Benefició del Gasto Público Social en los Noventa?" Universidad de Los Andes, Centro de Estudios de Desarrollo Económico, Facultad de Economía, Santa Fe de Bogotá.

Serra, P. 2000. "Subsidies in Chilean Public Utilities." Documento de Trabajo Serie Economía 70. Universidad de Chile, Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Santiago.

Shorrocks, A. F. 1983. "Ranking Income Distributions." *Economica* 50(197):3–17.

Vélez Echavarría, C. E. 1993. Gasto Social y Desigualdad, Logros y Extravíos: Estudio de la Incidencia del Gasto Público Social en Colombia. Santa Fe de Bogotá: Departamento Nacional de Planeación.

World Bank. 1994. World Development Report 1994: Infrastructure for Development. New York: Oxford University Press.

19. There were many observations with a classification of 0 (possibly an indicator of missing information) or 9 (possibly industrial or commercial premises). One household with no adults was also dropped from the sample.

Policy Selectivity Forgone: Debt and Donor Behavior in Africa

Nancy Birdsall, Stijn Claessens, and Ishac Diwan

We assess the dynamics behind the high net resource transfers by donors and creditors to Sub-Saharan African countries. Analyzing the determinants of overall net transfers for a panel of 37 recipient countries in 1978–98, we find that country policies mattered little. Donors—especially bilateral donors—actually made greater transfers to countries with high debt, largely owed to multilateral creditors, when policies were "bad." We conclude that comprehensive debt relief has the potential, though not the certainty, to restore selectivity in support of good policies. That would make development assistance more effective going forward—and increase public support in donor countries.

A large body of literature has developed in recent years on country factors and other factors that influence the effectiveness of aid and of the development aid business more generally. Two major findings have emerged. First, aid is more effective when the recipient country's policy and institutional environment satisfies some minimal criteria. Second, aid and debt relief have not been targeted particularly toward countries with adequate policies and institutions (Burnside and Dollar 2000).[1] In this article, we concentrate on understanding the dynamics behind the second finding. To do so, we analyze the donor and official creditor side of the aid process. Our specific hypothesis is that the growing debt of poor countries over the past two decades and the composition of that debt have affected the granting of new loans and grants by donors and official creditors.

Nancy Birdsall is President of the Center for Global Development; her e-mail address is nbirdsall@cgdev.org. Stijn Claessens is Professor of International Finance Policy at the University of Amsterdam; his e-mail address is stijn.claessens@uva.nl. Ishac Diwan is World Bank Country Director for Ethiopia and Sudan; his e-mail address is idiwan@worldbank.org. The authors are grateful to Jean-Claude Berthelemy, Alan Gelb, Bernhard Gunter, Ratna Sahay, and Axel van Trotsenburg and to various participants in seminars at the World Bank, the Carnegie Endowment for International Peace, the 2001 World Bank Economists' Forum, the 2001 WIDER Debt Relief Conference, and the 2003 American Economic Association meeting for their comments. They thank François Bourguignon and three referees for their extensive comments, Ying Lin (World Bank) for his excellent work with the data, and Brian Deese for other research help. The article was previously titled "Will HIPC matter? The Debt Game and Donor Behavior in Africa." Address questions and comments to the corresponding author, Stijn Claessens.

1. The first finding is not without controversy, however, particularly regarding the robustness of the connection between "good policy" and aid effectiveness (Hansen and Tarp 2001). See Hansen and Tarp (2000) for a review of the literature and World Bank (1998) for further information on aid effectiveness.

THE WORLD BANK ECONOMIC REVIEW, VOL. 17, NO 3 409–435

DOI: 10.1093/wber/lhg029

Analyzing the behavior of donors and creditors as it relates to the accumulated debt burden is important because it can shed light on a critical policy question. Will the official program of debt reduction for heavily indebted poor countries (HIPCs) not only reduce debts but also affect future donor behavior, particularly the ability and willingness of donors to direct aid to its best uses? Or will more debt reduction simply invite another round of business as usual (new loans and new debt accumulation) implicated in the first debt buildup?

To address these questions, we investigate donor and creditor behavior using data on net transfers for 37 Sub-Saharan countries over the 1980s and 1990s. Just as others have documented, our analysis confirms that the quality of a country's policy framework has mattered little in determining overall net transfers. Importantly, we find that the buildup of debt stocks owed to the multilateral creditors hindered the targeting of resource flows to countries that had better policies and improved institutional environments. We find in particular that more indebted countries received more net transfers—and that among countries highly indebted to multilaterals, those with policies below the median in quality received on average some 2 percentage points of gross domestic product (GDP) *more* in net transfers over the period. We also find that donors are selective for country policies in low debt countries but not so in high multilateral debt countries. These findings are robust to the use of different estimation techniques and alternative measures for the quality of policy.

The findings suggest that the buildup of debt, especially to multilaterals, has undermined the ability of the donors to be selective for country policies—that is, to transfer less where the policy setting is poor. They imply that debt reduction for high multilateral debt countries can allow the behavior of the donor community to shift to a low debt regime mode, a regime that in the past has allowed selectivity. Debt reduction can be interpreted in short as a way out for a donor community otherwise locked into a pattern of nonselectivity in the high multilateral debt countries.

Section I describes the overall setting for development assistance and documents the accumulation of debt by Sub-Saharan countries over the 1980s and 1990s to different classes of creditors. Section II describes the data used and provides the major trends and raw statistics. Section III specifies the hypothesis, describes the empirical analysis, and discusses the major findings. Section IV concludes.

I. Development Assistance and Debt Accumulation in Africa, 1980s and 1990s

Over the past 25 years Sub-Saharan countries have been major recipients of overseas development assistance. Aid in the form of grants and loans from bilateral and multilateral donors has amounted to about $350 billion dollars (in nominal terms). In some countries gross aid flows were 60 percent or more of GDP in some years; in many countries flows often exceeded the government's tax revenues. With a few exceptions, these countries have had low rates of per

capita GDP growth. Despite high levels of lending and grant programs, the growth rate of per capita GDP for the region as a whole was negative over the last two decades (about −2 percent a year in the 1980s and −1 percent in the 1990s). Average GDP per capita in constant prices was lower in 2000 than in 1960. And the number and proportion of poor people actually increased: 40 percent of the 600 million people in Sub-Saharan Africa lived on less than $1 a day in 2000 (World Bank 2000).

Meanwhile, much of the high level of development assistance took the form of loans, producing a growing stock of debt—from about $60 billion in 1980 to $230 billion in 2000. Annual debt service paid also increased, from an average of $6 billion a year in the early 1980s to about $11 billion in the late 1990s. Growth in debt service was much less than growth in debt, however, due to debt restructuring and increases in the concessionality of resources provided. In the 1990s especially, repeated rounds of debt rescheduling and debt service relief by the official donor community—and an increase in the proportion of donor transfers as grants—kept debt service from rising more.

While many other countries have had external debt problems, two features of the debt problem in Africa are notable. First, because of the preponderant role of official creditors and donors (as opposed to commercial creditors), net transfers have been always positive and large. Total disbursements in the form of new loans and grants have always exceeded countries' actual debt service. Indeed net transfers—the difference between new disbursements and debt service paid—have been 10 percent or more of GDP for most countries for the two decades. Second, the proportion of total debt owed to the International Monetary Fund (IMF), World Bank, and other multilaterals (African Development Bank, European Investment Bank) has been constantly growing as bilateral donors switched from loans to grants and increasingly forgave outright portions of debt owed them. Between 1980 and 1998 the share of multilateral debt in total debt increased from about one-seventh of total debt to almost one-third, and the share of the multilaterals in total debt service increased from about one-tenth to one-third.[2]

These features highlight the important differences between the debt problems of the African countries today and those of Latin American and other middle-income countries in the 1980s, extensively analyzed (see Eaton and Fernandez 1995 for a review of this literature). Most debts of the Latin American countries were then due to commercial rather than official creditors. Each commercial creditor was individually interested in maximizing the value of its claim on the country. This desire to extract payments led to a "tax" on the country, a debt overhang. High levels of debt were leading to disincentives to adjust economic policies, and new investors were deterred from committing resources to the country (Diwan and Rodrik 1992). The literature then stressed the potential beneficial effects of a reduction in the face value of debt for creditors because it

2. Total debt service paid also increased, from about 7 to 15 percent of the value of exports.

could increase the incentives for a debtor to adjust and enhance a country's ability to attract new (type of) investors and fresh funds. Ex post, the debt overhang was resolved through debt reduction (the Brady plan) that seems to have had some of these beneficial effects (Claessens and Diwan 1994).

