THE WORLD BANK
ECONOMIC REVIEW

Volume 20 • 2006 • Number 1

Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality
Alain de Janvry and Elisabeth Sadoulet

Child Labor and School Achievement in Latin America
Victoria Guimarães, Peter F. Oravem, and Mario A. Sánchez

The Long-Run Economic Costs of AIDS: A Model with an Application to South Africa
Clive Bell, Shantayanan Devarajan, and Hans Gersbach

Robust Multidimensional Spatial Poverty Comparisons in Ghana, Madagascar, and Uganda
Jean-Yves Duclos, David Sabn, and Stephen D. Younger

The Impact of Regulatory Governance and Privatization on Electricity Industry Generation Capacity in Developing Economies
John Cubbin and Jon Stern

An Empirical Analysis of State and Private-Sector Provision of Water Services in Africa
Colin Kirkpatrick, David Parker, and Yin-Fang Zhang

www.wber.oxfordjournals.org
The World Bank Economic Review is a professional journal for the dissemination of World Bank-sponsored and other research that may inform policy analysis and choice. It is directed to an international readership among economists and social scientists in government, business, international agencies, universities, and development research institutions. The Review seeks to provide the most current and best research in the field of quantitative development policy analysis, emphasizing 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 illustrate how professional research can shed light on policy choices. Consistency with World Bank policy plays no role in the selection of articles.

Articles are drawn from work conducted by World Bank staff and consultants and by outside researchers. Non-Bank contributors are encouraged to submit their work.

Before being accepted for publication, articles are reviewed by three referees—one from the World Bank and two from outside the institution. Articles must also be endorsed by two members of the Editorial Board before final acceptance.


Instructions for authors wishing to submit articles are available online at www.wber.oxfordjournals.org. Please direct all editorial correspondence to the Editor at wber@worldbank.org.
Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality
Alain de Janvry and Elisabeth Sadoulet

Child Labor and School Achievement in Latin America
Victoria Gunnarsson, Peter F. Orazem, and Mario A. Sanchez

The Long-Run Economic Costs of AIDS: A Model with an Application to South Mica
Clive Bell, Shantayanan Devarajan, and Hans Gersbach

Robust Multidimensional Spatial Poverty Comparisons in Ghana, Madagascar, and Uganda
Jean-Yves Duclos, David Sahn, and Stephen D. Younger

The Impact of Regulatory Governance and Privatization on Electricity Industry Generation Capacity in Developing Economies
John Cubbin and Jon Stern

An Empirical Analysis of State and Private-Sector Provision of Water Services in Mica
Colin Kirkpatrick, David Parker, and Yin-Fang Zhang
Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality

Alain de Janvry and Elisabeth Sadoulet

Conditional cash transfer programs are now used extensively to encourage poor parents to increase investments in their children's human capital. These programs can be large and expensive, motivating a quest for greater efficiency through increased impact of the programs' imposed conditions on human capital formation. This requires designing the programs' targeting and calibration rules specifically to achieve this result. Using data from the Progresa randomized experiment in Mexico, this article shows that large efficiency gains can be achieved by taking into account how much the probability of a child's enrollment is affected by a conditional transfer. Rules for targeting and calibration can be made easy to implement by selecting indicators that are simple, observable, and verifiable and that cannot be manipulated by beneficiaries. The Mexico case shows that these efficiency gains can be achieved without increasing inequality among poor households.

Conditional cash transfer programs targeted to poor households have become widely used, in particular to induce beneficiary households to invest in their children's human capital. The approach presumes that the supply side of social services for education and health is in place and that stimulating demand through income effects is insufficient to induce major changes in human capital investment (Bourguignon, Ferreira, and Leite 2002). Instead, a condition that transforms the income effect into a price effect needs to be attached to the cash transfer. In this case, receiving the transfer requires meeting school attendance and health practice requirements.

This approach has been hailed as a major innovation in organizing poverty-reduction programs. Well-known programs that follow this approach include Progresa (now called Oportunidades) in Mexico, Bolsa Escola and Bolsa Familia in Brazil, Red de Protección Social in Nicaragua, Programa de Asistencia...
Familiar in Honduras, the Program of Advancement through Health and Education in Jamaica, Food-for-Education in Bangladesh, and Subsidio Unico Familiar in Chile (Ravallion and Wodon 2000; Skoufias 2000; Morley and Coady 2003; Rawlings and Rubio 2005). Some of these programs have become very large and expensive. In 2004 Oportunidades served 4 million families at a cost of $2.2 billion. In 2001 Bolsa Escola covered 4.8 million families at a cost of $700 million. While few programs have been rigorously evaluated, Progresa was found to have a positive impact on education (Schultz 2004) and health (Gertler 2004). However, almost no analysis has been conducted on the effectiveness of alternative program designs in achieving these results, despite the large sums spent to obtain them. This issue is addressed here by analyzing whether better targeting of qualifying poor households and better calibration of the levels of cash transfers could help raise program efficiency.

Conditional cash transfer programs have a dual objective: immediate poverty reduction through transfers and long-term poverty reduction through investment in human capital. Efficiently meeting the first objective requires the transfers to be accurately targeted to poor households—a difficult task not addressed here (van de Walle 1998; Alderman 2001, 2002; Ravallion 2003). Meeting the second objective requires accurately selecting among poor households to minimize the efficiency leakages from payments to children already highly likely to attend school without a transfer (as opposed to children who will attend school only because of the transfer) and offering a level of transfer that is sufficient to meet the opportunity cost of the change in behavior, thus securing a high uptake while minimizing project costs. The specific concerns addressed here are the definition of targeting and calibration rules that can be easily implemented and that are cost-effective and transparent, and the potential tradeoffs between efficiency gains through implementation of these rules and higher inequality in transfers among poor households.

Progresa is used to explore these alternative program designs. Efficiency gains of 29-44 percent over the current program are found to be possible without increasing inequality among poor households.

I. The Efficiency Issue in Progresa

Progresa is a conditional cash transfer program for human capital formation targeted to poor rural children. It consists of three closely related components to address education, health, and nutrition issues: (a) a monetary transfer to each child under age 18 who regularly attends school between the third year of primary school and the third year of secondary school and who regularly visits a health center, (b) basic healthcare for all family members, and (c) nutritional supplements for children and women in need.

Progresa was introduced in 1997 and by 2000 fully covered marginal rural municipalities, reaching 2.6 million families. The budget for 2000 was $950 million—44 percent of it for education transfers, which benefited approximately
1.6 million children in primary school and 800,000 in secondary school (Coady 2000).

The transfers that Progresa families receive significantly increase their income—by an average of 22 percent. Progresa has explicitly targeted poor households in marginal rural areas of Mexico. The purpose here is not to question this targeting, which corresponds to Progresa’s objective of transferring resources to poor families. The purpose is to explore whether, for a given budget constraint, targeting and calibrating transfers to poor households can more efficiently increase school participation. Consequently, Progresa’s education component is used as a laboratory to explore alternative targeting and calibration rules. The idea is to derive lessons from this richly informed experiment that can be applied to Progresa and to other conditional cash transfer programs where severely limited budgets make targeting critical.

To measure its impact, Progresa selected a sample of 506 marginal communities containing 24,000 households and 17,000 children eligible for transfers and surveyed them a year before the program started and, subsequently, every six months over three years. Information was collected on individual, household, and community characteristics. The sample design consists of the random selection of 320 treatment communities and 186 control communities. The analysis is restricted to children in school in October 1997. Twelve percent of eligible children had dropped out of school by that time, some several years earlier, and while the program helped bring some of them back to school, this one-time effect at the onset of the program is not the focus of the analysis. The sample was further restricted to the 2,242 poor children who graduated from primary school in the summer of 1998 and faced the decision of whether to continue to secondary school. These data are used to estimate a model of the school enrollment decision that captures, in particular, the impact of Progresa transfers, paying particular attention to heterogeneity of conditions among children. Alternative targeting and transfer programs are then simulated and their efficiency is compared.

Focusing on Entry into Secondary School

A simple analysis of the overall Progresa budget suggests that an efficient program for school enrollment should focus strictly on the transition from primary school to secondary school, a point already suggested in the International Food Policy Research Institute evaluations (Skoufias 2000; Coady 2000; Schultz 2004).

The conditional transfer offered to each child is calculated according to the program’s rules. The program has a schedule of education transfers that increase as children progress to higher grades and that are higher for girls than for boys in secondary school (table 1). There is, however, a maximum total

---

1. The 'conditional transfer' is the exogenous amount that a child would receive from Progresa if he or she was in a treatment community and attended school. It depends on the gender and grade of the child and the household's demographics. At the household level the conditional transfer is the total amount that the household would receive if it was in a program community and complied with all Progresa rules.
### Table 1. Budget for Educational Transfers: Progresa in Sample Villages, 1998

<table>
<thead>
<tr>
<th>School Years that Children Could Attend</th>
<th>Number of Eligible Children(^a)</th>
<th>Schedule of Annual Conditional Transfer (US$)</th>
<th>Continuation Rate (percent)</th>
<th>U.S. Dollars(^b)</th>
<th>Share of Total (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary 3</td>
<td>1,909</td>
<td>70</td>
<td>98.2</td>
<td>114,229</td>
<td>11.8</td>
</tr>
<tr>
<td>Primary 4</td>
<td>1,811</td>
<td>80</td>
<td>97.8</td>
<td>120,260</td>
<td>12.4</td>
</tr>
<tr>
<td>Primary 5</td>
<td>1,613</td>
<td>100</td>
<td>97.1</td>
<td>135,626</td>
<td>14.0</td>
</tr>
<tr>
<td>Primary 6</td>
<td>1,476</td>
<td>135</td>
<td>97.4</td>
<td>166,035</td>
<td>17.2</td>
</tr>
<tr>
<td>Secondary 1</td>
<td>1,416</td>
<td>200/210(^c)</td>
<td>76.7</td>
<td>189,602</td>
<td>19.6</td>
</tr>
<tr>
<td>Secondary 2</td>
<td>752</td>
<td>210/235(^c)</td>
<td>96.1</td>
<td>134,884</td>
<td>14.0</td>
</tr>
<tr>
<td>Secondary 3</td>
<td>551</td>
<td>220/255(^c)</td>
<td>96.7</td>
<td>106,028</td>
<td>11.0</td>
</tr>
<tr>
<td>Total</td>
<td>9,528</td>
<td></td>
<td></td>
<td>966,664</td>
<td>100</td>
</tr>
</tbody>
</table>

\(^a\)Only children enrolled in school in 1997.

\(^b\)Taking into account the cap on total household conditional transfers. With 10 monthly conditional transfers per school year and an exchange rate in October 1998 of 10 pesos per US$, all transfers can be read as either pesos per month or US$ per year.

\(^c\)Conditional transfers for boys/girls.

**Source:** Authors' analysis based on the 1998 Progresa survey.
conditional transfer that each household can receive, set at $625 a year in 1998 (including $100 for nutrition). In the sample, 13.4 percent of eligible children are subject to the household transfer cap. Using the proportionality rule that Progresa applies, the conditional transfer corresponding to each child is calculated by scaling down all the education transfers in a household subject to the cap by the same factor. Among the children graduating from primary school, 28 percent are subject to the cap, and the conditional transfers vary from $100 to the full $200/$210, with an average of $169. The budget for education transfers in the sample treatment communities is calculated using these conditional transfers and the enrollment status of each child (table 1). The budget would be 17 percent higher, with no cap on total household conditional transfer. Taking into account these caps, transfers to primary school children account for 55.4 percent of the total budget for education transfers and the first year of secondary school for almost 20 percent.

Other studies show that Progresa's conditional transfers increase continuation rates at all grades (Behrman, Sengupta, and Todd 2001; de Janvry, Finan, and Sadoulet 2001; Schultz 2004). However, school continuation rates were already very high in both primary and secondary school before the Progresa intervention (figure 1). The increase in continuation rates that result from the conditional transfers is only around 1 percentage point in primary school and half a percentage point in the second and third years of secondary school. This suggests that, from an efficiency standpoint, Progresa is unnecessarily expensive for primary school, with 96 school-attending children paid for each child that is retained in school by the conditional transfer incentive, for an implied effective cost per additional child attending primary school of $9,600. Assisting the 34 percent of children who drop out of school at each grade level would require a very different program. Eliminating all transfers to primary school students would have saved 55.4 percent of the education transfer budget, or more than $230 million of the $950 million budget in 2000.

The critical problem in terms of education achievement occurs at entry into lower secondary school. Thus the analysis here continues only for secondary school.

The Efficiency of Progresa's Education Transfers

There are two sources of inefficiency in a conditional cash transfer program that need to be optimally reduced. The first is paying people for what they were already going to do. With Progresa, this is obvious in primary school. But the problem also arises in secondary school: 64 percent of the poor children who graduate from primary school would enter secondary school without a transfer. Reducing this efficiency leakage requires being able to predict who will not

---

2. This cap was introduced so that the program does not induce a fertility response.
3. For households subject to the cap all conditional transfers are scaled down by a common factor so that they add up to the cap. This prevents the households from keeping a child out of school without penalty.
continue in school, using a model for the probability that a given type of child will enroll in school. Because such predictions are inevitably noisy, there is no way to completely avoid this inefficiency. The challenge, however, is to reduce it by not targeting children most likely to attend school anyway.

The second source of inefficiency comes from offering transfers that are either too high or too low relative to the minimum amount needed to induce the conditional action. As shown later, the simple difference estimation of the impact of Progresa indicates that the program raised the enrollment rate from 63.6 to 76.6 percent. The conditional transfers offered were thus sufficient for the 13 percent of children in the sample who were attracted to enroll in secondary school and would not have done so otherwise. Could a smaller conditional transfer have had the same effect? For the 23.4 percent that did not take the conditional transfer, would a higher transfer have induced them to accept the offer? If so and if these children can be identified, should the conditional transfer offered to them be increased?

4. This inefficiency concept is analogous to the issue of fungibility with inframarginal transfers, whereby beneficiaries substitute other commodities for those subsidized by the program, meaning that the program has no real effect on total consumption of the targeted commodity.
If there were a clear opportunity cost to children's time in school, the subsidy could be calibrated to match it. For Progresa the transfer is 40 percent of what children of the same age earn when they work. But the opportunity cost of children's time at school is not easy to calculate. Less than 30 percent of the children who drop out at the end of primary school work during the subsequent 18 months (45 percent of boys and 10 percent of girls), increasing to 35 percent (55 percent of boys and 12 percent of girls) the following year. Lack of money or need to work is the most common reason given for not continuing school (57 percent), but other important reasons are that the child does not like school or does not learn (23 percent) and that the school is too far away (13 percent).

The Progresa randomized experiment allows the estimation of the children's response function, which is necessary to design the transfer schedule that maximizes return to the program. Since no experiments were conducted to observe the response to different levels of conditional transfers, the cap on total conditional transfers to a household is used to infer the marginal response to varying conditional transfer amounts.

Dealing with these two sources of inefficiency requires an accurate predictive model of the probability of attending school as a function of the characteristics of the child, the household, and the community and of the amount of the conditional transfer offered. The analysis here concentrates on entry into secondary school since that is where the conditional transfer can induce an important behavior change that results in efficiency gains.

The conditional transfers offered to children in second and third years of secondary school are not questioned for two reasons. First, these conditional transfers are part of the expected benefits of entering secondary school, and the measured impact of Progresa thus includes their effect. Second, while very high continuation rates are observed in secondary school, they are for a selected group of children who voluntarily entered secondary school without any subsidy in 1998, before the program was in place. Other children who are induced to enter secondary school with a conditional transfer are very unlikely to continue at the same rate if the subsidies were discontinued. No experimental design allows this particular continuation rate to be studied here because Progresa always supported the first three years of secondary school. The safe bet is that whatever transfer is provided in the first year needs to be provided in all three years of secondary school (as is currently the case), and while in 1998 many fewer Progresa children were in the second and third years of secondary school than in the first year (because it was the first year of the program), these numbers should even out three years after program implementation. Thus the results of the analysis of the first year are applied to all three years of secondary school.

II. A Model of Optimal Cash Transfer

Let $P(X,Y)$ denote the probability that a child with characteristics $X$ and eligible for a conditional transfer $T$ will enroll in school. Eligibility is denoted
by the index function $I \in [0,1]$. Children's characteristics are distributed according to the density function $f(X)$.

The allocation problem consists in choosing the eligibility status $I(X)$ and, if eligible, the conditional transfer $T(X)$ offered to a child with characteristics $X$, to maximize the gain in enrollment over the population:

$$\max_{I(X), T(X)} \int [P(X, T) - P(X, 0)] f(X) dX$$

subject to a budget constraint:

$$\int P(X, T) T f(X) dX \leq B$$

where $B$ is the budget available for the program. The first order condition for the optimal conditional transfer is that, for any eligible child $(I = 1)$,

$$P_T - \lambda (P_T T + P) = 0$$

where $P_T = \partial P / \partial T$ and $\lambda$ is the Lagrange multiplier associated with the budget constraint. This relationship states that the ratio of cost $(P_T T + P) dT$ to enrollment benefit $(P_T dT)$ of a marginal increase in the conditional transfer $(dT)$ is equal for all children. Hence, the cost of the marginal child brought to school is equal across children types $X$. Note that the cost has two terms. The first term, $P_T T dT$, is the transfer cost to the $P_T dT$ marginal children brought to school by the increase in conditional transfer. The second term, $P dT$, is the cost of giving the increase in transfer, $dT$, to all $P$ children of the same type $X$, even though they enrolled in school with the initial transfer $T$. This is the marginal equivalent of the decomposition of the cost of transfers:

$$P(X, T) T = [P(X, T) - P(X, 0)] T + P(X, 0) T$$

where the first term is the cost of the transfers to the children brought to school by the conditional transfer and the second term the cost of transfers to the children with similar observable characteristics who would have gone to school anyway.