The situation of the African countries is quite different (Claessens and others 1997). Although the debt stocks were rising, they did not impose any actual debt servicing burden because countries received large positive net transfers and did not need to repay their creditors.[3] But with debt service payments rising, especially to multilaterals, higher disbursements by donors (the multilaterals themselves or the bilaterals) were needed to maintain net transfers. The rising debt levels and the increase in the share of the multilaterals meant that by the mid-1990s the donors and creditors—not the indebted countries—were caught in a debt trap. The donors wanted to avoid having the indebted countries, among the poorest in the world, fall behind in debt service to the multilaterals. Arrears to the multilaterals would have meant the curtailment of future lending by not only the multilaterals but all other donors because continuing aid flows required an active multilateral lending program. From the donors' point of view, arrears would make visible the failure of the past aid transfers.

As debt stock considerations started to drive new disbursements, the quality of policy and degree of poverty in countries became less relevant. The need to maintain high new disbursements in highly indebted countries may have meant that donors no longer had enough freedom to differentiate new disbursements by the quality of policy and degree of poverty. In other words, indebtedness took some of the donor community's ability to be selective for policy and poverty, and aid flows started to respond more to debt stocks, less to policy and poverty.[4]

II. Data and General Analysis of Donor Behavior

To assess creditor and donor behavior, we analyze debt indicators and net transfers for a sample of countries in Sub-Saharan Africa over 1977–98. We want to assess donor behavior toward countries in the region independent of whether countries eventually became HIPC-eligible or not, a classification that occurred around 1998, and without the classification affecting donor behavior. We therefore use a sample that includes both HIPC and non-HIPC countries and stop our analysis at 1998, avoiding any sample selection problem. We include in

3. The fact that net transfers have generally been positive has not been sufficiently taken into account in the often heard arguments that countries are spending more on debt service than on social programs. Birdsall and Williamson (2002) point out that it may, however, still be true that the local taxes needed to pay debt service are not really offset by the often tied aid transfers for specific projects.

4. Easterly (1999) develops a model to explain why countries with certain characteristics end up with high debt. His model has the strong implication that countries pursue bad policies to receive future debt reduction. The model does not examine the behavior of the creditors to these countries, however, in relation to the debt composition.

TABLE 1. Sample Countries

HIPC debt initiative	Non-HIPC
Benin	Botswana
Burkina Faso	Gabon
Burundi	Lesotho
Cameroon	Mauritius
Central African Republic	Nigeria
Chad	Seychelles
Comoros	Swaziland
Congo	Zimbabwe
Congo, Dem. Rep	
Côte d'Ivoire	
Ethiopia	
Gambia, The	
Ghana	
Guinea	
Guinea-Bissau	
Kenya	
Liberia	
Madagascar	
Malawi	
Mali	
Mauritania	
Niger	
Rwanda	
Senegal	
Sierra Leone	
Sudan	
Togo	
Uganda	
Zambia	
29	8

Note: HIPC classification as of fall 2002.
Source: World Bank (various years).

our analysis the 37 Sub-Saharan countries with all necessary data. Some countries (including Eritrea, Angola, Somalia, and Tanzania) are excluded because of lack of data in many individual years. Of the 34 African countries eligible for HIPC treatment (in 2002), 29 are included in our sample of 37 (table 1).[5]

Data

The variable of interest for this analysis is the amount of net transfers a country receives from abroad related to debt or grants—that is, the amount of net movement of real resources to the country from official sources on account of debt or grants. We exclude from our analysis flows related to foreign direct,

5. The HIPC sample includes mainly Sub-Saharan countries. Because the bulk of aid flows has also gone to Sub-Saharan countries, we focus on the behavior of donors in these countries.

portfolio, and other nondebt investments. Net transfers are defined, in accord with the World Bank's *Global Development Finance* (GDF), as the amount of resources the countries receive in the form of grants and new debt disbursements net of repayments on old debt. In other words,

$$NT = G + NB - (P + R) = G + NB - TDS \tag{1}$$

where NT = net transfers, G = the amount of grants (free gifts) the country receives, NB = new debt disbursements, P = principal repayment on existing debt, R = interest payment on existing debt, and TDS = total debt service paid, the sum of principal repayment P and interest payments R. We restrict ourselves to resources directed to the government from mainly official (nonprivate) sources. These sources include the World Bank, the IMF, and other multilaterals as well as bilateral donors and donor agencies.

All data on debt and net transfers are from the World Bank's GDF statistics (various years). This data set, published annually, relies on debtors' reports, cross-verified with creditor sources. It provides the statistics on debt, disbursements, and repayments. In accord with these data, all amounts related to debt are on a cash basis—that is, they represent actual payments, not scheduled amounts, so that arrears or debt and debt service reduction do not confound the data. GDF does not collect data on grants itself but relies on donors' official development assistance and Organisation for Economic Co-operation and Development (OECD)–Development Assistance Committee reports for the grant information.

Grants do not give rise to repayment obligations and are thus not affected by the difference between obligations due and actual payments, the arrears. However, the grants data do include some elements of official debt and debt service reduction because donors have included debt forgiveness in their reported figures. In particular, the amount of grants reported by donors includes some debt forgiveness that may not imply any actual net transfers to the debtor.[6] The quality of the data on official debt reduction is known to be very poor, however, and we cannot correct the grants figure to derive a more accurate actual net resource figure (see Renard and Cassimon 2001).

We limit the impact of outliers by dropping observations for years when net transfers to a country were more than 60 percent of gross national product (we also used a lower threshold of 30 percent, but results were not qualitatively different). Nor do we always have all the independent variables that we later need in our regressions. This means that some country-year observations drop out as well. For our 37 countries, we end up with 848 country-year observations.

6. This could lead to a negative relationship between net transfers and the stock of debt, because debt stocks are reduced and although grants are higher, which would downward bias the coefficients on debt stock in our regression analysis of net transfers. Because we find positive coefficients, any misreporting only strengthens our results.

Indebtedness Categories

We turn now to a systematic analysis of donor and creditor behavior, addressing some specific questions. Were donors' net transfers to countries related to recipient countries' debt stock? Were donors and creditors providing higher net transfers to countries with better policies? For a given policy framework, were donors transferring more to countries with higher levels of poverty? In short, was there selectivity by donors and creditors as a function of countries' (changing) policies and degree of poverty? Or did the mounting debt stock and the resulting debt "crisis" lock donors into defensive lending to high debt countries, depriving them of selectivity and leverage for recipient country policies?

We start by dividing our sample of country-year observations into three subgroups of indebtedness, first distinguishing low and high debt countries, and then creating within the high debt group a further subdivision into low and high multilateral debt group (table 2). Specifically, each country-year observation is considered a low (or high) debt regime if the country's debt to GDP ratio in that year is below (or above) 62.8 percent (the median debt to GDP ratio of the whole sample). The high debt group is further subdivided into low and high multilateral debt groups, according to whether the share of multilateral debt in total debt for that country and year is below or above 41.2 percent (the median share of multilateral debt in total debt for the whole sample). In other words, the high debt, low multilateral category comprises country-year observations for which the total debt to GDP ratio is greater than 62.8 percent but the share of multilateral debt in total debt is less than 41.2 percent, whereas the high debt, high multilateral category constitutes those country-year observations in which the total debt to GDP ratio is also greater than 62.8 percent but the share of multilateral debt in total debt is greater than 41.2 percent. The high debt, high multilateral category comprises country-year observations in which the total debt to GDP ratio is also greater than 63.8 percent but the share of multilateral debt in total debt is greater than 41.2 percent.

The total sample in this classification varies slightly, between 34 and 37, because debt data are not available for each year for each country. The size of the subsamples also varies over time. The number of low debt countries was 32 in 1977, but dropped to only 9 in 1998. For the whole group the number of high debt countries grew sharply, from 2 in 1977 to 27 in 1998. The number of high multilateral debt countries in the high debt group increased sharply, from 0 in 1977 to 20 in 1998. The number of low multilateral debt countries in the high debt group grew from 2 in 1977 to 14 in 1987 and then dropped to 7 in 1998.

We next use this classification to compare the behavior of net transfers across the three debt regimes. We start by plotting the average ratio of net transfers to GDP for the three debtor classes (figure 1). Net transfers as a share of GDP have been greater in almost all years for the countries in the two high debt categories, especially for those in the high multilateral debt category. Net transfers have declined over time for the countries in all regimes, somewhat more in the high multilateral debt regime and especially so in the low multilateral debt regime.

TABLE 2. Debt Classification of Countries, 1977–98

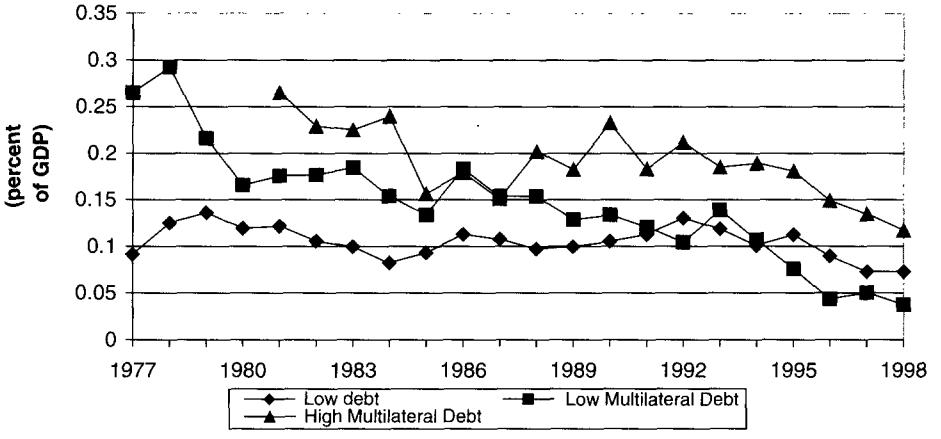
Country	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Benin	L	L	L	L	L	L	L	L	ML	L	ML	L	ML	MH	MH	MH	MH	MH	MH	MH	MH	MH
Botswana	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Burkina Faso	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Burundi	L	L	L	L	L	L	L	L	L	L	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Cameroon	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	ML	ML	ML	ML	ML
Central African Republic	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	MH	MH	MH	MH	MH
Chad	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	MH	L	L	MH	MH
Comoros	L	L	L	L	L	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Congo, Dem. Rep.	L	L	L	L	L	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML
Congo, Rep.	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML
Côte d'Ivoire	L	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML
Ethiopia	.	.	.	.	L	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML
Gabon	L	L	L	L	L	L	L	L	L	L	ML	ML	ML	L	ML	ML	ML	ML	ML	ML	ML	ML
Gambia, The	L	L	L	L	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Ghana	L	L	L	L	L	L	L	L	L	L	MH	L	L	L	L	MH	MH	MH	MH	MH	MH	MH
Guinea	.	.	.	.	.	.	.	.	.	ML	ML	ML	ML	ML	ML	ML	MH	MH	MH	MH	MH	MH
Guinea-Bissau	L	L	L	.	ML	ML	MH	ML	.	ML	.	.	.	MH	MH	MH	MH	MH	MH	MH	MH	MH
Kenya	L	L	L	L	L	L	L	L	MH	L	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	L	L
Lesotho	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	MH
Liberia	L	L	L	L	ML	MH	MH	MH	MH	MH	MH	.	.	.	.	.	.	.	.	.	.	.
Madagascar	L	L	L	L	L	L	L	ML	ML	ML	ML	ML	ML	ML	MH	MH	MH	MH	MH	MH	MH	MH
Malawi	L	L	L	L	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Mali	L	L	L	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	MH	ML	ML	MH	MH	MH	MH	MH

Mauritania	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	MH	MH	MH	MH	MH
Mauritius	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Niger	L	L	L	L	L	L	L	ML	ML	ML	ML	MH	MH	MH	L	L	MH	MH	MH	MH	MH	MH
Nigeria	L	L	L	L	L	L	L	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	L	L
Rwanda	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	.	MH	MH	L	L
Senegal	L	L	L	L	L	ML	ML	ML	ML	ML	ML	MH	MH	L	L	L	MH	MH	MH	MH	MH	MH
Seychelles	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Sierra Leone	L	L	L	L	L	L	L	L	L	MH	ML	L	L	ML	ML	ML	MH	MH	MH	MH	MH	MH
Sudan	L	L	L	L	L	ML	ML	ML	ML	L	L	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML
Swaziland	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
Togo	L	ML	ML	ML	ML	ML	ML	ML	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Uganda	.	.	.	L	MH	L	L	L	L	L	L	L	L	L	MH	MH	MH	MH	L	L	L	L
Zambia	ML	ML	ML	ML	ML	ML	ML	ML	ML	ML	MH	ML	MH	MH	MH	MH	MH	MH	MH	MH	MH	MH
Zimbabwe	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	MH	L	L	MH
Summary																						
Number of low debt	32	30	30	29	26	23	23	19	17	18	13	15	14	15	14	13	11	7	8	9	11	9
Number of low multilateral debt	2	4	4	5	7	9	8	13	12	13	14	12	12	11	10	11	9	8	8	8	7	7
Number of high multilateral debt	0	0	0	0	3	4	5	4	6	6	9	8	9	10	12	12	16	20	20	19	18	20
Total	34	34	34	34	36	36	36	36	35	37	36	35	35	36	36	36	36	35	36	36	36	36

Note: The three debt classification are constructed as follows: L, the low debt regime characterizes years where a country's debt to GDP ratio is below 62.8 percent; ML, the high debt, low multilateral category regime characterizes years where a country's total debt to GDP ratio is greater than 62.8 percent and the share of multilateral debt in total debt is less than 41.2 percent; and MH, the high debt, high multilateral regime characterizes years when a country's total debt to GDP ratio is greater than 62.8 percent and the share of multilateral debt in total debt is greater than 41.2 percent.

Source: Authors' classification based on World Bank (various years).

FIGURE 1. Average Ratio of Net Transfers to GDP by Creditor Class, 1977–98



Note: See note to table 2 for debt classification.
Source: World Bank (various years).

Table 3 shows the means (and standard deviations) for the net transfers variable for all countries and for the three debt regimes, averaging over all years. Consistent with figure 1, net transfers have been systematically higher for the high debt countries, especially the high multilateral cases. For the whole period high multilateral debt countries received net transfers equal to some 18 percent of GDP, compared with some 10 percent of GDP for the low debt group. The low multilateral debt cases (a subdivision of the high debt regime) received some 15 percent of GDP.

III. CREDITOR AND DONOR SELECTIVITY: EMPIRICAL ANALYSIS

The aggregate statistics indicate that countries that found themselves with higher debts, especially to multilateral institutions, received larger net transfers than other countries. Higher net transfers to more indebted countries need not mean, however, that there was some kind of inefficiency. For one, some unobserved country characteristics, such as a high level of poverty or good policies, may have led donors to provide high net transfers. These high net transfers in the past may have been in the form of high lending in the past, resulting in a high debt stock. So countries highly indebted today may have been receiving more net transfers in both the past and present because they have a higher degree of poverty or better policies. We would call the higher transfers inefficient, however, if they were going to highly indebted countries independent of the quality of their domestic policy, institutional capacity to productively absorb flows, and poverty level. In that case, the high indebtedness, particularly to multilateral creditors, could be a barrier to selectivity in lending across countries.

TABLE 3. Means and Standard Deviations for Sample Countries, 1977–98 Average

					High debt			
	All		Low debt		Low multilateral debt		High multilateral debt	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Net transfers/GDP	0.135	0.119	0.104	0.087	0.146	0.151	0.183	0.117
CPIA	2.763	0.813	2.963	0.831	2.488	0.698	2.895	0.670
PVTDSGDP	0.547	0.339	0.323	0.160	0.837	0.308	0.651	0.306
GDPCAP	753	1,082	981	1,314	635	919	321	128
POP	10,824,872	16,612,134	8,527,659	14,559,811	16,741,602	23,105,792	7,601,202	7,240,667
No. observations	848		394		239		215	

Note: The net transfers variable is in U.S. dollars and is scaled by U.S. dollar GDP. CPIA is the policy measure from the World Bank. *PVTDSGDP* is the present value measure of all future scheduled debt service payments relative to GDP. $GDPCAP_{ij}$ is GDP per capita, measured in thousands of dollars. *POP* is population size. See note to table 2 for debt classifications.