Given the optimal conditional transfer amount conditional on eligibility, the optimal eligibility rule is defined by:

$$Z = 1 \text{ if } [P(X, T) - P(X, 0)] T + \lambda P(X, T) T \geq 0$$
$$1 = 0 \text{ otherwise}$$

The optimal allocation of a budget $B$ is thus the solution to the system of equations (2), (3), and (5).

In the particular case of a linear probability model, which is used in the empirical work here, the conditional expectation of the enrollment probability is:
where $\delta_0I + X\beta T$ is the total impact of $T$ and $X$, which includes a constant term, is the marginal impact of $T$. The presence of an intercept $\delta_0I$ is due to the fact that only conditional transfers of $100–210$ are observed, meaning that the linearity of the conditional transfer effect cannot be extended below that range to a 0 conditional transfer.

The optimal conditional transfer and eligibility criteria defined in equations (3) and (5) are:

$$T = \max \left[ \frac{1}{2\lambda} - \frac{1}{2} \frac{X\beta + \delta_0}{X\delta}, 0 \right]$$

where $\lambda$ is the solution to the budget constraint in equation (2). The expression shows that both eligibility and the optimal conditional transfer for any given child are a function of the ratio

$$\frac{X\beta + \delta_0}{X\delta} = \frac{EP(X, 0) + \delta_0}{EP_T}$$

The first term in the numerator is the expected probability that children with characteristics $X$ would attend school even without a conditional transfer, and the denominator is the marginal effect of the conditional transfer on the expected enrollment probability. Children will thus be eligible for and be offered high conditional transfers when they have a low initial probability of enrollment or a high enrollment response to a conditional transfer. This optimal conditional transfer is a function of all the characteristics $X$ that predict enrollment, albeit in a very nonlinear form.

Whether any program could use such a complex formula to compute conditional transfers is questionable. But it is a useful benchmark because it gives the maximum efficiency that could be reached with the observables $X$, and it will thus be computed in the empirical analysis that follows. Next, however, the optimal program is constrained to be linear and to use a small number of observable characteristics.

**An Implementable Conditional Cash Transfer Program**

To be useful for program implementation, eligibility rules need to be simple and transparent. Indicators used to determine eligibility, and the level of conditional transfers must be few, easily observable and verifiable, and nonmanipulatable by households. Simplicity and transparency are also important to ensure the political acceptability of a subsidy program (Schady 2002). Progresa uses grade and gender to set the schedule of conditional transfers (table 1). The objective is thus to simplify the formula in equation (7) established for the optimal conditional cash transfer program to a linear index on the basis of a few characteristics $Z$ of the children.
The allocation problem lies in choosing the eligibility status and, if eligible, the conditional transfer to offer to each child to maximize the gain in enrollment over the population [equation (1)], subject to the budget constraint [equation (2)], and using simple linear formulas for eligibility and conditional transfer:

\[ T = Z \alpha \quad \text{and} \quad I = I[Z \gamma \geq \gamma_{\text{min}}] \]

where \( Z \) is a subset of characteristics of the children and \( \alpha \), \( \gamma \), and \( \gamma_{\text{min}} \) are parameters to be determined. As in the model above, optimal eligibility is defined by the sign of the optimal conditional transfer value:

\[ I = 1 \iff T = \max(Z \alpha, 0) > 0. \]

The parameters \( \alpha \) are a solution to the maximization of a quadratic function:

\[ \max_{\alpha} \sum_{i \in E} m_i Z_i \alpha - \lambda \left[ B - \sum_{i \in E} (P_{0i} + \delta_0 + m_i Z_i \alpha) Z_i \alpha \right] \]

where \( E \) is the set of eligible children, \( m_i = X_i \hat{\delta} \) is the marginal effect of the conditional transfer on child \( i \)'s school enrollment, \( P_{0i} = X_i \beta \) is child \( i \)'s enrollment probability without the transfer, and \( \lambda \) is the Lagrange multiplier on the budget constraint. The conditional transfer formula in equation (9) is thus a simple linear combination of a few observed characteristics \( Z \). It is similar to the scoring system used in many welfare programs, whereby characteristics \( Z \) command scores \( \alpha \) that add up to an aggregate score \( Z \alpha \). In this case \( Z \alpha \) determines not only eligibility but also the conditional transfer amount.

Important empirical questions are whether the use of this simple scoring schedule is close enough to the optimal conditional cash transfer schedule and what type I (exclusion) and type II (inclusion) errors are made in this implementation. These questions are addressed after the schedules are established.

III. Predicting Enrollment

In this section a predictive model of entry into secondary school is built. Although a probit and a logit perform better at high and low probabilities, a linear model is used here to avoid imposing heterogeneity on the impact of the conditional transfer through the functional form, since this will be an important determinant of the targeting rule.\(^5\) The sample of children finishing primary school and eligible for a Progresa transfer (defined as poor according to

\(^5\) In the simulation exercises that follow, the problem of predicted negative probability is never encountered (the majority of children have predicted probabilities above 0.4), but some predictions are above 1, even without conditional transfer and more when applying conditional transfers. For simulation purposes these will be set equal to 1.
the Progresa welfare index) in both the control and treatment communities is used, and randomization in the selection of communities ensures that being in a treated community is orthogonal to the children's characteristics. The average treatment effect can thus be obtained simply by comparing the average enrollment rates of children in the two types of communities. However, the actual amount of conditional transfer offered to a child is not orthogonal to its characteristics because being subject to the cap rule and the corresponding household scaling factor are both a function of the children's age structure, which is likely correlated with household preferences that influence schooling decisions. The impact of the continuous treatment effect is thus estimated, controlling for the conditional transfer level.

The empirical equivalent to equation (6) is written as:

\[ S_i = \delta_0 I_i + \delta I_i T_i + \beta_0 T_i + u_i \]

or

\[ S_i = \delta_0 I_i + \delta I_i T_i + \beta_0 T_i + X_i \beta + u_i, \]

with control variables \( X_i \)

where \( S_i \) is a binary variable indicating the enrollment status of child \( i \), \( I_i \) is a dummy variable that indicates whether child \( i \) lives in a treatment community, \( T_i \) is the conditional transfer that child \( i \) is eligible for under the program, and the control variables \( X_i \) are child, household, and community characteristics.

The program increases the probability that qualifying poor children will continue to secondary school by 13 percentage points [table 2, column (1)]. As expected this is slightly higher than the estimated 8–9 percentage point increase in enrollment probability conditional on completed primary school (i.e., including children who had dropped out of school before the program started) obtained in other studies (e.g., Schultz 2004).

The value of the conditional transfer (variable \( I_i T_i \)), which varies across children because of the cap on household transfer that affects 26 percent of the qualifying children, is used in column (2). Results show that the marginal effect of the conditional transfer is high (1.42 percentage points per $10). Note that the imposed linear form gives a meaningful positive effect only for conditional transfers above $100, which is not really restrictive because current conditional transfers are much higher. Adding a large number of child, household, and community controls in column (3) indicates that the main correlates of a child's secondary school enrollment are age of the child (negative), mother’s literacy and the household's maximum education level (positive), the number of agricultural workers and self-employed workers in the household (negative), total expenditure (positive), and distance to school (negative). State effects are also important. Both models predict that the current $200 conditional transfer increases the probability of enrollment by

---

6. The quality of the randomization is verified and documented in Behrman and Todd (1999).
Table 2. Linear Probability Model of Enrollment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment community</td>
<td>0.718</td>
<td>0.130*** (0.019)</td>
<td>-0.146 (0.171)</td>
<td>-0.172 (0.156)</td>
<td>-0.091 (0.162)</td>
<td>-0.159 (0.156)</td>
</tr>
<tr>
<td>Conditional transfer*treatment *(dollar variable)</td>
<td>1.215</td>
<td>0.142 (0.088)</td>
<td>0.156* (0.080)</td>
<td>0.061 (0.084)</td>
<td>0.095 (0.083)</td>
<td></td>
</tr>
<tr>
<td>Conditional transfer*treatment *male</td>
<td>0.601</td>
<td></td>
<td>0.003 (0.019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional transfer*treatment *(age - 12)</td>
<td>1.239</td>
<td></td>
<td>0.020*** (0.007)</td>
<td>0.016** (0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional transfer*treatment *father indigenous</td>
<td>0.419</td>
<td></td>
<td>0.037** (0.019)</td>
<td>0.028 (0.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional transfer*treatment *no secondary school in village</td>
<td>0.945</td>
<td></td>
<td>0.022 (0.022)</td>
<td>0.037* (0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child and household characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional transfer *(dollar per year)</td>
<td>1.940</td>
<td>-0.015 (0.069)</td>
<td>-0.072 (0.069)</td>
<td>0.006 (0.065)</td>
<td>-0.063 (0.069)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.507</td>
<td></td>
<td>0.057 (0.037)</td>
<td>0.073*** (0.029)</td>
<td>0.057 (0.037)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>13.012</td>
<td></td>
<td>-0.090*** (0.008)</td>
<td>-0.130*** (0.010)</td>
<td>-0.110*** (0.011)</td>
<td></td>
</tr>
<tr>
<td>Father is indigenous</td>
<td>0.354</td>
<td></td>
<td>0.027 (0.040)</td>
<td>0.059** (0.029)</td>
<td>-0.006 (0.045)</td>
<td></td>
</tr>
<tr>
<td>Birth order</td>
<td>2.014</td>
<td></td>
<td>0.016 (0.015)</td>
<td></td>
<td></td>
<td>0.014 (0.015)</td>
</tr>
<tr>
<td>Head is male</td>
<td>0.930</td>
<td></td>
<td>-0.037 (0.044)</td>
<td></td>
<td></td>
<td>-0.033 (0.044)</td>
</tr>
<tr>
<td>Has no father</td>
<td>0.114</td>
<td></td>
<td>-0.015 (0.045)</td>
<td></td>
<td></td>
<td>-0.011 (0.045)</td>
</tr>
<tr>
<td>Father is literate</td>
<td>0.670</td>
<td></td>
<td>0.054* (0.028)</td>
<td></td>
<td></td>
<td>0.053* (0.028)</td>
</tr>
<tr>
<td>Father's education</td>
<td>2.491</td>
<td></td>
<td>0.000 (0.006)</td>
<td></td>
<td></td>
<td>0.000 (0.006)</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Homogenous Impact</th>
<th>Heterogeneous Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has no mother</td>
<td>0.050</td>
<td>0.058 (0.075)</td>
<td>0.057 (0.074)</td>
</tr>
<tr>
<td>Mother is literate</td>
<td>0.621</td>
<td>0.055** (0.027)</td>
<td>0.057** (0.027)</td>
</tr>
<tr>
<td>Mother's education</td>
<td>2.351</td>
<td>−0.002 (0.006)</td>
<td>−0.003 (0.006)</td>
</tr>
<tr>
<td>Mother is indigenous</td>
<td>0.372</td>
<td>0.059 (0.039)</td>
<td>0.037 (0.039)</td>
</tr>
<tr>
<td>Mother's age</td>
<td>36.192</td>
<td>0.001 (0.001)</td>
<td>0.001 (0.001)</td>
</tr>
<tr>
<td>Number of children 0–10 years old</td>
<td>2.586</td>
<td>0.002 (0.006)</td>
<td>0.002 (0.006)</td>
</tr>
<tr>
<td>Number of children 11–19 years old</td>
<td>2.781</td>
<td>−0.014 (0.012)</td>
<td>−0.014 (0.012)</td>
</tr>
<tr>
<td>Number of agricultural workers</td>
<td>1.274</td>
<td>−0.031*** (0.009)</td>
<td>−0.030*** (0.009)</td>
</tr>
<tr>
<td>Number of nonagricultural workers</td>
<td>0.314</td>
<td>−0.02 (0.014)</td>
<td>−0.019 (0.014)</td>
</tr>
<tr>
<td>Number of self-employed</td>
<td>0.194</td>
<td>−0.039** (0.018)</td>
<td>−0.037** (0.018)</td>
</tr>
<tr>
<td>Number of unpaid family workers</td>
<td>0.332</td>
<td>−0.012 (0.010)</td>
<td>−0.013 (0.010)</td>
</tr>
<tr>
<td>Number of other working adults</td>
<td>0.100</td>
<td>−0.037 (0.028)</td>
<td>−0.035 (0.028)</td>
</tr>
<tr>
<td>Household's maximum education</td>
<td>4.975</td>
<td>0.018*** (0.004)</td>
<td>0.017*** (0.004)</td>
</tr>
<tr>
<td>Total expenditure ($100 per year)</td>
<td>8.055</td>
<td>0.004** (0.002)</td>
<td>0.004** (0.002)</td>
</tr>
<tr>
<td>Dwelling has dirt floor</td>
<td>0.696</td>
<td>0.047** (0.020)</td>
<td>0.043** (0.020)</td>
</tr>
<tr>
<td>People per room in dwelling</td>
<td>5.206</td>
<td>−0.003 (0.004)</td>
<td>−0.002 (0.004)</td>
</tr>
<tr>
<td>Dwelling has water</td>
<td>0.327</td>
<td>0.056*** (0.020)</td>
<td>0.058*** (0.020)</td>
</tr>
<tr>
<td>Rainfed land (hectares)</td>
<td>2.059</td>
<td>−3×10^{-4} (0.002)</td>
<td>8×10^{-5} (0.002)</td>
</tr>
<tr>
<td>Irrigated land (hectares)</td>
<td>0.066</td>
<td>−0.009 (0.015)</td>
<td>−0.007 (0.015)</td>
</tr>
<tr>
<td>Herd size</td>
<td>0.878</td>
<td>−0.005 (0.006)</td>
<td>−0.005 (0.006)</td>
</tr>
<tr>
<td>Community characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No secondary school in community</td>
<td>0.775</td>
<td>0.014 (0.045)</td>
<td>−0.221*** (0.033)</td>
</tr>
<tr>
<td>Distance to secondary school</td>
<td>1.031</td>
<td>−0.130*** (0.026)</td>
<td>−0.129*** (0.025)</td>
</tr>
<tr>
<td>(in kilometers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No school in community*girl</td>
<td>0.384</td>
<td>−0.028 (0.041)</td>
<td>−0.029 (0.041)</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guerrero</td>
<td>0.174</td>
<td>-0.124*** (0.044)</td>
<td></td>
<td></td>
<td>-0.124*** (0.044)</td>
<td></td>
</tr>
<tr>
<td>Michoacan</td>
<td>0.136</td>
<td>-0.168*** (0.045)</td>
<td></td>
<td></td>
<td>-0.166*** (0.045)</td>
<td></td>
</tr>
<tr>
<td>Puebla</td>
<td>0.159</td>
<td>-0.137*** (0.043)</td>
<td></td>
<td></td>
<td>-0.140*** (0.043)</td>
<td></td>
</tr>
<tr>
<td>Queretaro</td>
<td>0.050</td>
<td>-0.268*** (0.054)</td>
<td></td>
<td></td>
<td>-0.279*** (0.054)</td>
<td></td>
</tr>
<tr>
<td>San Luis Potosi</td>
<td>0.133</td>
<td>-0.163*** (0.045)</td>
<td></td>
<td></td>
<td>-0.161*** (0.045)</td>
<td></td>
</tr>
<tr>
<td>Veracruz</td>
<td>0.281</td>
<td>-0.103** (0.041)</td>
<td></td>
<td></td>
<td>-0.103** (0.041)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.636*** (0.015)</td>
<td>0.666*** (0.134)</td>
<td>2.009*** (0.204)</td>
<td>2.428*** (0.186)</td>
<td>2.301*** (0.231)</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>2,242</td>
<td>2,242</td>
<td>2,242</td>
<td>2,242</td>
<td>2,242</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.02</td>
<td>0.02</td>
<td>0.23</td>
<td>0.17</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** 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' analysis based on the 1998 Progresa survey.
the same 14 percentage points (table 3), which confirms that the controls are orthogonal to the treatment.  

Columns (4) and (5) explore heterogeneity of impact across categories of children, with and without controls, focusing on the aspects of heterogeneity that may be usable for targeting purposes. They are the child’s age, father’s ethnicity, and whether there is a secondary school in the community. Progresa recognizes gender differences, which are not found to be important in explaining differential impacts of transfers on the decision to continue into secondary school. The results show that age, ethnicity, and presence of a school in the community all have large impacts on enrollment, both directly as controls and in affecting the impact of the conditional transfer. The results in column (5) are used as the predictive model to evaluate the impact of targeting on enrollment.

Heterogeneity implies large differences in the impact of a conditional transfer on enrollment across categories of children (table 3). For a 12-year-old boy with a nonindigenous father and a school in the community, the $200 conditional transfer increases the probability of enrollment by only 3–4 percentage points. If he is two years behind normal progress, the conditional transfer increases the probability of enrollment by 10–12 percentage points. For a 12-year-old boy with an indigenous father or no secondary school in the community, the conditional transfer increases the probability of enrollment by 9–11 percentage points. Combining the features of being a boy, 14-years-old, with an indigenous father, and in a community with no secondary school means that a $200 conditional transfer raises the probability of school enrollment by 23–24 percentage points. These large differences suggest that there can be efficiency gains by using some of these dimensions of heterogeneity to target conditional transfers—in the same way that Progresa uses gender differences to set conditional transfer levels.

One concern is that identifying the impact of the size of a conditional transfer on enrollment derives from observing children who are offered less than the full transfer due to the cap on total household conditional transfers. These children are, by definition, from households with more eligible children. To verify that the enrollment model for these households does not differ significantly from the model for smaller households, the estimation was compared with a model

7. Another interesting result in column 3 is the relative magnitude of the impacts of a conditional transfer (Iₖ) compared with a nonconditional transfer (household total expenditure variable) on enrollment. While the result is only suggestive because total expenditure is endogenous, the $200 conditional transfer is associated with an increase in the probability of enrollment 17 times higher than an equal nonconditional transfer.

8. The age is centered on 12 years old, the median age for entry into secondary school, so that the coefficient on the direct variable is the impact on a 12-year-old.

9. The lack of significance and very low point estimate of the male*treatment interaction variable (0.003 with standard error 0.002) are robust to many specifications, including both fewer and more interaction terms and several control variables. The term is thus dropped from the estimation in column 5. The often reported difference between boys and girls comes from estimations of enrollment rather than continuation rates. Coady (2000) finds that most of that difference comes from Progresa’s very strong impact on girls’ re-entry into the school system during the program’s first year.
Table 3. Heterogeneity: Impact of Conditional Transfers on the Probability of School Enrollment by Type of Child

<table>
<thead>
<tr>
<th>Type of Child</th>
<th>Homogenous Impact</th>
<th>Heterogenous Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Community Dummy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variable (1)</td>
<td></td>
</tr>
<tr>
<td>Overall effect</td>
<td>0.130</td>
<td>0.035</td>
</tr>
<tr>
<td>Boy, 12 years old, with nonindigenous father, with secondary school in the</td>
<td>0.140</td>
<td>0.031</td>
</tr>
<tr>
<td>community ($200)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boy, 14 years old</td>
<td>0.115</td>
<td></td>
</tr>
<tr>
<td>Boy, with indigenous father</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td>Boy, with no secondary school in the community</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Boy, 14 years old, indigenous father, with no school in the community</td>
<td>0.243</td>
<td></td>
</tr>
</tbody>
</table>

Column numbers correspond to those in Table 2.

Source: Marginal effects based on results from table 2.
estimated for these children alone. The estimation is, as expected, more precise with the whole sample, but the parameters are neither individually nor globally significantly different between the two estimations (the p-values for the test of equality of the parameters on the conditional transfer variables are 0.49 without heterogeneity and 0.16 with heterogeneity), which confirms that the conditional transfer parameter was correctly identified. The orthogonality of the conditional transfer to all other variables was also verified by estimating different models for children in the treatment and control communities, and the parameters were checked to ensure that they are neither individually nor globally significantly different between the two estimations. Thus, the model can be used to predict behavior in the absence of a conditional cash transfer program.

IV. Comparing Conditional Cash Transfer Schedules

Three targeting and calibration schedules are now analyzed to determine whether they can help raise the efficiency of conditional transfers in inducing school enrollment (table 4). Each program has the same total budget as the current Progresa program. This budget is calculated by predicting for each sample child the expected uptake (predicted probability) \( EP \) and summing the expected transfers \( EP.T \) to reach a total annual outlay of $322,000 for the 2,242 sample children.\(^\text{10}\) The upper panel of table 4 reports the enrollment rates for all children and by category of children according to their 'risk level', that is, their predicted enrollment rates without any conditional transfer, or their eligibility status in the program. The lower panel reports some aggregate targeting and cost outcomes for the different schedules. These results are also represented in a graph of the enrollment probability of each program against the initial enrollment probability without a conditional transfer program (figure 2). The distance from the diagonal to each curve thus represents the gain in enrollment from the program with the corresponding schedule.

Emulating Progresa: A Universal Uniform Conditional Cash Transfer Program

The school enrollment rate without conditional transfer is 63.2 percent [table 4, column (1)]. Progresa's current universal conditional transfers with a cap and with differential values for boys and girls raise the probability of enrollment to 75.7 percent, a gain of 12.5 percentage points. The universal uniform conditional transfers program without a cap and without gender differences used as a benchmark for the subsequent simulations also raises the probability of enrollment to 75.7 percent [column (2)]. Under this program the conditional transfer

\(^{10}\) Another interesting exercise would be to define an efficient allocation of the total education budget of the current Progresa program. It would consist of reallocating the primary school budget to secondary school, thus doubling the budget for secondary school. A simulation of this budget reallocation shows that it would lead to almost universal secondary education with enrollment rates of 90.4–91.7 percent, depending on the rule used for transfer calibration.
<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Share of Total (percent)</th>
<th>Predicted enrollment rates (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No Program</td>
</tr>
<tr>
<td>All children</td>
<td>2,242</td>
<td>100.0</td>
<td>63.2</td>
</tr>
<tr>
<td>Probability of enrollment without conditional transfer (percent)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–40</td>
<td>354</td>
<td>15.8</td>
<td>27.8</td>
</tr>
<tr>
<td>40–60</td>
<td>583</td>
<td>26.0</td>
<td>50.9</td>
</tr>
<tr>
<td>60–70</td>
<td>376</td>
<td>16.8</td>
<td>64.9</td>
</tr>
<tr>
<td>70–80</td>
<td>392</td>
<td>17.5</td>
<td>74.6</td>
</tr>
<tr>
<td>80–100</td>
<td>537</td>
<td>24.0</td>
<td>90.5</td>
</tr>
<tr>
<td>Eligible students</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without conditional transfer</td>
<td></td>
<td></td>
<td>63.2</td>
</tr>
<tr>
<td>With conditional transfer</td>
<td></td>
<td></td>
<td>75.7</td>
</tr>
<tr>
<td>Noneligible students</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligibility (percent)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average annual transfer amount (US$)</td>
<td></td>
<td></td>
<td>193.6</td>
</tr>
<tr>
<td>Annual cost per additional child enrolled (US$)</td>
<td></td>
<td></td>
<td>1,151</td>
</tr>
<tr>
<td>Efficiency gain over universal uniform conditional transfer schedule (percent)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"Average over children who take the transfer.

Source: Authors' analysis based on the 2998 Progresa survey.
Figure 2. Impact of Alternative Conditional Cash Transfer Schedules on Enrollment Probabilities

Source: Authors’ calculations based on the 1998 Progresa survey.

per child is about $194 a year.11 Because many children receive a transfer even though they would attend school without one, the cost per additional child enrolled is $1,151 a year. Figure 2 shows that gains in enrollment due to the program are largest for children with a low probability of enrollment and decline as the enrollment probability rises.

Table 4 reports these gains, with enrollment probability rising from an average 27.8 to 47.2 percent, or 19.4 percentage points, for the children with probability of enrollment lower than 40 percent, while the gain is only 5.6 percentage points for those in the 80–100 percent category [columns (1) and (2)]. Gains are hence progressive in terms of the initial likelihood of going to school, even with a uniform conditional cash transfer program. This is the Progresa achievement that has been widely acclaimed in the literature. However, can better results be obtained by redefining the targeting and calibration of conditional transfers?

11. The level of the conditional transfer is determined by the Progresa budget, taking into account the uptake that it is predicted to induce.
An Optimal Variable Conditional Cash Transfer Program

The second simulation implements the optimal variable conditional transfer program established in the model under the same budget constraint and taking into account heterogeneity in probability of enrollment and responses to transfers across children. Both eligibility and the optimal conditional transfer value are simultaneously determined. This is done by offering the conditional transfer defined in equation (7) to children of characteristics X. To compute the conditional transfer values, the estimated values for \( \beta \), \( \delta_0 \), and \( \delta \) in table 2, column (5), are used, and the shadow value \( \lambda \) of the budget constraint that balances the budget is found by tâtonnement. The resulting conditional transfers vary from $100 to $350, depending on the child's characteristics. Under this optimal variable program the conditional transfers to children with a low probability of attending school are increased, and children with high probabilities of attending school are targeted less because efficiency leakages are particularly high for them.

The best predictor model for enrollment [table 2, column (5)] is again used to predict enrollment for every child in the sample. Students eligible to receive a conditional transfer have an average predicted enrollment rate of 78.9 percent, compared with 55.8 percent, had they not been offered the conditional transfer. The noneligible students have an average predicted enrollment rate of 89 percent. Overall, the predicted school enrollment rate is now 81.1 percent, an efficiency gain of 43.6 percent over the universal uniform conditional transfer schedule. The optimal variable conditional transfer schedule almost equalizes the probability of enrollment among children with very different initial probabilities to values close to 70 percent (figure 2). The largest gains in probability of enrollment are thus captured by those with the lowest initial probabilities.

Figure 3 shows the distribution of children by initial enrollment probability without a conditional cash transfer program, superimposing the distribution of those that are eligible in the optimal schedule (shaded) and showing, by difference, the distribution of noneligible (unshaded). Under the optimal conditional transfer schedule eligibility is concentrated on children with low initial probabilities, while noneligible children all have initial probabilities above 70 percent. The optimal calibration of conditional transfers also favors those with low initial probabilities, trying to induce them to enroll in school with a higher conditional transfer. The conditional transfers decline as the probability of being enrolled in school without a transfer rises. There are, however, relatively few children with predicted low enrollment probability; most are in the 40–80 percent range.

Some 77.5 percent of children are eligible for a conditional transfer under the optimal variable schedule, with the average transfer about $237, a 22 percent increase over the average transfer under the universal uniform schedule (table 4). The optimal schedule thus suggests raising the amount of the conditional transfer for children who are less likely to attend school while reducing it for children who are already likely to attend school without a conditional transfer. Since there are
still efficiency leakages among eligible children, the cost per additional child enrolled is $802, down from $1,151 under the universal uniform conditional transfer schedule. Cost saving per additional child enrolled is around 30 percent.

An Implementable Conditional Cash Transfer Program

Having established the optimal conditional transfer schedule as an efficiency benchmark, the analysis now turns to the definition of simpler implementable conditional transfer schedules, based on a linear combination of a few observable characteristics. For a given set of variables $Z$ the implementable schedule is the solution to the optimization problem defined in equations (9) and (10). The values for $\beta$, $\delta_0$, and $\delta$ are taken from column (5) of table 2, with the parameters $\alpha$ and $\lambda$ solved for iteratively, followed by exploration of combinations of characteristics $Z$ that are easily observable, verifiable by others, and nonmanipulatable by the household. An efficiency criterion for selection requires choosing characteristics that are important correlates of enrollment (to target the children least likely to enroll without a conditional transfer) or that indicate high sensitivity of enrollment to a conditional transfer. In addition to these

12. Starting with general eligibility, the optimization problem in equation 10 is solved for $\alpha$ as a function of $1$, with $1$ adjusted to balance the budget. These parameters are used to compute transfers and define eligibility. This procedure is iterated until there is convergence, that is, no change in eligibility between two consecutive iterations. This is always achieved in fewer than five iterations.
features, actual implementation of a program requires these criteria to be legally and politically acceptable. This is clearly an issue that every program should address in its own particular context.

In the base model the conditional transfer schedule depends only on gender and birth order of the child, presence of a secondary school in the community, distance to a secondary school if there is not one in the community, and state dummy variables, which are all strong correlates of enrollment. A few alternative specifications are reported later. Note that age of the child is not used because an eligibility criterion based on age could induce perverse behavior, such as parents delaying their children's entry to secondary school to benefit from a larger conditional transfer. The child's birth order, which cannot be manipulated, turns out to capture part of this information. Each variable can be easily observed and verified. In fact, instead of secret eligibility formulas (as currently used for poverty), which offer little room for recourse and accountability, self-registration is possible, with easy verification.

Results are reported in table 5. The birth order parameter indicates that the conditional transfer is highest for the oldest child and decreases by $12 for each younger sibling. Girls would optimally receive a premium of $25. The main source of variation in conditional transfer is related to distance to school, with a large premium for children who need to travel some distance to school and an additional amount for each kilometer traveled. The program also exhibits some variation across states, with a difference of $87 between the extreme cases of Queretaro and Guerrero.

Examples of eligibility and conditional transfer amounts computed with this simple points system are reported in the lower part of table 5. Children with a school in their own community (23 percent of the sample) are not eligible. Their predicted enrollment rate without conditional transfer is 80.5 percent, which is also the rate observed in control communities with a school. By contrast, all the children who do not have a school in their community are eligible for a conditional transfer. A boy who is the oldest child and lives 3 kilometers from a school would receive a conditional transfer of $213, whereas the third child would receive only $190. If the oldest child is a girl, she would receive $239. Cumulating all the disadvantages, a girl who is the oldest child and lives 6 kilometers from a school would receive the highest transfer: $266.

This conditional transfer schedule results in an efficiency loss relative to the optimal variable schedule, the cost to be paid for simplicity, and transparency (table 4). Although the number of eligible children is about the same as with the optimal conditional transfer schedule (77.4 percent), the eligibility criteria are not the same. The implementable schedule includes 9 percent of the children not

---

13. The average distance to school for the 77 percent of children who do not have a school in their community is 3.1 kilometers. Enrollment rates decrease very sharply with distance to school in the control communities, reaching a low of 43 percent for the 19 percent of children who live more than 4 kilometers from school.
TABLE 5. Optimal Implementable Conditional Cash Transfer Schedule

<table>
<thead>
<tr>
<th>Conditional transfer formula</th>
<th>Base Model</th>
<th>With Illiteracy</th>
<th>Geographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>(US$ per year)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth order</td>
<td>-12</td>
<td>-12</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-25</td>
<td>-25</td>
<td></td>
</tr>
<tr>
<td>No secondary school</td>
<td>-476</td>
<td>-502</td>
<td>447</td>
</tr>
<tr>
<td>in the community</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to secondary</td>
<td>50</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>school $[\ln(1+kms)]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother illiterate</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father illiterate</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guerrero</td>
<td>-295</td>
<td>-351</td>
<td>-295</td>
</tr>
<tr>
<td>Hidalgo</td>
<td>-278</td>
<td>-327</td>
<td>-283</td>
</tr>
<tr>
<td>Michoacan</td>
<td>-246</td>
<td>-288</td>
<td>-253</td>
</tr>
<tr>
<td>Puebla</td>
<td>-267</td>
<td>-317</td>
<td>-273</td>
</tr>
<tr>
<td>Queretaro</td>
<td>-208</td>
<td>-257</td>
<td>-214</td>
</tr>
<tr>
<td>San Luis Potosí</td>
<td>-260</td>
<td>-303</td>
<td>-260</td>
</tr>
<tr>
<td>Veracruz</td>
<td>-285</td>
<td>-333</td>
<td>-288</td>
</tr>
</tbody>
</table>

Examples of transfers by children types in state of Guerrero (US$ per year)

- Oldest, male, with literate parents, and school 3 kilometers away: Not eligible Not eligible Not eligible
- Oldest, male, with illiterate parents, and school 3 kilometers away: 213 184 218
- Third child, male, with literate parents, and school 3 kilometers away: 190 162 218
- Oldest, female, with illiterate parents, and school 3 kilometers away: 239 265 218
- Oldest, female, with illiterate parents, and school 6 kilometers away: 266 292 245

<table>
<thead>
<tr>
<th>Efficiency gain over universal uniform schedule (percent)</th>
<th>Base Model</th>
<th>With Illiteracy</th>
<th>Geographic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29.4</td>
<td>31.0</td>
<td>28.5</td>
</tr>
</tbody>
</table>

eligible under the optimal schedule (type II error) and excludes 9 percent of the children eligible under the optimal schedule (type I error). Enrollment of eligible children rises from 58.2 percent without a conditional transfer to 79.1 percent with one. Enrollment is 80.5 percent for noneligible children and 79.4 percent for all children. This implies a 29.4 percent efficiency gain over the universal uniform conditional transfer schedule. Cost per additional child enrolled is $889, still about 23 percent cheaper than under the universal uniform schedule but nearly 11 percent more expensive than under the optimal variable schedule.

Alternative implementable schedules that vary the characteristics used to establish eligibility and conditional transfer amounts are now explored.
Adding mother's and father's literacy status, important predictors of school enrollment, raises the efficiency gain to 31 percent above the universal uniform schedule. While it can be argued that such subsidies (here computed as $26 if the mother is illiterate and $30 if the father is illiterate) may give the wrong signal and bias, the return to education, they can also be seen as a way to compensate for the handicap that children of uneducated parents have and to help children catch up.

At the other extreme, how efficient would it be to define a conditional transfer schedule at the community level (although only for poor children)? Doing so leads to an important efficiency gain of 28.5 percent over the uniform conditional transfer schedule. This geographical targeting is interesting because it shows that in the particular case of rural Mexico an important efficiency gain could be obtained by redesigning the conditional transfer program as a school transportation subsidy. A simple transportation subsidy would capture 65 percent of the efficiency gain that the optimal conditional transfer program would garner. The question then arises of how this intervention compares with a supply-side policy that would bring schools closer to where people live. This is beyond the scope of this article, but Coady (2000) estimates that the cost of raising enrollment through a supply-side intervention that increases the number of rural schools would be more than seven times as much as the current Progresa program.

These specific implementable schedules are illustrations of the idea that designing a relatively simple conditional transfer program, with a points system that is transparent and easily verifiable, is indeed feasible and could ensure large efficiency gains.