Source: Authors' computations based on World Bank (various years), for debt data; IMF, various years, *International Financial Statistics*, for GDP per capita; and World Bank, various years, *World Development Indicators*, for population.

Policy Variables

We therefore need to control for country characteristics in our net transfers comparisons. For policy, we use an explicit measure of the quality of the policy environment in each country in each year: the World Bank's Country Policy and Institutional Assessment (CPIA). This measure, calculated annually by World Bank country specialists, has 20 components measuring macroeconomic, sectoral, social, and public sector institutions and policies on a scale of 1 to 6. It is determined on the basis of criteria standardized across countries. A separate World Bank unit makes a considerable effort to ensure consistency and comparability across countries and over time.[7]

The CPIA has been used by other researchers in investigating the determinants of aid flows, most notably Burnside and Dollar (2000) and Collier and Dollar (2002). It is considered a meaningful measure of the quality of the policy framework in a country in each year. The CPIA also has the advantage of including not only criteria related to public policy effort but also those related to institutional capacity and governance (rule of law, anticorruption efforts). It may thus be more closely related to a country's capacity to absorb transfers effectively than traditional measures of policy effort (such as trade and financial liberalization, privatization, and inflation).[8]

A disadvantage of the CPIA index is that it may be influenced by incentives to affect (indirectly) the lending behavior of the World Bank. Specifically, country managers may want (and be able) to improve their countries' ratings to justify a larger lending program for their country. To the extent this is true, however, the link between policy as measured by the CPIA and lending will be overstated, and any result showing that policy is not a factor will thus be a stronger result. An additional disadvantage is that the actual CPIA is available only to the public at the country level in more aggregated form, so that our results can be replicated only inside the World Bank.

We therefore use as a robustness test the policy index created by Burnside and Dollar (2000), based on publicly available information. Specifically, the index weights the following three variables: budget surplus as a share of GDP, the rate of inflation, and the degree of openness of the economy, as measured by Sachs and Warner (1995).[9] This alternative policy index is a somewhat cruder measure of the quality of policy and institutional environment and we can expect results to be less strong. Nor can the policy index be created for all countries and all years because some of the component variables are not always available.

7. The ratings have an element of judgment that may be affected by specialists' separate knowledge of a country's actual or likely overall prospects. This makes the ratings potentially endogenous to, for example, growth, though probably less to net transfers in a particular year.

8. In the absence of any good argument for alternative weighting of the components, we use the average. Collier and Dollar (2002) show that their results regarding aid allocation and poverty are not sensitive to reweighting the components.

9. The equation developed by Burnside and Dollar (2000) is: Policy $= 1.28 + 6.85 *$ Budget Surplus $- 1.40 *$ Inflation $+ 2.16 *$ Openness.

Because of missing data the sample of country-year observations for the policy index is much smaller, 484 compared with 848 for the CPIA sample.

To investigate the relationship between net transfers and policy by different categories of indebtedness, we first use the CPIA index to classify countries into two groups, those with a CPIA of less than 3, so-called bad policy countries, and those with a CPIA equal to or greater than 3, so-called good policy countries. We do this for the whole group as well as separately for the three debt categories. The split for the whole group is 46 percent/54 percent—that is, about half of the country-year observations are bad policy country-years and half are good policy country-years. We then calculate the mean net transfers as a share of GDP for each subclassification. We find that there are some statistically significant differences in net transfers between bad and good policy countries (table 4, top).

For all countries combined, bad policy countries receive on average some 1.5 percentage points of GDP more in net transfers than good policy countries do. This difference is also statistically significant.

There are no statistically significant differences in net transfers between bad and good policy countries for the groups of low debt and low multilateral debt countries. The difference in net transfers for the overall sample seems to be caused by the behavior of net transfers to the high multilateral debt countries, for which there is a difference between bad and good policy countries of 4.1 percentage points of GDP. This difference is statistically significant, with a t-statistic of 2.6. In other words, countries highly indebted to multilaterals received more net transfers when their policies were worse than average.

When we use the sample for the alternative policy index, we find similar effects for the debt breakdowns but different effects for the further breakdown by the policy index (table 4, bottom). The comparisons in net transfers between the debt groups are qualitatively similar to the overall comparisons: the low debt group receives net transfers equal to 8.4 percentage points of GDP and the low multilateral debt group 9 percentage points, whereas the high multilateral debt group receives 15.7 percentage points, a difference of 7 percentage points over the low debt group.

The breakdown between good and bad policy countries does not indicate the same results. For the high multilateral debt group, bad policy countries actually receive some 2.5 percentage points less in net transfers. For none of the debt classifications, however, is there a statistically significant difference between the bad and good policy countries, possibly because the sample is smaller and the index cruder, or because we have not yet controlled for other variables.

Other Country Characteristics

So far, the analysis has investigated only the simple interrelationships between the quality of policy, indebtedness, and net transfers. The raw statistics suggest that the degree and nature of indebtedness affect the relationship between policy and net transfers under the CPIA, although not under the alternative index. But the data remain suggestive because they do not yet control for other country characteristics and they vary by policy index and sample. The data in table 3 already indicate that

TABLE 4. t-Test on Net Transfers/GDP

	All		Low debt		Low multilateral debt		High multilateral debt	
	NT	Difference *t*-test	NT	Difference *t*-test	NT	Difference *t*-test	NT	Difference *t*-test
Based on CPIA								
All	0.135		0.103		0.146		0.183	
Good	0.127	0.015	0.103	0.001	0.145	0.001	0.163	0.041
Bad	0.142	1.810	0.103	−0.110	0.147	−0.220	0.204	2.610
Based on alternate policy variable								
All	0.103		0.084		0.090		0.157	
Good	0.103	0.000	0.080	0.008	0.079	0.016	0.168	−0.025
Bad	0.103	−0.051	0.088	0.868	0.095	0.803	0.143	−1.384

Note: The *t*-test measures the statistical significance of the difference in net transfers between the good and bad policy countries. The net transfers variable is in U.S. dollars and is scaled by U.S. dollar GDP. CPIA is the policy measure from the World Bank. See note to table 2 for debt classification.

Source: Authors' computations based on World Bank (various years).

it can be important to control for other country characteristics because there are large differences in GDP per capita and population size between the three groups of indebtedness countries. To properly assess the relationship between net transfers and the quality of policy, we extend our analysis by performing a multivariate regression controlling for several country characteristics.

The first control variable is the degree of indebtedness. Because the stock of debt is simply a transformation of accumulated past borrowing, it could be that high debt reflects country circumstances or policies can explain large net transfers in the past. These characteristics could include former colonial ties (Alesina and Dollar 2000), strategic interests of donors, openness of the economy, policy stance, or other country factors. These country characteristics may lead countries that received high net transfers in the past, and thus have high debt stocks, to continue to receive net transfers today. Including past indebtedness will control for some of this tendency.

Second, we want to control for the degree of poverty in the country because donors may provide net transfers in relation to a country's degree of poverty irrespective of policy indebtedness and other country factors. Controlling for the incidence of poverty is difficult, however, because poverty data are not available for long time periods. We have only a small, cross-country sample of poverty data for the 1990s. We therefore use instead GDP per capita (in dollars) as a proxy for the incidence of poverty (as well as the overall level of development of the country). The high multilateral debt countries are much poorer on average than the low debt countries, $320 per capita versus $980 per capita (see table 3). Perhaps they receive more net transfers simply because they are poor, and the relationship between net transfers and indebtedness arises because of greater poverty levels in high debt countries.

Third, we want to control for the size of economy Because there is a tendency for smaller countries to receive relatively more aid, as reported by Burnside and Dollar (2000) among others. (This relationship may arise because small countries tend to be more open and thus more exposed to international shocks, because donors expect greater influence on policy in smaller countries, or for other reasons.) In our specification, we use the log of population as a control variable for size.