Comparing Direct Costs and Efficiency Leakages under the Three Schedules

A key determinant of the relative efficiency of different targeting schedules is their efficiency leakages, namely the magnitude of the transfers that go to children who would attend school without the conditional transfer. This is analyzed in a comparative fashion in figures 4 and 5, where the total transfer cost for each category of children is divided into direct costs (transfers to children that would not otherwise have enrolled, represented in black) and efficiency-leakage costs (transfers to children that would have enrolled anyway, represented with stripes). Differences among the figures are quite telling. Under the universal uniform conditional transfer schedule, leakages are particularly high, especially among children with a high probability of attending school without a conditional transfer. Some 83.2 percent of the total cost is efficiency leakages, leaving an effective direct cost of only 16.8 percent. The optimal variable conditional transfer schedule reduces efficiency leakages by targeting children with a low probability of attending school and increasing the magnitude of the conditional cash transfers (figure 5). Efficiency leakages are reduced to 64.9 percent of total costs, implying an effective direct cost of 36.1 percent. The implementable conditional transfer schedule (not pictured) has an efficiency leakage of 72.5 percent of total costs. Because targeting is simplified and
transparent, it is a compromise between the universal and the optimal variable conditional transfer schedules. Its effective direct cost is 27.5 percent.

In conclusion, the optimal variable conditional transfer schedule could offer a significant efficiency gain in school enrollment and could be implemented through a secret formula like the one Progresa currently uses to target poverty. It may, however, be too complex to administer, and secrecy is not a desirable feature as it makes recourse almost impossible. However, the implementable variable conditional transfer schedule— with its transparent targeting— also results in substantial efficiency gains relative to Progresa's current universal uniform conditional transfers.

V. EFFICIENT CONDITIONAL CASH TRANSFER PROGRAMS AND EQUITY

Are these optimal and implementable schedules regressive or progressive? In other words, are efficiency gains in enrollment achieved at an equity cost? Conditional transfers driven by efficiency gains indeed raise the issue that maximally efficient programs may be inequitable (Das, Do, and Özler 2005). For this reason eligibility is restricted to poor households. However, when poor households are further targeted, are the resulting transfers regressive or progressive?

Before looking at the distributive effect of targeting among poor households, it is interesting to note that the Progresa transfers themselves were not...
particularly efficient in reducing poverty or inequality. Indeed, with poverty measured by consumption per capita, the transfers are almost uniformly distributed across levels of per capita consumption (de Janvry and Sadoulet 2003). This article, however, discusses the issue of tradeoff between efficiency and equity using the Progresa welfare index measured in 1997, rather than the income-consumption level, since this is what Progresa uses as a poverty indicator. Figure 6 shows the average distributed transfer with households ranked by the Progresa welfare index. The average transfer distributed by Progresa shows a clear upward trend and thus regressivity among poor households. This is because of the low uptake rate in low welfare classes. By contrast, the average distributed transfer decreases across welfare levels with the optimal variable conditional transfer schedule (from $160 to $140) and is uniform with the implementable schedule. Efficiency gains in implementing conditional transfer programs designed to maximize the effect of the conditionality are thus not achieved at the cost of rising inequality among poor households.

VI. Conclusions

This article questions whether efficiency gains can be achieved in conditional cash transfer programs by improving targeting among poor households and
better calibrating conditional transfers. The efficiency objective is to maximize the impact of the condition imposed on the transfer, in this case gains in school enrollment among poor children. Using the data from the Progresa randomized experiment, the analysis focused on the crucial education decision for children in poor Mexican rural communities—namely whether to continue schooling at the secondary level.

Achieving efficiency gains by targeting and calibrating conditional transfers requires focusing on children who have a high probability of not enrolling in school without a conditional transfer and who have a high response to the amount offered, within the overall program budget constraint. Implementing this program requires predicting school enrollment as a function of the conditional transfer offered and of child, household, and community characteristics. Heterogeneity in responses shows that age, ethnicity, and presence of a school in the community lead to large differences in enrollment. Three alternative targeting and calibration schedules were then compared: the current Progresa schedule of universal uniform conditional transfers; an optimal schedule of variable conditional transfers; and a...
schedule of implementable conditional transfers, where the criteria used for targeting and calibration are easily observable, verifiable by others, and nonmanipulatable by the household. In setting up a new program, a pilot experiment would need to be used to estimate the enrollment probability model necessary to establish the targeting and calibration formulas.

The optimal schedule is found to offer a 44 percent efficiency gain over the universal conditional transfer schedule and the implementable schedule a 29 percent gain. The optimal schedule reduces efficiency leakages (receipt of transfers by children who would attend school without a conditional transfer) from 83 percent of total costs to 65 percent, and the implementable schedule to 73 percent. These efficiency gains are not achieved at the cost of rising inequality among poor households.

The overall conclusion is thus that large efficiency gains can be achieved in implementing what are in many countries very expensive conditional cash transfer programs for human capital formation among poor households if rules for targeting and calibration of conditional transfers are designed to maximize the effect of the conditionality.

References


Child Labor and School Achievement in Latin America

Victoria Gunnarsson, Peter F. Orazem, and Mario A. Sánchez

Child labor's effect on academic achievement is estimated using unique data on third and fourth graders in nine Latin-American countries. Cross-country variation in truancy regulations provides an exogenous shift in the ages of children normally in these grades, providing exogenous variation in the opportunity cost of children's time. Least squares estimates suggest that child labor lowers test scores, but those estimates are biased toward zero. Corrected estimates are still negative and statistically significant. Children working 1 standard deviation above the mean have average scores that are 16 percent lower on mathematics examinations and 11 percent lower on language examinations, consistent with the estimates of the adverse impact of child labor on returns to schooling.

About one of eight children in the world is engaged in market work. Despite general acceptance that child labor is harmful and despite international accords aimed at its eradication, progress on lowering the incidence of child labor has been slow. Although often associated with poverty, child labor has persisted in some countries that have experienced substantial improvements in living standards. For example, Latin America, with several countries in the middle- or upper-middle-income categories, still has child labor participation rates that are similar to the world average.

Countries have adopted various policies to combat child labor. Most have opted for legal prohibitions, but these are only as effective as the enforcement. As many child labor relationships are in informal settings within family enterprises, enforcement is often difficult. Several countries, particularly in Latin America, have initiated programs that offer households an income transfer in exchange for keeping children in school and out of the labor market.

Victoria Gunnarsson is a research officer in the Fiscal Affairs Department at the International Monetary Fund; her email address is vgunnarsson@imf.org. Peter F. Orazem is a professor at Iowa State University; his email address is pfo@iastate.edu. Mario A. Sánchez is a social development specialist at the Inter-American Development Bank; his e-mail address is mariosan@iadb.org. The authors thank Wallace Huffman, Robert Mazur, three referees, the editor, and seminar participants at Iowa State and Minnesota for numerous comments and suggestions.
Presumably, governments invest resources to lower child time in the labor market in anticipation that the child will devote more time to the acquisition of human capital. The government's return will come from higher average earnings and reduced outlays for poverty alleviation when the child matures. However, despite a huge acceleration in the research on child labor, there is surprisingly little evidence that relates child labor to schooling outcomes in developing countries. Most children who work are also in school, suggesting that child labor does not lower school attainment. Additionally, studies that examine the impact of child labor on test scores have often found negligible effects, although most of these are in industrial country contexts. More recently, Heady (2003) and Rosati and Rossi (2003) have found some evidence that child labor lowers primary school test scores in developing countries.

This article builds on these last two studies by examining the link between child labor and school achievement in nine countries in Latin America. This article benefits from more detailed data sets that allow controls for child, household, school, and community variables, and it uses an empirical strategy that controls for the likely endogeneity of child labor. The results are consistent: in all nine countries, child labor lowers performance on the tests of language and mathematics proficiency, even when controlling for school and household attributes and for the joint causality between child labor and school outcomes. To the extent that lower cognitive attainment translates into lower future earnings, as argued by Glewwe (2002), these results suggest that there is a payoff in the form of higher future earnings from investing in lowering the incidence of child labor.

I. Literature Review

Most studies that analyze the relationship between time at work and school attainment have focused on high-school or college students in industrial countries. These studies have generally found little evidence that part-time work combined with schooling affects school achievement. When adverse effects are found, they are apparent only at relatively high work hours. Important exceptions include recent studies by Tyler (2003) and Stinebrickner and Stinebrickner (2003) that found that after controlling for the likely endogeneity of child labor, working while in school led to much larger implied declines in high-school math scores and in college grade point averages than had been found previously. Post and Pong (2000) also found a negative association between

1. Two excellent recent reviews of the recent literature are by Basu and Tzannatos (2003) and Edmonds and Pavcnik (2005).
work and test scores in samples of eight graders in many of the 23 countries they studied.³

There are several reasons why the experience of older working students may differ from that of young children working in developing countries. Young children may be less physically able to combine work with school, so that young working children may be too tired to learn efficiently in school or to study afterwards. Children who are tired are also more prone to illness or injury that can retard academic development. It is possible that working at a young age disrupts the attainment of basic skills more than it disrupts the acquisition of applied skills for older students. School and work, which may be complementary activities once a student has mastered literacy and numeracy, may not be compatible before those basic skills are mastered.

Past research on the consequences of child labor on schooling in developing countries has concentrated on the impact of child labor on school enrollment or attendance. Here the evidence is mixed. Patrinos and Psacharopoulos (1997) and Ravallion and Wodon (2000) found that child labor and school enrollment were not mutually exclusive and could even be complementary activities. However, Rosenzweig and Evenson (1977) and Levy (1985) found evidence that better-developed child labor markets lowered school enrollment. There is stronger evidence that child labor lowers time spent in human capital production, even if it does not lower enrollment. Psacharopoulos (1997) and Sedlacek and others (forthcoming) reported that child labor lowered years of school completed, and Akabayashi and Psacharopoulos (1999) discovered that child labor lowered study time.

Nevertheless, school enrollment and attendance are not ideal measures of the potential harm of child labor on learning because they are merely indicators of the time input into schooling and not the learning outcomes. Even if child labor lowers time in school, it may not hinder human capital production if children can use their limited time in school efficiently. This is particularly so if schools are of such poor quality that not much learning occurs in any case. By contrast, the common finding that most working children are enrolled in school may miss the adverse consequences of child labor on learning if child labor is not complementary to the learning process at the lower grades.

A more accurate assessment of the impact of child labor on human capital production requires the measures of learning outcomes, such as test scores, rather than education inputs, such as time in school, to determine whether child labor limits or enhances human capital production. Moreover, evidence suggests that cognitive skills, rather than years of schooling, are the fundamental determinants of adult wages in developing countries (Glewwe 1996, Moll 1998).

³ The study included several developing countries, including Colombia, Iran, South Africa, Thailand, and the Philippines, which had the largest estimated negative effects of child labor on school achievement. However, the estimates do not control for school attributes or possible joint causality between school achievement and child labor.
Therefore, identifying the impact of child labor on school achievement will yield more direct implications for child labor's longer-term impacts on earnings and poverty status later in the child's life.

Direct evidence of the impact of child labor on primary school achievement is rare. Heady (2003) found that child work had little effect on school attendance but a substantial effect on learning achievement in reading and mathematics in Ghana. Rosati and Rossi (2003) reported that in Nicaragua and Pakistan, more hours of child labor are associated with poorer test scores. Both of these studies have weaknesses related to data limitations. Heady treated child labor as exogenous, but it is plausible that parents send their children to work in part because of poor academic performance. Rosati and Rossi had no information on teacher or school characteristics, although these are likely to be correlated with the strength of local child labor markets.

This study makes several important contributions to knowledge of the impact of child labor on schooling outcomes in developing countries. It shows how child labor affects test scores in nine developing countries, greatly expanding the scope of existing research. Because the same examination was given in all countries, the study can illustrate how the effect of child labor on cognitive achievement varies across countries that differ greatly in child labor incidence, per capita income, and school quality. Because the countries also differ in the regulation and enforcement of child labor laws, cross-country variation in schooling ages and truancy laws can provide plausible instruments for endogenous child labor. Finally, because the data set includes a wealth of information on parent, family, community, and school attributes, the impact of child labor on schooling outcomes can be estimated while holding fixed other inputs commonly assumed to explain variation in schooling outcomes across children.

The results are consistent. Child labor lowers student achievement in every country. The conclusions are robust to alternative estimation procedures and specifications. The inescapable conclusion is that child labor has a significant opportunity cost in the form of forgone human capital production, a cost that may not be apparent when looking only at enrollment rates for working children.

II. Empirical Model

Ben Porath (1967) laid out the classic model of human capital investments over the life cycle. There are diminishing marginal returns to time in school because of concavity in the human capital production process and because the opportunity cost of allocating time to further skill acquisition increases as skills are accumulated. In addition, finite life spans limit the length of time to capture returns from schooling as age increases, further decreasing the marginal returns to time in school as age rises. All of these factors suggest that time invested in human capital production will decrease as an individual ages. However, early in life,
children may specialize in schooling if the present value of the return is sufficiently high relative to its current marginal cost.\footnote{The main predictions are not altered if leisure is added to the model. It will still be optimal to invest more intensively in human capital early in life and to decrease investment intensity with age. In addition, because the cost of leisure is the value of work time, individuals will consume the least leisure when wages are highest. In the application here, children will consume less leisure as they age, and so older children will still be expected to work more than younger children. \textit{Heckman} (1976) presented a detailed model of human capital investment, leisure demand, and consumption over the life cycle. \textit{Huffman} and \textit{Orazem} (2006) present a much-simplified model that generates the predictions discussed in the text.}

Of interest here is the tradeoff parents face in deciding whether a child should specialize in schooling or should divide time between school and work. By age $t$, the child has completed $E_t$ years of schooling. In addition, the child has matured for $t$ years. The opportunity cost of a child's school time is assumed to rise with $E_t$ and $t$ and is also a function of local labor market conditions $Z_t$. The returns to time in school will depend on how much the child is expected to learn, $Q_t$. A vector of observable parent, home, school, and community variables, $H_t$, may affect tastes for child labor as well as the productivity of child time in school through $Q_t$. The child's labor supply function will be of the form

$$C_t = c(E_t, t, Z_t, Q_t, H_t, \varepsilon_t)$$

where $\varepsilon_t$ is a random error.

The human capital production process is assumed to depend on past human capital accumulations, current factors that would make the child's time in school more productive, and the time spent in school. Letting $Q_t$ be an observable measure of cognitive skills produced in school, the human capital production process will be of the form

$$Q_t = q(E_t, t, C_t, H_t, \eta_t)$$

where $\eta_t$ is a component of cognitive ability that the parents can observe but not the econometrician.

Because the decision on whether or how much the child works is based in part on the parents' knowledge of $\eta_t$, and because student outcomes are influenced by child labor, $\text{Var}(\varepsilon_t, \eta_t) \neq 0$, and ordinary least squares estimation of equation (2) will be biased. Short of a randomized experiment that assigns children into working and non-working groups, the best candidate to resolve the problem will be to find variables that shift the probability that a child works but do not directly affect child learning in school. Needed are variables that alter the local labor market for child labor, $Z_t$, to provide exogenous shifts in the child labor equation in estimating equation (2).

Factors Shifting the Probability of Child Labor

Elements of the vector $Z_t$ are required that alter the local labor market for children but do not affect test scores. Because the probability of working rises
with age, factors that alter the age at which a child would normally be in a given grade will also affect the probability that the child will be working. In Latin America, the age at which children are expected to start school varies across countries from 5 to 7 years of age. The age at which a child may legally leave school also varies from 12 to 16 years of age. As a consequence, children must attend school as few as 5 years in Honduras to as many as 10 years in Peru.

These differences in laws regulating school attendance and child labor alter the age at which children would normally enter grades 3 and 4 and thus the opportunity costs of being in those grades. Children starting school earlier will be younger at grade 3 and more likely to attend school full time without working. Third and fourth graders in countries with the lowest working ages are more likely to appear legal, even if they are under 12 years of age. Therefore, children in countries with low truancy ages will be more likely to be working while attending school.

An alternative measure of the opportunity cost of attending school would be the local market wage for children. Because most child labor is unpaid work for family enterprises, however, market wages would not adequately capture the value of time outside of school even if such information were available. In their place is used the presumed upward relationship between the marginal productivity of child labor and the child's age, assumed to be driven largely by physical stature. Interactions between measures of a country's school starting age or truancy age and a child's age are used to capture exogenous variation across countries in the probability that third and fourth graders work. These shifts in the net return to time in school provide the needed exogenous shift in $C$.

Within countries the largest source of variation in demand for child labor occurs across rural and urban areas. There are more uses for child labor in rural markets, and so labor force participation rates are higher for rural children than for urban children in all the countries in this study. That source of variation is captured with interactions between child age and a dummy variable indicating rural residence for boys and girls.

How these elements of $Z_t$ affect the probability of engaging in child labor is illustrated in figures 1–3.

Factors Affecting School Outcomes

Estimation of equation (2) follows the educational production function literature in that $Q$ is measured by test scores that are explained by variables characterizing the student's parents, household, teacher, school, and community.

5. Rosenzweig (1980) found that in a sample of adults, wages for day labor in India were primarily driven by stature and not by acquired education. Wage patterns reported by Ray (2000) for boys and girls in Pakistan and Peru suggest rising opportunity costs of child time as age increases.

6. Angrist and Krueger (1991) used variation in compulsory school starting ages across states to instrument for endogenous time in school in their analysis of returns to schooling using U.S. Census data. Tyler (2003) used variation in state child labor laws to instrument for child labor in his study of U.S. high-school test scores. This study began with a large number of interactions, but the resulting variables were highly collinear, and so a parsimonious subset of the fuller specification was used.
Estimates of educational production functions are subject to numerous biases. Among the most commonly discussed is the lack of adequate control for the student's innate ability. Many studies have attempted to correct for the problem by using two test scores taken at different times. If ability has an additive effect on school achievement, the difference between the two output measures will be purged of the ability effect. The data for the current study include only tests taken at one point in time, so the differencing option is not available. However, there are reasons why undifferenced data may yield satisfactory or even preferred estimates to the differenced data. As Glewwe (2002) argued, if measures of \( H_t \) vary slowly over time, the value of the differenced measure of achievement is minimal. This is more likely to be true at the earliest stages of schooling, when there is less variation in curriculum, educational materials, or teacher training. Furthermore, the use of parental attributes such

7. See Glewwe (2002) for a comprehensive review of the problems associated with estimating educational production functions.
8. Ability bias has also been the subject of numerous studies estimating returns to schooling. The consensus is that the bias is small (Card 1999). If earnings and cognitive skills are closely tied, as argued by Glewwe (2002), the role of ability bias should be small in educational production estimates also.
as education and income should partially control for inherited ability. Finally, if there is considerable measurement error in estimates of $Q_\alpha$, the level of $Q_\alpha$ may be measured more reliably than the change in $Q_\alpha$. In any event, the results of the production function estimation in this study should be interpreted as cumulative as of grade 3 or 4 rather than the additional learning obtained in that grade.

III. DATA

In 1997 the Latin-American Laboratory of Quality of Education (LLECE) carried out the First Comparative International Study on Language, Mathematics, and Associated Factors for third and fourth graders in Latin America. LLECE initially collected data in 13 countries, but the required information for the regression analysis for this study was available only for nine countries: Argentina, Bolivia, Brazil, Chile, Colombia, Dominican Republic, Honduras, Paraguay, and Peru.  

The data set is composed of a stratified sample designed to ensure sufficient observations of public, private, rural, urban, and metropolitan students in each

---

9. Costa Rica was included in the initial data collection, but LLECE dropped those data because of consistency problems. Cuba was excluded because of missing data on child labor. Mexico and Venezuela lacked required information on child age.
country. Data were collected on 40 children from each of 100 schools in each country for a total of 4,000 observations per country. Half of the students were in the third grade and half in the fourth grade. For budgetary reasons LLECE had to use a priori geographic exclusions to limit the transportation and time costs of data collection. Very small schools with too few third and fourth graders and schools in remote, difficult to access, or sparsely inhabited regions were excluded. Because of the cost of translating examinations, schools with bilingual or indigenous language instruction were also excluded. As the excluded schools would cater to relatively more disadvantaged populations, our results should be viewed as applying to school populations that are less rural, from more majority ethnic groups, and somewhat more advantaged than average for all Latin-American children.

**Test Scores**

Survey instruments consisted of tests administered to the sample of children of the sampled schools, and self-applied questionnaires to school principals, teachers, parents (or legal guardians) of the tested children, and the children. In addition, surveyors collected information on the socioeconomic characteristics of the

---

10. For a detailed description of the a priori exclusions in each country, see Table III.6 of the Technical Bulletin of the LLECE.
community. A description of the variables used in the analysis is provided in appendix table A-1, and summary statistics are reported in appendix table A-2.\(^{11}\)

All children were tested in mathematics, and all were tested in Spanish except the Brazilian children who were tested in Portuguese. The tests and questionnaires were given only to children who attend school, so no information was obtained on children who are not in school. Therefore, the results can be applied to enrolled children only. If working children who perform most poorly in school drop out to work full time, the estimate of the consequences of child labor on schooling outcomes may miss some of those most harmed by child labor while including children who can work and still perform well in school. However, 95 percent of children aged 9–11 are enrolled in Latin America, so the bias is likely to be modest.\(^{12}\) In settings where primary enrollment rates are much lower, the bias could be substantial, however.

\textit{Child Labor}

Child labor is measured by children's responses to a question asking whether they are engaged in work outside the home.\(^{13}\) The concentration on paid work outside the home avoids some definitional problems related to distinguishing unpaid work for home enterprise from household chores. However, it is also apparent in the application that child labor in the home does not have the same apparent negative consequences on student achievement as does work outside the home.

A comparison of the intensity of child labor participation rates in nine countries for children who report that they work inside or outside the home and average language and mathematics test scores shows an unvarying pattern (table 1).\(^{14}\) Children who work only some of the time outperform those who work often. Children who almost never work outperform those who work sometimes or often. The differences are almost always statistically significant. The advantage is large for children who almost never work over those who often work, averaging 22 percent on the mathematics examination and 27 percent on the language

11. For some reason, language scores were reported for 2 percent fewer students than were mathematics scores. The missing scores appear to be due to random reporting errors, as there were no large differences between the sample means of the group taking the mathematics and language tests. The means are reported from the sample taking the mathematics examination.

12. Sedlacek and others (2005) presented data on enrollment by age for 18 Latin-American countries. Even for the poorest quintile of children, enrollment rates are more than 90 percent for children aged 9–11.

13. As pointed out by a referee, it would be better to have information on hours of work rather than these more-vague measures of work intensity. The instrumental variables procedure described later is an attempt to correct for biases because of measurement error in child labor.

14. The averages are reported for the subset of countries for which data were available on both language and mathematics test scores and for which responses could be matched for working inside and outside the home. Only partial information was available for Mexico and Venezuela, but the pattern of average test scores for children working outside the home in Mexico and Venezuela was the same—children working more outside the home had significantly lower average test scores. Data limitations prevented generating the corresponding average test scores for children working inside the home for those two countries.
<table>
<thead>
<tr>
<th>Country</th>
<th>Language Test Scores (Maximum Score = 19)</th>
<th>Mathematics Test Scores (Maximum Score = 32)</th>
<th>Language Test Scores (Maximum Score = 19)</th>
<th>Mathematics Test Scores (Maximum Score = 32)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Working Outside the Home</td>
<td>Working Inside the Home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>12.0</td>
<td>15.7</td>
<td>13.9</td>
<td>17.9</td>
</tr>
<tr>
<td>Sometimes</td>
<td>13.0 (8.3)**</td>
<td>17.2 (9.6)**</td>
<td>14.3 (2.9)*</td>
<td>18.6 (3.9)*</td>
</tr>
<tr>
<td>Almost never</td>
<td>14.3 (19.2)**</td>
<td>18.4 (17.2)*</td>
<td>14.7 (5.8)*</td>
<td>19.9 (11.2)'</td>
</tr>
<tr>
<td>Bolivia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>9.7</td>
<td>14.2</td>
<td>11.2</td>
<td>15.9</td>
</tr>
<tr>
<td>Sometimes</td>
<td>10.1 (4.1)*</td>
<td>14.7 (3.5)</td>
<td>11.2 (0.0)</td>
<td>16.0 (0.6)</td>
</tr>
<tr>
<td>Almost never</td>
<td>11.3 (16.5)'</td>
<td>15.1 (6.3)**</td>
<td>11.8 (5.4)</td>
<td>17.2 (8.2)'</td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>11.2</td>
<td>14.4</td>
<td>13.0</td>
<td>16.9</td>
</tr>
<tr>
<td>Sometimes</td>
<td>11.7 (4.3)*</td>
<td>15.5 (7.6)**</td>
<td>13.4 (3.1)'</td>
<td>18.0 (6.5)**</td>
</tr>
<tr>
<td>Almost never</td>
<td>13.5 (20.5)**</td>
<td>17.9 (24.3)**</td>
<td>13.0 (0.0)</td>
<td>17.5 (3.6)</td>
</tr>
<tr>
<td>Chile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>11.3</td>
<td>13.3</td>
<td>13.4</td>
<td>16.7</td>
</tr>
<tr>
<td>Sometimes</td>
<td>12.1 (7.1)**</td>
<td>14.8 (11.3)**</td>
<td>13.7 (2.2)</td>
<td>17.3 (3.6)*</td>
</tr>
<tr>
<td>Almost never</td>
<td>13.5 (19.5)*</td>
<td>16.1 (21.1)**</td>
<td>14.0 (4.5)</td>
<td>17.7 (6.0)*</td>
</tr>
<tr>
<td>Colombia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>9.9</td>
<td>13.9</td>
<td>11.7</td>
<td>15.7</td>
</tr>
<tr>
<td>Sometimes</td>
<td>11.1 (12.1)**</td>
<td>15.3 (10.1)**</td>
<td>12.2 (4.3)'</td>
<td>15.8 (0.6)</td>
</tr>
<tr>
<td>Almost never</td>
<td>12.4 (25.3)**</td>
<td>15.9 (14.4)**</td>
<td>12.2 (4.3)</td>
<td>16.1 (2.5)</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>9.6</td>
<td>12.8</td>
<td>10.3</td>
<td>13.2</td>
</tr>
<tr>
<td>Sometimes</td>
<td>9.6 (0.0)</td>
<td>13.2 (3.1)</td>
<td>10.8 (4.8)</td>
<td>13.8 (4.5)</td>
</tr>
<tr>
<td>Almost never</td>
<td>10.8 (12.5)**</td>
<td>13.2 (3.1)</td>
<td>10.2 (-1.0)</td>
<td>12.4 (-6.1)</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Country</th>
<th>Working Outside the Home</th>
<th>Working Inside the Home</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Language Test Scores</td>
<td>Mathematics Test Scores</td>
</tr>
<tr>
<td></td>
<td>(Maximum Score = 19)</td>
<td>(Maximum Score = 32)</td>
</tr>
<tr>
<td>Honduras</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>8.9</td>
<td>11.7</td>
</tr>
<tr>
<td>Sometimes</td>
<td>9.4 (5.6)*</td>
<td>12.3 (5.1)**</td>
</tr>
<tr>
<td>Almost never</td>
<td>11.6 (30.3)**</td>
<td>14.5 (23.9)**</td>
</tr>
<tr>
<td>Paraguay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>10.2</td>
<td>12.9</td>
</tr>
<tr>
<td>Sometimes</td>
<td>11.3 (10.8)**</td>
<td>14.9 (15.5)**</td>
</tr>
<tr>
<td>Almost never</td>
<td>12.1 (18.6)**</td>
<td>16.4 (27.1)**</td>
</tr>
<tr>
<td>Peru</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>8.7</td>
<td>11.0</td>
</tr>
<tr>
<td>Sometimes</td>
<td>9.5 (9.2)**</td>
<td>11.2 (1.8)**</td>
</tr>
<tr>
<td>Almost never</td>
<td>11.2 (28.7)**</td>
<td>12.9 (17.3)**</td>
</tr>
<tr>
<td>All countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Often</td>
<td>9.9</td>
<td>13.1</td>
</tr>
<tr>
<td>Sometimes</td>
<td>10.8 (9.0)**</td>
<td>14.2 (8.4)**</td>
</tr>
<tr>
<td>Almost never</td>
<td>12.6 (27.3)**</td>
<td>16.0 (22.1)**</td>
</tr>
</tbody>
</table>

"Difference from often working group significant at the 0.05 confidence level.

**Difference from often working group significant at the 0.01 confidence level.

**Note:** Results are the simple mean test score over all children in the child labor group in the county. Numbers in parentheses are the percentage difference relative to children who often work outside the home when not in school. For definitions of "often," "sometimes," and "almost never," see Table A-1.

**Source:** Authors' computations based on data from the 1997 survey by the Latin-American Laboratory of Quality of Education, as described in the text; UNESCO (2002).
examination. The test advantage for occasional child laborers is smaller but still significant at 8.4 percent for mathematics and 9 percent for languages.

Children were asked a similar question about how intensively they worked inside the home. It seems that working inside the home is less costly for human capital development in schools. Across all countries, those who work often inside the home have average test scores only 7 percent lower than those who almost never work inside the home and only 4 percent lower than those who sometimes work inside the home. The test score gaps for those working outside the home were considerably larger. Furthermore, in only three of the nine countries were average test scores significantly higher for children almost never working inside the home relative to those often working inside the home. In three other countries, those often working inside the home had higher average test scores than did those rarely working inside the home.

Nevertheless, there is a more basic reason for not analyzing the implications of working inside the home on student achievement: more than 95 percent of students reported working inside the home sometimes or often, with nearly identical incidence of work reported for girls and boys and for urban and rural children. This lack of meaningful variation means that the pattern of test scores against work intensity inside the home is unlikely to be reliable. In fact, attempted empirical models could not distinguish statistically between children who did and those who did not work inside the home—everyone was predicted to participate in household labor. It is possible that work inside the home is damaging to schooling outcomes, but our data lack sufficient variation in measured household work to capture the effect. For these reasons, we concentrate our analysis on child labor outside the home.

**Exogenous Variables**

The presumed positive relationship between age and the value of child time working outside the home is used to identify the child labor equation. This relationship varies across urban and rural areas and between boys and girls. It also appears to shift as children reach 10 years of age. This effect is allowed with a spline defined as follows. A dummy variable, \( d_{10} \), takes the value of 1 for children under 10 years of age and 0 otherwise. For children aged 10 and older, the age effect is captured by interactions between \( (1 - d_{10}) \) and age.

The countries included in the data differ in their legal regulations governing the age at which children enter school and when they can leave school. Information on compulsory schooling laws for each country was obtained from the UNESCO (2002). In the empirical specification, these laws shift the age–child labor relationship beyond age 10, using interaction terms of the form \( \text{AGE} (1 - d_{10}) \text{LAGE} \), where LAGE is the legal age of school entry or school exit.\(^{15}\)

---

\(^{15}\) This is a more parsimonious specification than the one with all possible interaction terms. In particular, separate coefficients on the dummy variable \( (1 - d_{10}) \) and their interactions with age, gender, and rural residence did not add to the explanatory power of the child labor equation.
The child's value of time in school will depend on how much the child can learn. This will depend on home attributes that are complementary to child time in school, such as books and parental education, and on the quality of the school. Most of these measures are self-explanatory. However, some of the school variables merit comment. The measure of the classroom environment, inadequacy, is a weighted average of several measures of poor school infrastructure and supplies. Teachers were asked the extent to which they judged classroom lighting, temperature, hygiene, security, acoustics, and textbooks to be inadequate. The weighted sum of the responses is used as the aggregate index of school shortcomings, where the weights were taken as the first principal component from a factor analysis of the teachers' responses. The number of Spanish- or Portuguese-speaking students is included as a measure of the cost of providing schooling services. As the number of nonnative speakers of the language of instruction increases, resources must be diverted to second-language instruction, potentially limiting school productivity.

IV. Econometric Strategy

The results in table 1 suggest a strong negative effect of child market labor on school achievement, but the effect may be in the reverse direction—poor schooling outcomes leading to child labor. The direction of this bias is difficult to predict. The most plausible is that poor school performers are sent to work so that the least squares coefficient on child labor will be biased downward. However, both Tyler (2003) and Stinebrickner and Stinebrickner (2003) found biases in the opposite direction for older students, with better students more likely to work. Measurement error in the self-reported incidence of child labor could also bias the estimated coefficient of child labor on schooling outcomes. The cumulative direction of these sources of bias cannot be established, but both simultaneity and measurement error can be handled by the use of plausible instruments that alter the probability of engaging in child labor without directly affecting test scores.

The first step in the estimation process is to predict child labor. The categorical measure of child market work includes 0 (almost never work), 1 (sometimes work), and 2 (often work). Equation (1) was estimated with an ordered probit specification, using child, parent, school, and community variables to explain variation in market work. Predicted child labor from equation (1) is used as the measure of C in estimating equation (2). This two-stage estimation leads to consistent, but inefficient estimates of the parameters of the achievement equation. A bootstrapping method is used to correct for the inefficiency in the estimators in which 100 samples with replacement are drawn from the original data, subjected to the ordered probit estimation and then inserted into the second-stage achievement equation to simulate the sampling variation in the estimates. The bootstrap standard errors are reported for the test score equations.
Estimates from the probit child labor supply equation, reported in table 2, are needed to identify the effect of child labor on test scores but are also of interest in their own right. The estimation uses the dependent variables reported in table 1 except that data for Mexico and Venezuela are dropped because child's age was not reported. Because the two samples are not identical, separate estimates are reported for the samples of children taking the mathematics and language examinations. The coefficients on the age-interacted variables differ somewhat across the two samples, but the overall relationship between age and child labor is similar between the two samples. The other coefficient estimates are similar across the two samples.

Boys are more likely than girls to work outside the home, and rural boys and girls work more than their urban counterparts, who in turn work more than their metropolitan counterparts. Children of more-educated parents and

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mathematics Test Scores</th>
<th>Language Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exogenous variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.048 (0.009)**</td>
<td>-0.014 (0.009)</td>
</tr>
<tr>
<td>Boy</td>
<td>0.291 (0.036)**</td>
<td>0.163 (0.037)**</td>
</tr>
<tr>
<td>No preschool</td>
<td>-0.016 (0.019)</td>
<td>0.029 (0.019)</td>
</tr>
<tr>
<td>Parents/household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent education</td>
<td>-0.065 (0.007)**</td>
<td>-0.046 (0.008)**</td>
</tr>
<tr>
<td>Books at home</td>
<td>-0.080 (0.012)**</td>
<td>-0.071 (0.012)**</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish enrollment/100</td>
<td>-0.004 (0.002)**</td>
<td>-0.005 (0.002)**</td>
</tr>
<tr>
<td>Inadequate supply</td>
<td>0.062 (0.009)**</td>
<td>0.065 (0.009)**</td>
</tr>
<tr>
<td>Math/week (Spanish/week)</td>
<td>-0.014 (0.004)**</td>
<td>-0.010 (0.003)**</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>0.350 (0.033)**</td>
<td>0.290 (0.034)**</td>
</tr>
<tr>
<td>Urban</td>
<td>0.197 (0.033)**</td>
<td>0.121 (0.031)**</td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boy x rural</td>
<td>-0.019 (0.045)</td>
<td>0.144 (0.045)**</td>
</tr>
<tr>
<td>Boy x urban</td>
<td>-0.062 (0.043)</td>
<td>0.103 (0.044)**</td>
</tr>
<tr>
<td>Age x compulsory start (1 − d10)</td>
<td>0.004 (0.001)**</td>
<td>0.002 (0.001)**</td>
</tr>
<tr>
<td>Age x compulsory end (1−d10)</td>
<td>-0.002 (0.000)**</td>
<td>0.000 (0.001)**</td>
</tr>
</tbody>
</table>

LL | -21,623.743 | -21,179.099 |
Pseudo R² | 0.034 | 0.034 |
Number of observations | 20,699 | 20,290 |

"Significant at the 0.10 confidence level.
"+'Significant at the 0.05 confidence level.