To control for any other country differences and possible endogeneity in the relationships, we use a fixed-effects estimation technique. Fixed effects allow us to control for any unobserved country characteristics and to take care of any (remaining) endogeneity issues. We also employ ordinary least squares (OLS) as a test of robustness. We estimate the following model:

$$(2) \qquad NT_{ij} = \alpha + \beta_1 * PVTDSGDP_{ij} + \beta_2 * GDPCAP_{ij} + \beta_3 * LNPOP_{ij} + \beta_4 * CPIA_{ij} + \beta_5 * (\text{Debt classification interacted with } \textit{policy dummies}_{1,2,3} \text{ for country } i \text{ and year } j).$$

The net transfers variable for country i in year j, NT_{ij}, is scaled to GDP. *PVTDSGDP* is the present value measure of all future scheduled debt service

payments relative to GDP, which thus takes into account the concessionality of debt. This measure is preferred to such measures as the total debt stock to GDP, which ignores the concessionality of debt, or the ratio of annual debt service to GDP, which does not provide a measure of the full future debt burden. *PVTDSGDP* is taken from the GDF database. GDP per capita in thousands of dollars, *GDPCAP*, is used as a proxy for the overall level of development of the country and the incidence of poverty. Data on GDP per capita are from the IMF *International Financial Statistics*. Population size in logs, *LNPOP*, is to control for the size of the country and comes from the World Bank. CPIA is our policy measure variable already introduced (which we also substitute by the alternative policy index). The three debt dummy variables interacted with good and bad policy dummy variables are constructed consistent with tables 2 and 4: for all three debt groups (low debt; high debt, low multilateral; and high debt, high multilateral) we created a separate bad policy dummy variable if the CPIA for that country-year observation was less than 3 (or if the alternative policy index was less than its median).

Regression Results

The regression results show that net transfers are positively related to debt stocks, consistent with figure 1 and the raw statistics of average net transfers by country indebtedness classification of table 3 (table 5). This relationship may

TABLE 5. Basic Result for Net Transfers Regression

	CPIA				Alternative policy variable	
Variable	Fixed effect		OLS		Fixed effect	
PVTDSGDP	0.097	(8.26)	0.101	(9.47)	0.107	(6.23)
GDPCAP	−0.022	(−4.10)	−0.058	(−17.26)	−0.011	(−1.75)
INPOP	−0.114	(−7.97)	−0.041	(−17.12)	−0.091	(−4.95)
CPIA or policy	0.003	(0.63)	0.004	(0.60)	0.000	(0.04)
BPL	0.000	(0.00)	0.012	(1.13)	−0.004	(−0.56)
BPLM	0.009	(1.01)	−0.001	(−0.09)	−0.008	(−0.76)
BPHM	0.026	(2.86)	0.030	(2.54)	0.017	(1.77)
Constant	1.834	(8.42)	0.735	(16.33)	1.480	(5.21)
No. observations	848		848		484	
F-value	22.53		89.44		11.52	
R^2 (adjusted)			0.422			
Within	0.17				0.155	
Between	0.28				0.156	
Overall	0.17				0.046	

Note: Numbers in parentheses are *t*-statistics. The net transfer variable is in U.S. dollars and is scaled by U.S. dollar GDP. *PVTDSGDP* is the present value measure of all future scheduled debt service payments relative to GDP. *GDPCAP* is GDP per capita, measured in thousands of dollars. The coefficients for *GDPCAP* are multiplied by 1,000. *LNPOP* is population size, in logarithms. CPIA is the policy measure from the World Bank. See note to table 2 for debt classifications.

Source: Authors' computations based on World Bank, various years, GDF, for debt data; IMF, various years, *International Financial Statistics*, for GDP per capita; and World Bank, various years, *World Development Indicators*, for population.

reflect the defensive lending of donors, with high debt stocks triggering more net transfers to prevent arrears. But to prevent arrears it would be necessary only to roll over debt service due by providing an equal amount of new disbursements—that is, keep net transfers at 0. It would not be necessary to provide additional net transfers as debt burdens become larger.

The positive coefficient is also possible because some country characteristics not included in the regression may be correlated with past and current net transfers. But the fixed-effects regression technique should control for most of these characteristics. The positive relationship is not likely to be due to the fact that the net transfers figures we use are not corrected for any official debt reduction included by donors in the grant figures they report to the OECD, for that would lead to a downward bias.[10] Our preferred interpretation for the positive coefficient is that when debt stocks are large, countries manage to "bargain" for larger net transfers, perhaps because donors are more eager to prevent arrears or defaults that would increase the visibility of any failures in past lending programs.

We also find that countries tend to receive larger net transfers when they are poor—the coefficient on GDP per capita is statistically significant and negative.[11] There is also evidence of a small-country effect—the coefficient for population (in logs) is statistically significant and negative.

Surprisingly, we find no direct effects of the policy variable on net transfers—the coefficient for the CPIA is statistically insignificant.[12] It seems as if the quality of a country's policy does not affect the relative amount of net transfers it receives. When we analyze the effects of bad policy for the three debt groups, however, we find

10. Correcting the net transfers for debt reduction is difficult. Renard and Cassimon (2001, p. 20) write, "We have no strong basis to suggest a percentage by which the DAC figures must be reduced to give a correct expression of the costs of debt reduction to the donors." Even so, as also noted by Renard and Cassimon, the inclusion of official debt reduction in aid would upward bias the net transfers countries actually received. At the same time, the total debt reduction figure is reported in GDF and the debt stock is reduced by the official debt reduction. This would upward bias net transfers and downward bias *PVTDSGDP*, which would mean that the coefficient on *PVTDSGDP* would be downward biased, weakening the relationship. Besides biasing the coefficient on the debt stocks, there could be some other bias in our regression results as well. We therefore investigated whether the estimated relationship between net transfers and policy is affected in a systematic way by degree of debt reduction. Regression results (not reported) show that all estimated coefficients are robust to the inclusion of the GDF-reported debt reduction figures as another independent variable.

11. We multiply the coefficients on *GDPCAP* by 1,000 for presentational purposes. As noted, poverty data are not available for long time periods, but we do have a small cross-country sample of poverty data. When substituting poverty for *GDPCAP* for that sample, we find that countries with relatively greater poverty counts have larger net transfers.

12. Although this appears to contradict other findings, this insignificance of the policy variable mainly results from the fact that we do not use the nonlinear specification others have used. Collier and Dollar (2002) show that transfers (to all recipient countries) are nonmonotonic with respect to the CPIA; they rise for low and moderate levels of CPIA and then decline as CPIA improves. When we also allow for this nonlinearity, we find that about half of the countries lie on the upward sloping and another half on the downward sloping part of the curve. This means one cannot make a general statement on the effects of policy on net transfers.

that the high multilateral debt countries with bad policy receive about 2.5 percentage points *more* in net transfers as a share of GDP. For the other two groups, the coefficient for the bad policy dummy variables is not statistically significant. The 2.5 percentage points more in net transfers is less than the 4 percentage points reported earlier (table 4). Because the earlier comparisons did not control for country characteristics, some of the control variables may partly explain the higher net transfers.

The large remaining difference, equal to about one-fifth of the average net transfers all countries received over this period, nevertheless implies that in the high multilateral debt countries, bad policy is associated with more, not fewer net transfers. This result confirms our hypothesis that for countries highly indebted to the multilaterals, donors have not been able to practice selectivity for the policies countries have adopted (or for their institutional capability). Indeed, they have provided more (not fewer) resources where policies have been worse. Our interpretation: because of the large multilateral debts, donors have allowed poor policy to continue in these countries and actually provided more resources to accommodate the larger macro-imbalances, which in turn caused the higher debt stocks.

Robustness Tests

We confirm this result using a number of robustness tests. We start by reporting the OLS regression results (table 5). The regression results have the same signs but somewhat different statistical significance for the various control variables. The OLS results are generally stronger for the country control variables—no surprise, because the fixed-effects results control more for country differences. Again, the present value of debt to GDP has a positive statistically significant coefficient, and the level of GDP per capita and the population size (in logs) has a statistically significant negative coefficient. The policy variable, CPIA, again is not statistically significant, suggesting that the fixed-effects estimator does not hide permanent characteristics related to the quality of county policy. The bad policy dummy variable again remains statistically significant only for the high multilateral debt group and implying a 3 percentage point difference in net transfers. Our finding is thus robust to the estimation technique used.