Note: Numbers in parentheses are standard errors. Regressions also include dummy variables that control for missing values.

Source: Authors' computations based on data from the 1997 survey by the Latin-American Laboratory of Quality of Education, as described in the text; UNESCO (2002).
children who have access to more books in the home are less likely to work outside the home. School quality also affects the incidence of child labor. Schools with inadequate supplies encourage child labor. Children in schools with more non-Spanish or non-Portuguese language speakers among their peers are also more likely to work outside the home. Schools that offer more classes in Spanish or Portuguese and mathematics per week also lower the incidence of child labor. In general, these results suggest that better schooling inputs in the home and school lower the incidence of child labor. The exception is that attending preschool does not have a significant effect on child labor in this sample.

The joint test of the null hypothesis that the instrumental variables have no effect on child labor is easily rejected. Variation in truancy laws across countries and in the child labor market for boys within countries does shift the probability that children work. The impact of these laws on the average incidence of child labor is illustrated in figures 1 and 2. The effect was disabled below age 10. As the school starting age rises from ages 5 to 7, the probability of child labor rises about 6 percentage points for a 10-year-old, all else equal, and by 10 percentage points for a 14-year-old (figure 1). As the school-leaving age rises from 12 to 16 years old, the probability of child labor falls by 8.5 percentage points for a 10-year-old and by 11.5 percentage points for a 14-year-old (figure 2). These results suggest that truancy laws do have an effect on child labor on average.

Regional variation in the market for child labor shifts child labor supply for boys and girls (figure 3). The dummy variable spline effectively fixes child labor intensity for children under 10 years of age. After the age of 10, child labor intensity rises for both boys and girls. In each market, boys work more than girls. The higher market labor force participation for boys is consistent with the presumption that the marginal product of child labor is higher for boys than girls. However, rural girls have higher labor force participation than metropolitan boys.

VI. Child Labor and School Achievement

The results from estimating equation (2) both with and without controls for the endogeneity of child labor are reported in table 3. In the specification in table 3, when child labor is treated as exogenous, it takes the values of 0 (almost never work), 1 (sometimes work), or 2 (often work). When treated as endogenous, child labor is a continuous variable with domain over the real line taken as the fitted values from the ordered probit estimation in table 2. The rest of the regressors are the child, household, parent, and school variables used as regressors in table 2.  

16. Ages are truncated below 8 (0.4 percent of the sample) and above 15 (0.8 percent of the sample) because of insufficient observations to generate reliable child labor supply trajectories.

17. Similar estimates of the adverse effect of child labor on test scores were obtained when a school-specific fixed effect was used to control for the impact of variation in school and community variables instead of the vector of school and community variables.
### Table 3. Least Squares and Instrumental Variables Equations on Test Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mathematics Test Scores</th>
<th>Language Test Scores</th>
<th>Mathematics Test Scores</th>
<th>Language Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work outside</td>
<td>-1.184 (0.051)**</td>
<td>-1.087 (0.036)**</td>
<td>-7.603 (1.248)**</td>
<td>-3.980 (0.484)**</td>
</tr>
<tr>
<td>Beta coefficient</td>
<td>[-0.1591]</td>
<td>[-0.2041]</td>
<td>[-0.4081]</td>
<td>[-0.2951]</td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.097 (0.027)**</td>
<td>0.045 (0.019)**</td>
<td>0.309 (0.070)**</td>
<td>0.162 (0.024)**</td>
</tr>
<tr>
<td>Boy</td>
<td>0.731 (0.079)**</td>
<td>-0.165 (0.056)**</td>
<td>2.480 (0.358)***</td>
<td>0.679 (0.155)**</td>
</tr>
<tr>
<td>No preschool</td>
<td>-0.256 (0.093)**</td>
<td>-0.181 (0.066)**</td>
<td>-0.376 (0.088)**</td>
<td>-0.079 (0.040)**</td>
</tr>
<tr>
<td>Parents/household</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent education</td>
<td>0.327 (0.036)**</td>
<td>0.280 (0.026)**</td>
<td>-0.107 (0.106)</td>
<td>0.134 (0.042)**</td>
</tr>
<tr>
<td>Books at home</td>
<td>0.735 (0.061)**</td>
<td>0.497 (0.042)**</td>
<td>0.196 (0.100)**</td>
<td>0.258 (0.037)**</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish enrollment/100</td>
<td>-0.046 (0.008)**</td>
<td>0.022 (0.006)**</td>
<td>-0.079 (0.010)**</td>
<td>0.007 (0.005)</td>
</tr>
<tr>
<td>Inadequate supply</td>
<td>-0.329 (0.046)**</td>
<td>-0.357 (0.031)**</td>
<td>0.073 (0.096)</td>
<td>-0.140 (0.038)**</td>
</tr>
<tr>
<td>Math/week (Spanish/week)</td>
<td>0.027 (0.017)</td>
<td>0.022 (0.006)**</td>
<td>-0.073 (0.016)**</td>
<td>-0.049 (0.012)**</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.730 (0.107)**</td>
<td>0.240 (0.076)**</td>
<td>1.847 (0.225)**</td>
<td>0.794 (0.117)**</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.692 (0.122)**</td>
<td>-0.893 (0.087)**</td>
<td>1.641 (0.410)**</td>
<td>0.275 (0.202)</td>
</tr>
<tr>
<td>Constant</td>
<td>13.778 (0.446)**</td>
<td>10.657 (0.248)**</td>
<td>14.400 (0.453)**</td>
<td>8.045 (0.391)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.084</td>
<td>0.127</td>
<td>0.063</td>
<td>0.091</td>
</tr>
<tr>
<td>Number of observations</td>
<td>20,699</td>
<td>20,290</td>
<td>20,699</td>
<td>20,290</td>
</tr>
</tbody>
</table>

"Significant at the 0.10 confidence level.  
** Significant at the 0.05 confidence level.

Note: Regressions also include dummy variables controlling for missing values.

*aNumbers* in parentheses are standard errors.

*bNumbers* in parentheses are bootstrap standard errors.

The beta coefficients indicate the number of standard deviations the test score will change from a 1 standard deviation increase in child labor.

Source: Authors’ computations based on data from the 1997 survey by the Latin-American Laboratory of Quality of Education, as described in the text; UNESCO (2002).
The impact of child labor on test scores is negative and significant whether child labor is treated as exogenous or endogenous. Because of the difference in the scale of measured child labor across the two specifications, it is difficult to directly compare the magnitude of the implied effect of child labor on test scores. The results are compared in two ways. First, the implied effect of a 1 standard deviation increase above the mean in child labor is computed in each of the equations. When treated as exogenous, a 1 standard deviation increase in child labor causes both mathematics and language tests scores to fall by about 0.2 standard deviations. In other words, children working 1 standard deviation above the mean score on average 8 percent lower on mathematics examinations and 6 percent lower on language examinations than do otherwise identical children working at the mean level. When controlling for endogeneity, the effect increases to 0.4 standard deviation (16 percent) drop in the mathematics examination and a 0.3 standard deviation (11 percent) drop in the language examination. This finding that the magnitude of the child labor effect on academic achievement rises after controlling for endogeneity is consistent with results reported by Tyler (2003) and Stinebrickner and Stinebrickner (2003) for older U.S. students.

Second, the two sets of estimates are compared by tracing the predicted mathematics and language test scores at each decile of the reported and predicted child labor distributions (figures 4 and 5). At the breakpoints of the exogenous measure (going from child labor level 0 to level 1 at the 40th percentile and from level 1 to level 2 at the 74th percentile), the predicted test scores using the reported and corrected measures are close to one another. However, the relationship is steeper at the upper and lower tails of the distribution of predicted child labor, particularly for the mathematics test. The implication is that the impact of child labor on test scores is understated in the first two columns of table 3 by restricting the range of child labor to three discrete levels.

Glewwe's (2002) review of the human capital literature in developing countries argued that cognitive ability as measured by test scores is strongly tied to later earnings as an adult. Returns to schooling for those who worked as children would therefore be expected to be lower than for those who did not work, all else equal. Consistent with that expectation, Ilahi, Orazem, and Sedlacek (forthcoming) found that, holding constant years of schooling completed, Brazilian adults who worked as children received 4–11 percent lower returns per year of schooling completed. The estimates here suggest that child labor outside the home reduces achievement per year of schooling attended by 11–16 percent. Because many of the third and fourth graders in the sample will repeat the grade, the estimates are an upper-bound measure of the lost human capital per year

18. The Davidson–MacKinnon (1993, pp. 237–40) variant of the Hausman test easily rejected the assumption of exogeneity of child labor. The overidentification tests of the instruments failed to reject the null hypothesis of exogeneity at the 10th percentile in the language test sample and at the 5th percentile for the mathematics test sample.
completed, and so the results correspond closely in magnitude to the estimates of Ilahi, Orazem, and Sedlacek of adverse impacts of child labor on earnings.

Most of the other variables have similar effects across the two sets of estimates in table 3, with two main exceptions. The adverse effects of being a boy or being in a rural school disappear in the instrumented equations. Gender and rural residence are closely tied to the incidence of child labor. It is likely that the negative effects on test scores of being male and being in a rural area are related to the indirect effect of these variables on the higher probability that male and rural children work.

Parental education and availability of books in the home lose influence on test scores after controlling for the endogeneity in child labor. School attributes also become less important in explaining test scores. Again, these factors had strong negative effects on child labor, and so part of their positive effect on school outcomes presumably works through their impact on child school attendance and reduced time at work. The literature on the extent to which school quality can explain variation in school achievement has emphasized the large variation in coefficients for the same school inputs across studies and country settings (Hanushek and Luque 2003). The results here suggest that one reason for the uncertain impact of school attributes may be that school quality is more important in affecting school attendance and child labor than in directly affecting test scores.
Figure 5. Predicted Mathematics Test Scores by Child Labor Decile

Maximum = 32, Average score = 14.6

Note: Dashed lines shows 1 standard deviation confidence band for ordered probit estimates. Source: Authors' simulations based on results in table 3, column 3.

VII. Conclusions

Working outside the home lowers average school achievement in samples of third and fourth graders in each of the nine Latin-American countries studied. Child labor is shown to have significant adverse effects on mathematics and language test scores using various specifications correcting for possible endogeneity and measurement error in self-reported child labor intensity. Children who work even occasionally score an average of 7 percent lower on language examinations and 7.5 percent lower on mathematics examinations. There is some evidence that working more intensely lowers achievement more, but these results are more speculative in that empirical models were unable to distinguish clearly between working "sometimes" and working "often."

These adverse effects of child labor on cognitive ability are consistent in magnitude with the estimated adverse effects of child labor on earnings as an adult. Thus, it is plausible that child labor serves as a mechanism for the intergenerational transmission of poverty, consistent with empirical evidence presented by Emerson and Souza (2003) and the theoretical models of poverty traps advanced by Basu (2000), Basu and Van (1998), and Baland and Robinson (2000).

Such large effects suggest that efforts to combat child labor may have substantial payoffs in the form of increased future earnings or lower poverty rates once children become adults. How to combat child labor is less clear. The child labor
supply equations developed here suggest that truancy laws have some effect in lowering the incidence of child labor. However, most of the variation in child labor occurs within countries and not across countries, so policies must address local child labor market and poverty conditions as well as national circumstances in combating child labor. Policies that alter the attractiveness of child labor or bolster household income, such as income transfer programs that condition receipt on child enrollment or reduced child labor, are likely candidates. Recent experience with such programs in Brazil, Honduras, Mexico, and Nicaragua appears to support further development and expansion of such programs.

APPENDIX

TABLE A-1. Variable Description

<table>
<thead>
<tr>
<th>Endogenous variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math score</td>
<td>Mathematics test score (C)</td>
</tr>
<tr>
<td>Language score</td>
<td>Language test score (C)</td>
</tr>
<tr>
<td>Work outside</td>
<td>Index of how often student works outside the home (0–2) (C)</td>
</tr>
<tr>
<td>Often</td>
<td>Student reports that he or she often works outside the home (C)</td>
</tr>
<tr>
<td>Sometimes</td>
<td>Student reports that he or she sometimes works outside the home (C)</td>
</tr>
<tr>
<td>Almost never</td>
<td>Student reports that he or she almost never works outside the home (C)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exogenous variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Student age (years) (C)</td>
</tr>
<tr>
<td>d10</td>
<td>Dummy variable if student is below 10 years old</td>
</tr>
<tr>
<td>Boy</td>
<td>Dummy variable if student is a boy (C)</td>
</tr>
<tr>
<td>No preschool</td>
<td>Student did not attend preschool/kindergarten (C)</td>
</tr>
<tr>
<td>Parents/household</td>
<td></td>
</tr>
<tr>
<td>Parent education</td>
<td>Average education of parent(s) or guardian(s) (P)</td>
</tr>
<tr>
<td>Books at home</td>
<td>Number of books in student’s home (P)</td>
</tr>
<tr>
<td>School</td>
<td></td>
</tr>
<tr>
<td>Spanish enrollment</td>
<td>Total number of Spanish (Portuguese) speaking students enrolled (Pr)</td>
</tr>
<tr>
<td>Inadequate supply</td>
<td>Index of school supply inadequacy (Pr)</td>
</tr>
<tr>
<td>Math/week</td>
<td>Number of mathematics classes per week (Pr)</td>
</tr>
<tr>
<td>Spanish/week</td>
<td>Number of Spanish (Portuguese) classes per week (Pr)</td>
</tr>
<tr>
<td>Community</td>
<td></td>
</tr>
<tr>
<td>Metropolitan area</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>Dummy variable indicating if school is located in an urban area (2,500 to 1</td>
</tr>
<tr>
<td>Rural</td>
<td>Dummy variable indicating if school is located in a rural area (fewer than</td>
</tr>
</tbody>
</table>

| Instruments         |                                                                             |
| Legal structure     |                                                                             |
| Compulsory start    | Compulsory school starting age in the country (U)                           |
| Compulsory end      | Compulsory school ending age in the country (U)                             |

Note: C, child survey or test; P, parent’s survey; T, teacher’s survey; Pr, principal’s survey; S, survey designer’s observation of socioeconomic characteristics of school community; UNESCO estimate.

Source: Authors’ analysis based on data from the 1997 survey by the Latin-American Laboratory of Quality of Education, as described in the text; UNESCO (2002).
### TABLE A-2. Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Observations</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endogenous variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics score</td>
<td>20,699</td>
<td>14.62</td>
<td>5.87</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>Language score</td>
<td>20,290</td>
<td>11.30</td>
<td>4.22</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>Work outside</td>
<td>20,699</td>
<td>0.86</td>
<td>0.79</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Often</td>
<td>20,699</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sometimes</td>
<td>20,699</td>
<td>0.36</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Almost never</td>
<td>20,699</td>
<td>0.39</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Exogenous variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>20,699</td>
<td>9.95</td>
<td>1.59</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>d10</td>
<td>20,699</td>
<td>0.46</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Boy</td>
<td>20,699</td>
<td>0.50</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No preschool</td>
<td>20,699</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Parents/household</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent education</td>
<td>20,699</td>
<td>1.66</td>
<td>1.62</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Books at home</td>
<td>20,699</td>
<td>1.61</td>
<td>1.22</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish enrollment</td>
<td>20,699</td>
<td>439.51</td>
<td>548.82</td>
<td>0</td>
<td>452</td>
</tr>
<tr>
<td>Inadequate supply</td>
<td>20,699</td>
<td>3.68</td>
<td>2.73</td>
<td>0</td>
<td>7.93</td>
</tr>
<tr>
<td>Math/week</td>
<td>20,699</td>
<td>4.66</td>
<td>3.35</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>20,699</td>
<td>0.45</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Rural</td>
<td>20,699</td>
<td>0.35</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compulsory start</td>
<td>20,699</td>
<td>5.94</td>
<td>0.74</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Compulsory end</td>
<td>20,699</td>
<td>13.74</td>
<td>1.13</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

Source: Authors' computations based on data from the 1997 survey by the Latin-American Laboratory of Quality of Education, as described in the text; UNESCO (2002).

### REFERENCES


The Long-Run Economic Costs of AIDS: A Model with an Application to South Africa

Clive Bell, Shantayanan Devarajan, and Hans Gersbach

Primarily a disease of young adults, AIDS imposes economic costs that could be devastatingly high in the long run by undermining the transmission of human capital—the main driver of long-run economic growth—across generations. AIDS makes it harder for victims' children to obtain an education and deprives them of the love, nurturing, and life skills that parents provide. These children will in turn find it difficult to educate their children, and so on. An overlapping generations model is used to show that an otherwise growing economy could decline to a low-level subsistence equilibrium if hit with an AIDS-type increase in premature adult mortality. Calibrating the model for South Africa, where the HIV prevalence rate is over 20 percent, simulations reveal that the economy could shrink to half its current size in about four generations in the absence of intervention. Programs to combat the disease and to support needy families could avert such a collapse, but they imply a fiscal burden of about 4 percent of GDP.

While the costs of AIDS in terms of human suffering are undeniably large, estimates of the associated macroeconomic costs have tended to be more modest, whether their basis be an explicitly formulated economic-demographic model or cross-country regression analysis. Most earlier studies of the former kind that focus on Africa—the continent where the epidemic has hit the hardest—put the annual loss of GDP at about 1 percent.1 These estimates stem from a particular view of how the economy functions: the AIDS-induced increase in mortality, even if it reduces labor supply, also reduces the pressure of population on existing land and capital, thereby raising the productivity of labor. Even if there is an accompanying decline

Clive Bell is a professor of economics at the South Asia Institute of the University of Heidelberg; his email address is clive.bell@urz.uni-heidelberg.de. Shantayanan Devarajan is the chief economist of the South Asia Region at the World Bank; his email address is sdevarajan@worldbank.org. Hans Gersbach is a professor of economics at the Alfred-Weber-Institut of the University of Heidelberg; his email address is gersbach@uni-heidelberg.de. The authors thank Ramona Brühns for excellent programming and research assistance. They are also grateful to Bernhard Pachl, Lars Siemers, and participants in seminars at the University of California at Berkeley, Cornell University, the Center for Global Development, the London School of Economics, and the World Bank, as well as to three anonymous referees and the editor of this journal, for their valuable and constructive comments and suggestions. The authors are responsible for any remaining errors of analysis.

1. See, for example, Arndt and Lewis (2000); Cuddington (1993); Cuddington and Hancock (1994); Kambou, Devarajan, and Over (1992); and Over (1992).

www.worldbank.org/wber
in aggregate savings and investment (from the reallocation of expenditures toward medical care, for instance), the net impact on the growth of GDP per capita turns out to be small. Econometric investigations based on country panel data yield the same result. Bloom and Mahal (1997), for example, found no effect on GDP at all; on returning to the question later with new data, Bloom and others (2001) managed to extract a small adverse effect.

This article argues that the long-run economic costs of AIDS are almost certain to be much higher—and possibly devastating. In doing so, it joins company with some other authors (Corrigan, Glomm, and Méndez 2004, 2005; Ferreira and Pessoa 2003), who recently and independently have pursued an approach based on an overlapping generations framework. This approach involves a very different view of how the economy functions over the long run, one that emphasizes the importance of human capital and its transmission across generations. The accumulation of human capital—that is, the stock of knowledge and abilities embodied in the population—is the force that generates economic growth over the long run. The mechanism that drives the process is the transmission of knowledge and abilities from one generation to the next.

The implications of this model are particularly relevant to Africa, the continent with the lowest level of human capital and the highest prevalence of the disease. In many African countries, AIDS presents a formidable hurdle to long-run economic growth. The application of the model to South Africa, that Sub-Saharan outlier with relatively high levels of income and human capital (and HIV prevalence), reveals that in the absence of specific interventions, a decline from middle-income status is possible in the long run.

The argument establishing how AIDS can severely retard economic growth is made in three steps.² First, AIDS destroys existing human capital in a selective way, striking primarily young adults. Some years after they have been infected, it reduces their productivity by making them sick and weak. It then kills them in their prime, destroying the human capital formed in them through child-rearing, formal education, and learning on the job.

Second, AIDS weakens or even wrecks the mechanisms that create human capital in the next generation. The quality of child-rearing depends heavily on the parents’ human capital. If one or both parents die before their offspring reach adulthood, the transmission of knowledge and potential productive capacity across the two generations will be weakened. At the same time, the loss of income due to disability and early death reduces the lifetime resources available to the family, which can lead to the children spending much less time, if any, at school. The chance that the children will contract the disease in adulthood also makes investment in their education less attractive, even when both parents themselves remain uninfected. The weakening of these transmission processes

². The argument here is confined to those factors that are the most important. For a longer list of the epidemic’s economic effects and related discussion, see, for example, Bell, Devarajan, and Gersbach (2004) and Corrigan, Glomm, and Mendez (2005).
is insidious: its effects are felt only over the longer run, as the poor education of children today translates into low productivity of adults a generation hence.

Third, as the children of AIDS victims become adults with little education and limited knowledge received from their parents, they are less able to invest in their own children's education, and a vicious cycle ensues. If nothing is done, the outbreak of the disease can eventually precipitate a collapse of economic productivity. Early in the epidemic, the damage may appear to be slight, but as the transmission of capacities and potential from one generation to the next is progressively weakened and the failure to accumulate human capital becomes more pronounced, the economy will begin to slow down, with the growing threat of a collapse to follow.

The argument has two important implications for economic policy. The first is fiscal. By killing off mainly young adults, AIDS also seriously weakens the tax base and thus reduces the resources available to meet the demands for public expenditures, including those aimed at accumulating human capital, such as education and health services not related to AIDS. As a result, the state's finances will come under increasing pressure, exacerbated by the growing expenditures on treating the sick and caring for orphans.

The other effect is an increase in inequality. If orphaned children are not given the care and education enjoyed by those whose parents remain uninfected, the weakening of the intergenerational transmission mechanism will express itself in increasing inequality among the next generation of adults and the families they form. Social customs of adoption and fostering, however well established, may not be able to cope with the scale of the problem, thereby shifting the onus onto the government, which is likely to experience increasing fiscal difficulties and thus to lack the resources to assume this additional burden.

The policy objective, therefore, is to avoid such a collapse. The instruments available for this purpose are (a) spending on measures to contain the disease and treat the infected, (b) aiding orphans, in the form of income support or subsidies contingent on school attendance, and (c) taxes to finance the expenditure program. The central policy problem is to find the right balance among these interventions to ensure economic growth over the long run without excessive inequality.

This article relates to recent contributions to the literature as follows. Those that adopt an overlapping generations framework have chosen somewhat different points of emphasis. In the model here, higher mortality risk undermines the formation of human capital through three channels. First, if one or both parents die early, their children will have less productive capacity because less human capital is transmitted. Second, the loss of income caused by early death in a family reduces schooling. Third, the chance that the children will be infected as adults makes investment in their education less attractive. Corrigan, Glomm, and Méndez (2004, 2005) consider only the first two channels, but they allow for effects on the accumulation of physical capital, which are absent in the model.
here. Thus, this article complements theirs in the task of establishing how AIDS might influence the course of per capita income. As for the possible magnitude of these effects, Corrigan, Glomm, and Méndez (2004) calibrate their model to Sub-Saharan economies and find that for infection rates of around 15–20 percent, the growth rate of per capita income drops about 30–40 percent. Ferreira and Pessoa (2003) concentrate on the reduced returns to investment in schooling in a setting free of uncertainty, and estimate that the time devoted to it can decline by up to a half.

Young (2005) adopts a quite different perspective on how the AIDS epidemic impinges on the South African economy. He embeds a Beckerian household model, with endogenous participation, fertility, and education decisions, in a Solovian constant-savings-rate macroeconomic framework. In estimating the behavioral equations and simulating the evolution of the South African economy, two competing effects are emphasized. On the one hand, the epidemic is likely to have a negative impact on orphans' accumulation of human capital. On the other hand, high prevalence rates lower fertility. Young finds that even with the most pessimistic assumptions regarding educational attainment, the fertility effect dominates and future per capita consumption possibilities are enhanced. Although more channels through which the epidemic may harm human capital accumulation are considered here, fertility is exogenous. Sensitivity tests are conducted, however, and these reveal that changes in the level of fertility have only minor effects on the growth of productivity. Bruhns (2005) develops a closely related theoretical model in which households choose the level of fertility and applies it to Kenya. Her conclusions are broadly similar to the ones arrived at here.

Some econometric studies look at aspects of the link between AIDS and human capital. McDonald and Roberts (2004) estimate an augmented Solow model that incorporates both health and education capital. They employ a panel of 122 countries over a longer timespan than that of Bloom and others (2001) and conclude that the macroeconomic effects of HIV/AIDS have been substantial, especially in Africa, where the average marginal impact on income per capita of a 1 percent increase in the HIV prevalence rate is estimated to be $-0.59$ percent. Hamoudi and Birdsall (2004) provide indirect econometric evidence that AIDS reduces schooling in Africa. Using data from Demographic and Health Surveys conducted in 23 Sub-Saharan African countries and employing two specifications, they settle on the estimate that a reduction in life expectancy at birth of 10 years is associated with a fall of 0.6 years in the average schooling attained by that cohort. Given that life expectancy at birth in most countries in Southern and East Africa fell by at least 10 years over 1985–2000 (Dorrington and Schneider 2001) and that average schooling among the population aged 25–49 years was in the modest range of 3–6 years, this is a significant and disturbing finding. Although their measure of mortality differs from the one used here, their finding supports the general approach adopted here. Other

3. The term is theirs, but a close reading strongly suggests that they mean prevalence rates.
recent microeconomic work suggests that orphans indeed suffer various set-
backs. Gertler, Levine, and Martinez’ (2003) study of Indonesian children, for
example, shows that orphans are less healthy, less likely to go to school, and
overall less prepared for life. Case, Paxson, and Abeidinger (2002) found for a
group of African countries that the schooling of orphans depends heavily on how
closely related they are to the head of the adopting household.

In section I of this article, we tackle the question of how AIDS impinges on the
economy conceptually by extending the model of Bell and Gersbach (2001) to
deal with disease-ridden environments, in which premature adult mortality is
increased by the outbreak of an epidemic. Parents have preferences over current
consumption and the level of human capital attained by their children. The
decision about how much to invest in education is influenced by premature
adult mortality in two ways: the family’s lifetime income depends on the adults’
health status and the expected payoff depends on the level of premature mortal-
ity among the children when they reach adulthood.

In section II, we apply the model to South Africa. The choice of South Africa
as a test bed is a natural one on several grounds. First, the very nature of the
model demands that the available economic and demographic series be long and
fairly reliable if the base for calibration is to be solid. Second, South Africa is a
middle-income country that has experienced substantial growth over much of
the past half century. A collapse of the kind analyzed in section I, were it to
occur, would therefore mean that there is a long way to fall. Third, the epidemic
has progressed rapidly in South Africa, from a prevalence rate among the
population aged 15–49 years of about 1 percent in 1990 to just over 20 percent
in 2003 (UNAIDS 2004).

Finally, in section III, we examine policies to avert the long-run economic
decline caused by AIDS. Interventions in the spheres of health and education are
examined. Finding the right balance between these two sets of measures is the
central policy problem, and the results in this section attempt to illuminate how
the balance should be struck. In any case, the sheer magnitude of the problem
indicates that additional public spending of the order of 3–4 percent of GDP may
be needed to contain the epidemic and ward off its worst effects.

I. The Model

There are two periods of life, childhood and adulthood. On becoming adults,
individuals form families and have children. When the children are very young,
they can neither work nor attend school. Since investment in education is
assumed to be the only form of investment, the family’s full income is wholly
consumed in this phase. Only after this phase is over, do the adults learn whether
they will die prematurely — and thus leave their children as half or full orphans.
Early in each generation of adults, therefore, all nuclear families fall into one of
the following four categories: (a) both parents survive into old age, (b) the father
dies prematurely, (c) the mother dies prematurely, and (d) both parents die prematurely. These states are denoted by \( s_t \in S_t := \{1, 2, 3, 4\} \). The subjective probability that a family formed at the start of period \( t \) lands in category \( s_t \) is denoted by \( \pi_t(s_t)^4 \). Once their states have been revealed, families make their allocative decisions accordingly, and the formation of human capital takes place. What follows is a terse account of the main elements; the details are set out in appendix A1.

Human capital is formed by a combination of child-rearing, whose quality depends on the parents' combined human capital, \( A_t(s_t) \), and the child's formal education, \( e_t \), expressed as a fraction of school-going years. A child so reared in period \( t \) attains the following level of human capital in period \( t + 1 \):

\[
\lambda_{t+1} = \begin{cases} 
\frac{z(s_t)f(e_t)A_t(s_t) + 1}{\xi}, & s_t = 1, 2, 3 \\
1, & s_t = 4 
\end{cases}
\]

where \( z(s_t) \) represents the strength with which capacity is transmitted across generations, \( f(\cdot) \) can be thought of as the "educational technology," and the presence of the 1 in the upper branch grants this basic (normalized) level of human capital to wholly uneducated adults. \( \xi(\leq 1) \) is the level of human capital attained by full orphans who grow up without care or education.

Let an individual's output be proportional to his or her level of human capital, an assumption that is certainly plausible over the very long run. Then a household with \( n_t \) children that finds itself in state \( s_t \) will have a well-defined level of full income, which the adults can allocate between consumption and investment in the children's education. The latter pays off in the form of each child's human capital on reaching adulthood. The (surviving) parents' optimal level of such investment, \( e_t^0[A_t(s_t), s_t, \kappa_{t+1}^e] \), depends on the level of full income, the relative price of education, the strength of their altruism toward their children, and the expected level of premature adult mortality in period \( t + 1 \),

\[
\kappa_{t+1}^e \equiv \frac{1 + \pi_{t+1}(1) - \pi_{t+1}(4)}{2},
\]

as they subjectively estimate it in period \( t \). Substituting \( e_t^0 \) into equation 1 yields

\[
\lambda_{t+1} = \begin{cases} 
\frac{z(s_t)f(e_t^0[A_t(s_t), s_t, \kappa_{t+1}^e])A_t(s_t) + 1}{\xi}, & s_t = 1, 2, 3 \\
1, & s_t = 4 
\end{cases}
\]

Equation (2) describes a random dynamical system. Note that each child in any given family state \( s_t \) attains the same \( \lambda_{t+1} \) in adulthood with certainty, but he or she can wind up in any of the states \( s_{t+1} \in \{1, 2, 3, 4\} \) after reaching adulthood and forming a family in period \( t + 1 \), and the succeeding branches proliferate in the future. The attendant threat of growing inequality will occupy an important

---

4. The population is assumed to be large enough that this is also the fraction of all families in that state after all premature adult deaths have occurred. Observe that these probabilities change over the course of the epidemic (table 2).
place in the analysis of policy interventions, but there is no space to go into the dynamical properties of the system in any detail here. What follows is aimed only at clarifying certain of their qualitative features.

It suffices, for this particular purpose, to look at what happens when there is no premature adult mortality \( \pi_{t+1}(1) = 1 \) for all \( t \), so that the only state that is ever observed is \( s_t = 1 \). To derive the typical dynamics, it is assumed that altruism is not operative when the adults are uneducated, that is, \( \epsilon_t^a = 0 \) when \( \lambda_t \) is sufficiently close to 1. It can then be shown that the system has at least two stationary states with respect to human capital if \( z(1)f(1)2\lambda^a + 1 > \lambda^a \), where \( \lambda^a \) is the lowest level of an adult's human capital such that a two-parent household chooses full education for the children in such an environment (Bell and Gersbach 2001).

The resulting phase diagram is illustrated in figure 1, where \( \Lambda^d > 2 \) denotes the smallest endowment of the adults' human capital such that they begin to send their children to school and \( \Lambda^a(= 2\lambda^a) \) denotes the corresponding endowment at which children finally enjoy full-time schooling. As depicted, the system has just two stationary states with respect to human capital. One is the state of economic backwardness, defined as \( \Lambda = 2 \). This stable state is a poverty trap, wherein all generations are at the lowest level of human capital. The other is an unstable state \( (\Lambda, = \Lambda^* V_t) \), in which the parents' human capital is such that they

---

**Figure 1. Phase Diagram Without Premature Adult Mortality**
choose a positive level of education for their children, who then attain $A^*/2$ in adulthood. To be precise, and recalling equation (2), $A^*$ satisfies

$$\frac{A^*(1)}{2} = z(1)f(\{e_t^0(A^*(1), 1, 1)\}A^*(1) + 1$$

where $\pi_t(1) = 1$ for all $t$. Observe that, starting from any $A > A^*$, unbounded growth is possible if and only if $2z(1)f(1) \geq 1$, and that the growth rate then approaches $2z(1)f(1) - 1$ asymptotically.

These results reveal that the intrusion of premature adult mortality may affect the system's dynamics not only by changing the probabilities of the states but also by increasing the values of $A^d$, $A^a$ and $A^a$ for states 1, 2, and 3, respectively, and thereby increasing the range of human capital levels within which a progressive decline into backwardness will set in. This turn of events is now examined in more detail.