We next investigate whether using the alternative policy index changes our results. We do so again using the fixed-effects regression techniques (see table 5). We find weaker statistical significance, possibly because the sample is reduced by half, from 884 country-year observations to 484. The coefficient on the dummy variable for the countries highly indebted to multilaterals with bad policy now indicates that these countries receive some 1.7 percentage points more in net transfers, still statistically significant at the 10 percent level. All other policy dummy variables and the policy variable itself are not significant. whereas the control variables retain their sign and significance. This result suggests that the use of a cruder policy index does not qualitatively change our main result.

As another robustness test, we investigate whether the relationship between debt and net transfers may have been affected by the occurrence of shocks and

the role of arrears. Many countries in the sample have faced large terms of trade shocks, frequently cited as one of the reasons for the poor economic performance, the need for continuing aid flows, and the poor policy records. If GDP were to decline systematically because of these shocks, this would not be a problem in our regressions.

But there could be a bias in our regression results if there is a relationship between shocks and the debt buildup. The argument would be as follows. Say that a country faces a surprise drop in GDP and cannot pay its debt service due. The debt to GDP ratio will rise because of the drop in GDP and the rise in debt (as arrears are capitalized into the debt stock). The net transfers to GDP ratio will also be high because GDP is low, and net transfers are possibly high due to arrears that reduce debt service paid and increases in new disbursements by creditors and donors to mitigate the impact of the shock. Because there have been quite big shocks to GDP for many of the countries in our sample, this pattern of net transfers to GDP being higher for countries and years when debt to GDP is also high could bias our coefficients.

There could also be a bias through the policy response, at least as measured. If countries adjust poorly to shocks and end up with larger fiscal deficits and higher inflation, policy (as measured) could be worse when net transfers are higher. This is possible particularly for the alternative policy index, which relies heavily on the budget surplus and inflation for its construction. Furthermore, donors may be more willing to accommodate poor policies when countries face adverse shocks, creating a similar bias.

To check for this possibility, we estimated a trend measure of annual GDP by running for each country a simple regression of actual GDP on a time trend.[13] We then replaced actual GDP in the various ratios (NT to GDP, PVTDS to GDP, and GDP per capita) by trend GDP as predicted by this regression. This way, the ratios are not affected by any short-run shocks to GDP. To investigate whether net transfers respond to short-run shocks to GDP, possibly confounding the results on the policy and other variables, we also included the deviation of GDP from trend as an additional independent variable in the regression. Table 6 reports the results, using again the fixed effects regression technique.

We find that using the permanent level of GDP in the ratios and adding the departure from trend in each year does not change our main result. The coefficient on the bad policy dummy variable for the countries highly indebted to multilaterals indicates that these countries receive some 2.1 percentage points more in net transfers, statistically significant at the 5 percent level. The coefficients on the control variables *PVTDSGDP*, *GDPCAP*, and *LNPOP* do not change much in magnitude and actually become somewhat more significant. Interestingly, the deviation from trend GDP, though positive, is not statistically significant. This suggests that net transfers are not being adjusted much in response to short-run

13. We thank the editor for suggesting this methodology.

TABLE 6. Regression of Net Transfers on Fitted GDP and of Gross Flows on Actual GDP

	Net transfers		Gross flows			
Variable	Fixed effects		Fixed effects		OLS	
PVTDSGDP	0.134	(14.68)	0.196	(16.41)	0.187	(19.31)
GDPCAP	−0.039	(−6.55)	−0.015	(−2.71)	−0.052	(−17.01)
INPOP	−0.135	(−9.67)	−0.123	(−8.4)	−0.037	(−17.11)
Residual	0.000	(1.26)	NA	NA	NA	NA
CPIA	0.009	(1.86)	0.006	(1.23)	0.010	(1.75)
BPL	0.003	(0.36)	0.003	(0.38)	0.012	(1.22)
BPLM	0.012	(1.43)	−0.003	(−0.34)	−0.016	(−1.5)
BPHM	0.021	(2.33)	0.015	(1.59)	0.015	(1.38)
Constant	2.125	(9.98)	1.939	(8.76)	0.655	(16.11)
No. observations	848		848		848	
F-value	51.55		50.22		141.96	
R^2 (adjusted)					0.538	
Within	0.342		0.307			
Between	0.341		0.337			
Overall	0.238		0.224			

Note: Numbers in parentheses are *t*-statistics. The net transfers variable is in U.S. dollars and is scaled by U.S. dollar GDP. *GF*, gross flows, is new disbursements and grant in U.S. dollars and is scaled by U.S. dollar GDP. *PVTDSGDP* is the present value measure of all future scheduled debt service payments relative to GDP. *GDPCAP* is GDP per capita, measured in thousands of dollars. The coefficients for *GDPCAP* are multiplied by 1,000. *LNPOP* is population size, in logarithms. CPIA is the policy measure from the World Bank. See note to table 2 for debt classifications. NA = not applicable.

Source: Authors' computations based on World Bank, various years, GDF, for debt data; IMF, various years, *International Financial Statistics*, for GDP per capita; and World Bank, various years, *World Development Indicators*, for population.

shocks to GDP. Even if they are, this pattern does not affect the overall relationship we stress—high indebtedness and poor policy lead to larger net transfers. In other words, donors do not necessarily adjust net transfers in response to shocks, but they are more willing to accommodate poor policy when multilateral debts are large. In that sense, our conclusions are reinforced.

We further confirm the importance of debt stocks in determining donor behavior by investigating the behavior of gross flows, instead of net transfers. We can rewrite the left-hand side variable of equation 1 as $NT = G + NB - DS$, with net transfers being new grants plus debt disbursements minus debt service. Taking into account that debt service is necessarily strongly related to the outstanding stock of debt, it is likely that gross flows, $NB + G$, are much more sensitive to debt stocks than net transfers are. To check for the role of debt stocks in gross flows, in the base regression we replace the dependent variable net transfers by the variable gross flows, also scaled by GDP, and rerun the regression. Table 6 reports these results, using fixed-effects and OLS regression techniques.

As expected, debt stocks, *PVTDSGDP*, are a very important determinant of gross flows, with a coefficient on debt stocks of 0.2, compared with 0.1 in the

base regression for net transfers. This confirms that creditors and donors provide gross flows in response to debt service due—that is, they roll over gross claims, which itself is closely related to debt stocks. But there is an additional effect related to debt classification and policy that adds to this tendency to provide gross flows. Although the coefficient on the bad policy dummy is less significant, it still indicates that gross flows are 1.5 percentage points higher to countries with poor policies highly indebted to multilaterals. Again, the results suggest that donors respond perversely to debt stocks, disbursing regardless of policy and other country circumstances, and providing more to countries with poor policies highly indebted to multilaterals.

Finally, we show the perverse relationships between net transfers and policy in the countries highly indebted to multilaterals by plotting the simple, univariate scatter of net transfers and the CPIA for each of the three indebtedness groups (figure 2; the results of simple univariate regressions are also reported). As can be seen, only for the countries highly indebted to multilaterals is there a negative relationship between net transfers and the CPIA. For the other indebtedness groups, the relationship is nonexistent.

We also address the question whether some donors and creditors are more sensitive to debt stocks and more selective for policy than others are. We do so by distinguishing among net transfers of five classes of donors or creditors:

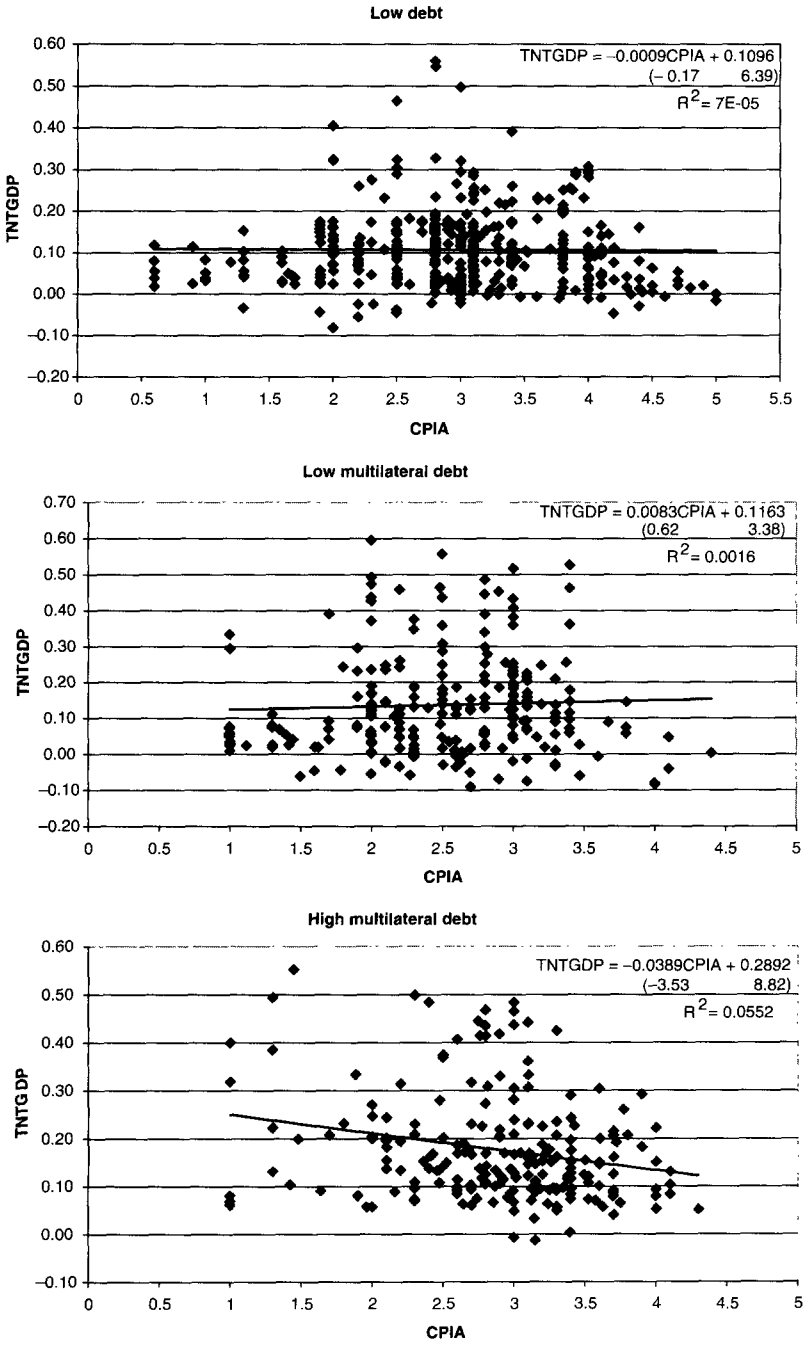
- The International Development Association (IDA) (concessional) and International Bank for Reconstruction and Development (IBRD) windows of the World Bank.
- The IMF.
- All multilaterals combined (IDA, IBRD, and IMF, as well as such other multilaterals as the African Development Bank and European Investment Bank).
- All bilaterals combined (including grants as well as loans).
- Private creditors (only the net transfers they provide to the governments are included).

Table 7 shows the results, using the same regression specification as in table 5 and employing the fixed-effects estimator but running separate regressions for each class of creditor.

For the net transfers from the IDA window and the bilaterals, the coefficients for the debt stock (*PVTDSGDP*) variable are statistically significant and positive, suggesting an element of defensive lending and bargaining by the country. (Again, the effects of lagged net transfers, due to some unknown or unmeasured country characteristic leading to both high debt stocks and high current transfers, is less likely because we used fixed effects.)

For the IBRD and IMF, however, the coefficients for the debt stock variable are statistically significant and negative, suggesting their greater concern with creditworthiness, thus cutting back new lending to more indebted countries. This is similar to the negative coefficient for the net transfers behavior of the private sector, which can safely be assumed to be mostly concerned with getting repaid.

FIGURE 2. Net Transfers and CPIAs by Creditor Class



Note: See note to table 2 for debt classifications.
Source: Authors computations based on World Bank, various years, GDF.

TABLE 7. Results for Fixed Effects Net Transfers Regression by Creditor Class

Variable	IDA	IBRD	IMF	All multilateral	Bilateral plus grant	Memo: Private
PVTDSGDP	0.027 (12.83)	−0.003 (−2.38)	−0.008 (−3.12)	0.025 (5.99)	0.085 (9.30)	−0.013 (−2.57)
GDPCAP	0.000 (0.10)	−0.001 (−1.29)	−0.001 (−0.98)	−0.003 (−1.66)	−0.020 (−4.71)	0.001 (0.57)
INPOP	0.010 (3.85)	−0.013 (−7.81)	−0.013 (−4.22)	−0.033 (−6.42)	−0.048 (−4.26)	−0.034 (−5.67)
CPIA	0.002 (2.54)	0.001 (1.30)	0.003 (2.46)	0.006 (3.59)	−0.003 (−0.89)	0.000 (0.12)
BPL	−0.001 (−0.62)	−0.001 (−1.40)	0.001 (0.54)	−0.001 (−0.18)	0.002 (0.35)	−0.002 (−0.50)
BPLM	−0.004 (−2.38)	0.000 (−0.51)	0.003 (1.51)	0.003 (1.08)	0.005 (0.72)	0.001 (0.15)
BPHM	−0.001 (−0.50)	0.001 (0.74)	0.003 (1.64)	0.003 (0.99)	0.021 (3.07)	0.001 (0.29)
Constant	−0.156 (−4.05)	0.191 (7.85)	0.200 (4.21)	0.491 (6.36)	0.812 (4.80)	0.531 (5.79)
No. observations	848	848	848	848	848	848
F-value	45.93	14.62	7.12	13.95	23.53	8.98
R^2 (adjusted)						
Within	0.29	0.11	0.06	0.11	0.17	0.07
Between	0.00	0.09	0.00	0.21	0.34	0.07
Overall	0.07	0.02	0.00	0.05	0.27	0.01

Source: Authors' computations based on World Bank (various years) for debt data; IMF, various years, *International Financial Statistics*, for GDP per capita; and World Bank, various years, *World Development Indicators*, for population.

For all multilaterals combined, the sign for the debt stock variable is statistically significant and positive, suggesting that IDA and the other multilaterals compensate for the repayment-oriented behavior of the net transfers for the IMF and the IBRD. Combined, these donors are mainly involved in defensive lending, or they relate their net transfers to some other country characteristic not yet controlled for. The coefficients for the control variables, *GDPCAP* and *LNPOP*, are generally the same as for the overall net transfers regressions, but not always statistically significant. The exception appears to be IDA, for which there is a large-country effect, rather than a small-country effect.

The coefficients on the policy variable CPIA are statistically insignificant for the net transfers from the IBRD and bilaterals, as well as for the net transfers from the private sector. Net transfers from IDA and the IMF, and all multilaterals combined, relate in a positive way to the quality of policy, with the coefficients also statistically significant. The positive relationship for IDA can be expected because the allocation of IDA resources is explicitly linked to the quality of the policy framework. For the bilaterals, in contrast, high indebtedness to multilaterals when combined with bad policy has perverse effects. The bad policy effect for the high multilateral debt category for the bilaterals is large, some 2.1 percentage points of GDP. For the IMF, there is also a positive effect on net transfers from belonging to the bad policy, high multilateral debt countries, but it is only 0.3 percentage points, significant only at the 10 percent level.

The IDA and the IBRD, as well as all multilaterals combined, have no significant bad policy effect, and IDA actually provides fewer net transfers to bad policy countries with high debt but a low share of multilateral debt. Because the bilaterals increase net transfers to GDP for the bad policy and high multilateral debt countries by some 2.1 percentage points, close to the total 2.5 percentage point effect for all net transfers (as reported in table 4), it seems that the higher net transfers going to bad policy countries is almost entirely due to the bilateral donors.[14] Although we found the quantitative effects of the IMF to be small, the IMF is very important—

14. We also conducted OLS regression, with results essentially the same overall as the net transfers regression. When using our alternative policy index, we find that policy is no longer significant for any creditor or donor class. Bilaterals and IMF still transfer some 1.3 and 0.4 percentage points, respectively, more to bad policy, high multilateral debt countries, whereas IDA transfers less to bad policy countries in all three indebtedness classes and IBRD transfers somewhat less to the bad policy, high multilateral debt countries. When using the trend GDP to calculate the ratios and adding the deviation from trend GDP as another regression variable, we find that the bilaterals transfer 1.9 percentage points more to the bad policy, high multilateral debt countries and IDA transfers less to bad policy countries in the high debt, low multilateral indebtedness class. We also find some evidence of smoothing of income shocks for net transfers from IDA and IMF as the coefficients on the residual GDP are positive for these two classes. Policy remains insignificant for all classes, however, except for net transfers from private creditors, where it is positive. Finally, regression results for gross flows (instead of net transfers) as dependent variable show that the coefficient for the debt stocks is the highest for the bilaterals and that the bilaterals provide 1.7 percentage points more in gross flows to the bad policy, high multilateral debt countries. IDA gross flows are less to bad policy, low multilateral debt countries but respond positively to policy, as do IMF gross flows. All these results are not reported but are available from the authors.

because in the Paris Club reschedulings and donor meetings, an IMF program is almost always required. Seemingly, the IMF provides the signal of accommodating bad policy more easily for countries that have high debts to the multilaterals.

As a last robustness test, for each class of indebtedness we ran separately the base regression without the policy-debt interaction dummy variables to investigate the behavior of each creditor or donor with respect to policy for each type of country (not reported). The sample sizes vary (394 for the low debt group; 239 for the high debt, low multilateral group; and 215 for the high debt, high multilateral group). We find net transfers from IDA to be significant positively related to policy for the low debt and the high multilateral debt group (though only at the 10 percent level for the latter). For the net transfers from the IMF, the policy variable is statistically significant and positive only for the low multilateral debt group, whereas for the bilaterals the policy variable is actually statistically significant negative for the low multilateral debt group. For all multilaterals combined, the policy variable is positive and significant for the low debt and the low multilateral debt groups but not for the high multilateral debt group. This confirms again that donors have more difficulty being selective about policy for the high multilateral debt countries, actually acting perversely, but in low debt countries they have been able to be selective.

IV. Conclusions and Implications

Net transfers remained positive over two decades in most Sub-Saharan countries, falling only somewhat in the 1990s. But with low growth in recipient countries, continuing net transfers meant a rising stock of debt relative to output. The bilateral donors tried to minimize the resulting burden of debt service by shifting to grants and offering repeated rounds of debt service relief.[15] The multilateral institutions were more restricted in ensuring repayments to themselves with new loans, because these loans would only increase the debt burden. At the same time, the multilaterals could not accept arrears. Bilateral donors in effect were ensuring that some of the poorest countries would not be pushed into arrears to the multilaterals, despite poor policies.

Countries "benefited" from this need to avoid arrears because they could bargain for more net transfers as their debt stocks increased. The core reason was that arrears to the multilaterals are a problem for all donors because they signal the end of business as usual for the other donors as well—and would be seen as a failure of the development assistance business in that country. Importantly, this need to maintain net transfers came at the cost of losing selectivity for country policy, especially for countries with bad policy highly indebted to

15. As some countries' GDP per capita fell (Côte d'Ivoire, Nigeria), some countries became eligible for cheaper IDA loans from the World Bank and concessional loans from the African Development Bank. That also helped minimize the burden of increasing debt service.

multilaterals. For these countries, the development community actually accommodated poor policies through higher net transfers.

This interpretation has important policy implications. It suggests that if debt levels are reduced enough in countries with high multilateral debt, the behavior of the donor community can then shift into a low debt regime for those countries—a regime that in the past has allowed selectivity for multilaterals, at least for IDA. Debt service reduction under HIPC (now enhanced under HIPC II) can thus be interpreted as a way out for a donor community otherwise locked into nonselectivity in the high multilateral debt countries.[16]

Debt relief, by encouraging selectivity and changing donor behavior for the better, can ensure more funds for countries currently indebted but with good policies. As others have shown, providing additional resources to countries with good policies would help enhance their growth and lead to more poverty reduction. Without additional donor resources, selectivity would imply fewer funds for countries with bad policies.[17] Ironically, the return to more selective transfers may avoid the full costs of the debt reduction program. That would be the case if the HIPC Initiative is "paid for" out of traditional donor financing, leading to lower future transfers to HIPC or other poor countries. The fact that some of the "grant" figures today already include an official debt reduction suggests that donors see grants and debt reduction to some extent as substitutes.

Though necessary, debt reduction is far from sufficient to ensure donor selectivity. There is a need, particularly following what may become "nonselective" debt relief, for greater emphasis on selectivity in future grantmaking and lending. This will require deep institutional changes on the part of the donors, in their bilateral programs and in their influence on the multilaterals. Fortunately, our analysis of the past behavior of creditors suggests that with debt reduction, this is at least possible. Donors can make the necessary break with past practice—and thus increase their contribution to the tremendous development challenges in Africa. Better donor behavior would also set the stage for more effective development assistance in the long run, making it politically possible to convince the public in donor countries to maintain and even raise development assistance budgets.

16. Debt relief under the HIPC Initiative is unlikely to directly free resources in high debt countries for spending on the poor. It has not been high debt burdens that have constrained resources transfers. High debt countries have been receiving more net transfers relative to other poor countries. Debt reduction alone does not free up resources for countries if debt is not being effectively serviced to begin with. But it is likely that in some countries transfers in the form of multiple donor projects, often tied to donor-supplied services, were not a good substitute for the direct increase in revenues that debt relief provides (through a decrease in tax-financed debt service). Birdsall and Williamson (2002) argue that in well-managed countries, debt relief is more efficient as a form of aid than are donor-financed projects.

17. The latter group could end up with reduced net transfers simply because the debt reduction will make it easier for donors and creditors to reduce what our evidence suggests is now forced defensive lending. Debt reduction can also more effectively create a virtuous circle by crowding in private flows to good policy/low debt countries (Birdsall and Williamson 2002). See also Berlage and others (2003).

References

Alesina, Alberto, and David Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5(1):33–63.

Berlage, Lodewijk, Danny Cassimon, Jacques Dreze, and Paul Reding. 2003. "Prospective Aid and Indebtedness Relief: A Proposal." *World Development* 31(10):1635–54.

Birdsall, Nancy, and Johan Williamson, with Brian Deese. 2002. *Delivering on Debt Relief: From* IMF *Gold to a New Aid Architecture*. Washington, D.C.: Center for Global Development and Institute for International Economics.

Burnside, Craig, and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):847–68.

Claessens, Stijn, and Ishac Diwan. 1994. "Recent Experience with Commercial Debt Reduction: Has the 'Menu' Outdone the Market?" *World Development* 22(2):201–13.

Claessens, Stijn, Enrica Detragiache, Ravi Kanbur, and Peter Wickham. 1997. "Heavily-Indebted Poor Countries' Debt: Review of the Issues." *Journal of African Economies* 6(2):231–54.

Collier, Paul, and David Dollar. 2002. "Aid Allocation and Poverty Reduction." *European Economic Review* 46(9):1475–500.

Diwan, Ishac, and Dani Rodrik. 1992. *External Debt, Adjustment, and Burden Sharing: A Unified Framework*. (Princeton Studies in International Finance, 73). Princeton University Department of Economics: Princeton, N.J.

Easterly, William. 1999. "How Did Highly Indebted Poor Countries Become Highly Indebted? Reviewing Two Decades of Debt Relief." Policy Research Working Paper 2225. World Bank, Development Research Group, Washington, D.C.

Eaton, Jonathan, and Raquel Fernandez. 1995. "Sovereign Debt." In Gene M. Grossman and Kenneth S. Rogoff, eds., *Handbook on International Economics*, vol. 3. North Holland: Elsevier.

Hansen, Hendrik, and Finn Tarp. 2000. "Aid Effectiveness Disputed." *Journal of International Development* 12(3):375–98.

———. 2001. "Aid and Growth Regressions." *Journal of Development Economics* 64(2):545–68.

Renard, Robrecht, and Danny Cassimon. 2001. "On the Pitfalls of Measuring Aid." WIDER Discussion Paper 2001/69. United Nations University/World Institute for Development Economics Research, Helsinki.

Sachs, Jeffrey, and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* (1):1–118.

World Bank. 1998. *Assessing Aid—What Works, What Doesn't Work, and Why*. World Bank Policy Research Report. New York: Oxford University Press.

———. 2000. *Can Africa Claim the 21st Century?* Washington, D.C.: World Bank.

———. Various issues. *Global Development Finance*. Washington, D.C.: World Bank.

The World Bank Economic Review
Author Index to Volume 17, 2003

Number 1: 1–143
Number 2: 145–316
Number 3: 317–442

The World Bank Economic Review
Title Index to Volume 17, 2003

Number 1: 1–143
Number 2: 145–316
Number 3: 317–442