**Disease, Increasing Inequality, and Economic Collapse**

The process by which the outbreak of an epidemic like AIDS may lead to economic collapse can be described as follows. At the start of period $t = 0$, a society of homogeneous, two-parent families, each with adult human capital endowment $2\lambda_0$, is suddenly assailed by a fatal disease. While the children are still young, all adults learn whether they are infected with the disease, and the survivors then choose the consumption--education bundle $(c_t(s_0), e_t(s_0))$ for $s_0 = 1, 2, 3$. How does the outbreak affect the subsequent development of the society? Children who are left as unsupported orphans ($s_0 = 4$) fall at once into the poverty trap. Even if both parents survive but have been such orphans in childhood, they cannot afford to send their children to school (as assumed above), and their succeeding lineage remains there. To discover what happens to the rest, the critical value function $\lambda^*(s, \kappa)$ is introduced for $s \in \{1, 2, 3\}$, which is defined for stationary fertility and mortality, $n_t = n$, $\forall t$ and $\kappa_t = \kappa$, $\forall t$. In this setting, it is natural to assume perfect foresight, namely $\kappa_{t+1}^* = \kappa_t = \kappa$, $\forall t$.

$$\lambda^*(s, \kappa) = z(s)f(\{e_t^0(A^*(s), s, \kappa)\})A^*(s) + 1$$

where $A^*(1) = 2\lambda^*(1)$, $A^*(2) = A^*(3) = A^*(2) = A^*(3)$, and $\kappa$ is a sufficient statistic of premature adult mortality in the stationary state, in which, by definition, all expectations are realized. $\lambda^*(s, \kappa)$ is the stationary-state level of human capital associated with a particular state $s$, that is, in any pair of generations, parent or parents and offspring share the same state. Equation (3) states that if adults with human capital $A^*$ find themselves in family state $s$ and the mortality environment $\kappa$, they will make choices for their children such that the latter will attain the same level of human capital on reaching adulthood.
The critical value function has two key properties, which are established in Bell, Devarajan, and Gersbach (2003):

1. \( \frac{\partial \lambda^*(s, \kappa)}{\partial \kappa} < 0, \quad s = 1, 2, 3 \)
2. \( \lambda^*(1, \kappa) \leq \lambda^*(2, \kappa) = \lambda^*(3, \kappa) \)

The first property implies that a permanent increase in premature adult mortality may cause a group that was earlier enjoying self-sustaining growth to fall into the poverty trap. The second property implies that single-parent families generally need higher individual levels of human capital than two-parent ones to escape the trap, in which case an increase in premature adult mortality also increases the share falling into the poverty trap by increasing the proportion of one-parent families.

In the long run, if nothing is done to support full orphans and the children of needy, one-parent households, the share of uneducated families will grow until, in the limit, the whole population is in a state of economic backwardness. Not only do some adults meet an early death but the whole society descends progressively into the poverty trap. Two questions arise. First, what are the chances that the AIDS epidemic will so increase the level of premature adult mortality as to precipitate a collapse? Second, what arrangements for support and insurance are there to prevent such a collapse? These questions are addressed with reference to South Africa in the next section.

II. AN APPLICATION TO SOUTH AFRICA

This section falls into two parts. In the first, we cover the results of the calibration rather than the procedure itself, the details of which can be found in Bell, Devarajan, and Gersbach (2003). The robustness of the calibration is examined using a sensitivity analysis of the critical value function. In the second part, we develop three benchmark simulations of the model so calibrated.

Calibration and Sensitivity Analysis

Beginning with the fundamental difference equation (1), the parameters \( z(s) \), the functional form \( f(e) \), and the boundary value of \( \lambda \) are needed. In view of the highly nonlinear nature of the system and the limited information available, the form \( f(e) = e \) is chosen. Since the unit time period of the model is a generation, with two overlapping generations, it is defensible to set the span of each at 30 years.

Inspection of the series for South African GDP reveals that the period from 1960 to 1975 was one of fairly steady and appreciable growth. This early subperiod is viewed as plausible initial basis for assessing how the post-apartheid
economy ought to be able to perform over the long haul. Denoting calendar years by the subscript \( k \) and ignoring child labor, \( \text{GDP} \) in year \( k \) is

\[
Y_k = \alpha L_k \lambda_k
\]

where \( L_k \) and \( \lambda_k \) denote the size of the labor force and the average level of efficiency in that year, respectively (table 1), and the parameter \( \alpha \) is the productivity of a unit of human capital. Since the labor force series begins in 1965, that year is the starting point for the calibration procedure. The series for \( e \) is quinquennial and takes the form of the average years of schooling among the population aged 25 years and older—for example, 4.06 years in 1960. Defining full schooling as 10 years (ages 6–15 years inclusive) yields an average value of \( e \) for those born between 1905 and 1935 of 0.406, which is denoted by \( e_{1960}^B \).

Employing equation (1) recursively, together with the relation between a family's earnings and its endowment of human capital and the series in table 1, yields the estimates \( \lambda = 0.818 \), \( \alpha = 3.419 \), and \( \lambda_{1965} = 2.696 \). The final step is to shift the starting point to 1960. As pointed out in the introduction, the AIDS prevalence rate rose from about 1 percent in 1990 to just over 20 percent a decade later. This is a strong argument for choosing 1990 as the date of the outbreak of the epidemic in South Africa, and hence 1960 as the starting point in the chosen 30-year framework. The interpolation from table 1 implies that \( \lambda \) grew at an annual rate of 0.58 percent between 1965 and 1990; thus \( \lambda_{1960} = \lambda_{1965} / (1.0058)^5 = 2.620 \).

Two comments on these estimates are in order. First, the parameter \( \alpha \) has the dimension of 1995 U.S. dollars per efficiency unit of labor per year. According to these estimates, therefore, a two-parent household in 1960 with two economically active adults and all the children attending school full-time would have had a family income of \( \alpha \lambda_{1960} \) or $17,915. In the event of a complete collapse that left the entire population uneducated, the family's income would be just $6,840

<table>
<thead>
<tr>
<th>Year</th>
<th>( Y_k ) (1995 U.S. Dollar)</th>
<th>( e_k^B )</th>
<th>( L_k )</th>
<th>( Y_k/L_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>( 49.2 \times 10^9 )</td>
<td>0.406</td>
<td>Not available</td>
<td>Not available</td>
</tr>
<tr>
<td>1965</td>
<td>( 68.4 \times 10^9 )</td>
<td>0.410</td>
<td>( 7.42 \times 10^6 )</td>
<td>9,220</td>
</tr>
<tr>
<td>1970</td>
<td>( 90.6 \times 10^9 )</td>
<td>0.447</td>
<td>( 8.24 \times 10^6 )</td>
<td>10,990</td>
</tr>
<tr>
<td>1975</td>
<td>( 113.0 \times 10^9 )</td>
<td>0.453</td>
<td>( 9.25 \times 10^6 )</td>
<td>12,230</td>
</tr>
<tr>
<td>1980</td>
<td>( 127.4 \times 10^9 )</td>
<td>0.461</td>
<td>( 10.34 \times 10^6 )</td>
<td>12,320</td>
</tr>
<tr>
<td>1985</td>
<td>( 132.4 \times 10^9 )</td>
<td>0.495</td>
<td>( 11.93 \times 10^6 )</td>
<td>11,100</td>
</tr>
<tr>
<td>1990</td>
<td>( 144.7 \times 10^9 )</td>
<td>0.500</td>
<td>( 13.58 \times 10^6 )</td>
<td>10,650</td>
</tr>
<tr>
<td>1995</td>
<td>( 151.0 \times 10^9 )</td>
<td>Not available</td>
<td>( 15.29 \times 10^6 )</td>
<td>9,880</td>
</tr>
<tr>
<td>2000</td>
<td>( 172.1 \times 10^9 )</td>
<td>Not available</td>
<td>( 16.98 \times 10^6 )</td>
<td>10,130</td>
</tr>
</tbody>
</table>

in the absence of child labor. Second, the estimate of \( z \) yields the value of the intergenerational growth factor when children attend school full-time, namely \( 22 = 1.636 \). This corresponds to an annual growth rate of productivity of about 1.64 percent over the long run, which seems rather modest in light of the East Asian experience, but quite in keeping with South Africa's recent performance.

The form of social organization has thus far remained conveniently in the background, but now that preferences must be specified, a definite choice is unavoidable. For much of the period in question, South Africa was quite rural, so one can make the case that there was widespread pooling of orphaned children, with all surviving parents caring for all children. This arrangement is a salient feature of the benchmark cases to be analyzed below. Let preferences over current consumption and the children's attained level of human capital on reaching adulthood be logarithmic:

\[
EU_t = 2[b \ln c_t(0) + n_t \left( \frac{\kappa_{t+1}}{\kappa_t} \right) \ln \lambda_{t+1}]
\]

where the state \( s_t = 0 \) denotes pooling, and a representative pair of surviving adults cares for \( n_t/\kappa_t \) children, all of whom are valued and treated identically. Given that the calibration is anchored to 1960, both \( \kappa_{60} \) and households' expectations in 1960 concerning the level of \( \kappa_{90} \) are needed. The realized value of \( \kappa_{90} \) was 0.86. The great reductions in mortality in those three decades benefited children far more than adults, however, so that it is defensible to set the expected value of \( \kappa_{90} \) at the actual value of \( \kappa_{60} \). Finally, it is assumed that in 1960 a representative couple, unaware of and untouched by AIDS in any way, chose the average years of schooling attained by the generation born between 1935 and 1965. This yields the value \( b = 33.45 \).

To complete the array of economic parameters, estimates of \( \beta \), the fraction of an adult's consumption to which each child has a claim, and \( \gamma \), a child's human capital when employed as a child laborer, are needed. Setting \( \beta \) at 0.5 seems

---

### Table 2. Family State Probabilities Corresponding to Premature Adult Mortality Rates

<table>
<thead>
<tr>
<th>Year</th>
<th>( \pi(1) )</th>
<th>( \pi(2) )</th>
<th>( \pi(3) )</th>
<th>( \pi(4) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20920</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990 (( D = 0 ))</td>
<td>0.855</td>
<td>0.101</td>
<td>0.039</td>
<td>0.005</td>
</tr>
<tr>
<td>2010 (( D = 1 ))</td>
<td>0.294</td>
<td>0.165</td>
<td>0.347</td>
<td>0.194</td>
</tr>
<tr>
<td>30920</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990 (( D = 0 ))</td>
<td>0.793</td>
<td>0.164</td>
<td>0.080</td>
<td>0.018</td>
</tr>
<tr>
<td>2010 (( D = 1 ))</td>
<td>0.112</td>
<td>0.180</td>
<td>0.272</td>
<td>0.436</td>
</tr>
</tbody>
</table>

*Note*: \( q_{s,x} \) denotes the probability that an individual will reach the age of \( s + x \) years, conditioned on reaching the age of \( x \) years. State probabilities do not correspond to the actual years shown but to the steady states associated with each disease environment (\( D = 0 \), 1).

*Source*: Authors' computations based on Dorrington and others (2001).
unobjectionable. A much lower value of \( \gamma \) is called for: \( \gamma = 0.2 \) yields a maximal level of annual earnings from a child's labor of \( \alpha \gamma = $685 \), which may be on the high side, but this is balanced by the fact that no direct costs of schooling have been included.

Turning to the demographic components of the model, the population roughly doubled between 1960 and 1990, so that in keeping with the assumptions in section I and the generation span of 30 years, each mother had, on average, four surviving children over that period. Whether AIDS will affect fertility in the future is unclear (some evidence points to a modest decline), but what is certain is that AIDS has already contributed to a marked rise in mortality among children under the age of 5 years (Dorrington and others 2001). Since there is also some evidence that fertility had started to fall by the early 1990s (World Bank 2002), it is assumed that each mother will have three surviving children from 1990 onward.

The overriding concern in calibrating the model demographically, therefore, is with premature mortality among adults. The benchmark case is that where there is no epidemic (\( D = 0 \)), which, in view of the low prevalence rate in 1990, is taken to be the age-specific mortality profile for that year, as set out in Dorrington and others (2001). The second reference case is that where the epidemic has reached maturity (\( D = 1 \)) in the absence of any effective measures to combat it. The corresponding profile is assumed to be Dorrington and others' forecast for 2010.

The next step is to calculate the corresponding state probabilities \( \pi_t(s_t) \), which requires an assumption about the incidence of the disease among couples. The probability of transmission within a union appears to be of the order of 10 percent a year under the conditions prevailing in East Africa (Marseille, Hofmann, and Kahn 2002), which, when cumulated over the median course of the disease from infection to death of a decade, implies that the probability of the event that both partners become infected, conditional on one of them getting infected outside the relationship, is about 0.65. Given the uncertainties involved, a less concentrated pattern of mortality within families has been assumed, namely that the incidence is independently and identically distributed. The resulting state probabilities are set out in table 2, where their values correspond not to the actual years shown but rather to the steady states associated with each disease environment (\( D = 0, 1 \)).

The appalling dimension—social, economic, and psychological—of the epidemic in its mature phase are plain. In the absence of AIDS, 85 percent of all children would grow up enjoying the care, company, and support of both natural parents, and fewer than 1 percent would suffer the misfortune of becoming full orphans (table 2). If the epidemic is left to run unchecked, it will leave almost 20 percent of the generation born from 2010 onward full orphans, about 50 percent will lose one parent in childhood, and a mere 30 percent or so will reach adulthood without experiencing the death of one or both parents. The epidemic will also reverse the usual pattern of excess mortality among fathers—from
about twice as high as among mothers to a third to a half lower. Given the mother's special role in securing the young child's healthy development, it can be argued that this reversal imparts additional force to the shock.

The final step is to undertake some sensitivity analysis. Since the decisive factor in the system's dynamics is how $\lambda_t$ lies in relation to the critical, steady-state values $\lambda^*(s, \kappa, n, z)$, an appropriate way of investigating the robustness of the calibration procedure is to examine the sensitivity of $\lambda^*(\cdot)$ to variations in the parameter values estimated or derived above. The values of $z$, $\alpha$, and $\lambda_{60}$ are estimated jointly, so one cannot be varied without modifying the others. Two types of sensitivity analysis can be performed. First, the three parameters can be varied within this straitjacket. Second, taking $\alpha$ and $\lambda_{60}$ as given, $z$ (as well as $\kappa$ and $n$) can be varied in such a way that the whole configuration is actually more optimistic than the one that emerged from the calibration (forexample, by setting $z$ in excess of 0.818). The second approach is chosen because it evaluates the robustness of the findings over a much wider domain and allows the parameters to take new values after 1990 (as already indicated for $\kappa$ and $n$).

Table 3 sets out the values of $\lambda^*(\cdot)$ for a variety of plausible parameter values. In keeping with the above discussion of fertility and mortality, the choices are $n = 3$ and $n = 4$, with $\kappa = 0.860$ and $\kappa = 0.338$, which correspond to $D = 0$ and $D = 1$, respectively. The intermediate value $\kappa = 0.6$ represents a less dramatic, or waning, epidemic. In addition to the calibrated value $z = 0.818$, somewhat more optimistic values can be considered, namely 0.9 and 1.0, as well as the possibility that the future value of $z$ may be reduced by the higher dependency ratio that will attend higher premature adult mortality (say, $z = 0.7$). Beginning with the calibrated values $n = 4$, $\kappa = 0.860$, and $z = 0.818$, equation (3) yields $\lambda^*(s, \cdot) = 2.06$, 2.10, and 4.33 for $s = 0$, 1, and 2, respectively, the first two of which lie comfortably below $\lambda_{60}$. Since the fraction of one-parent households under nuclear family arrangements was a modest 14 percent (see table 2), it can be assumed that the implicit burden of supporting them and full orphans was both tolerable and actually taken up. It follows that regardless of the family arrangements actually in force, the South African economy had already been launched on a path toward steady-state growth before the epidemic broke out in the early 1990s.

The reduction in fertility from $n = 4$ to $n = 3$ after 1990 has only a very slight effect on $\lambda^*$ under both family arrangements. The fall in $n$ implies a smaller weight on the term for altruism toward children, but this is just outweighed by the correspondingly smaller claims that fewer children make on the family's resources—whether they are raised under pooling or within a nuclear family. Indeed, this effect is small in all the parameter constellations in table 3, which leads to the conclusion that plausible changes in fertility do not play an important role in determining the qualitative nature of the system's dynamics.

The other striking feature of table 3, by contrast, is the sensitivity of $\lambda^*(s, \cdot)$ to $\kappa$. In all variations for $\kappa = 0.338$ (that is, $D = 1$), $\lambda^*(s, \cdot) > \lambda_{90} = 3.14$, which
points to a progressive economic collapse in the face of an undiminished continuation of the epidemic and in the absence of any countervailing intervention. If $\kappa = 0.6$ and $n = 3$, this fate is avoided under both family arrangements (assuming, as above, that needy families will be supported) when $z$ takes the value 0.9 or higher. When $z$ takes the calibrated value 0.818, however, the pooling arrangement only barely escapes the trap, whereas the two-parent nuclear family ($s = 1$) barely slips into it. Summing up, these results suggest that even allowing for some uncertainty about the calibrated values of $z$ and $\lambda_{60}$ and about the estimated value of $\kappa$ in the steady state corresponding to $D = 1$, as well as the behavior of fertility, the current course of the epidemic poses a very real threat to the long-term growth of the South African economy.

Simulations

Three simulations of the course of the economy for the period after 1990 form the set of benchmarks.

**Benchmark 1: Pooling, No AIDS.** The corresponding trajectory of the variable $A_0$, about which all else revolves, is plotted in figure 2. As noted above, the key feature of this story is that steady-state growth is ultimately attained. Starting from the modest level of 0.5 in 1960, education becomes virtually full-time in the generation born from 2020 onward, by which point, income per head is two-thirds higher than in 1960, with another increase of 80 percent in the next generation. The burden of child-dependency is limited throughout: 0.65 adopted children per couple in addition to the four of their own before 1990 and 0.49 in addition to the three of their own thereafter. This is the relatively happy counterfactual into which AIDS intrudes at $t = 0$ (1990).

| $n = 3$ | $s = 0$ | 2.58 | 2.14 | 1.90 | 1.67 | 3.74 | 3.11 | 2.77 | 2.43 | 6.71 | 5.60 | 5.00 | 4.40 |
| $s = 1$ | 2.61 | 2.18 | 1.94 | 1.71 | 3.84 | 3.23 | 2.91 | 2.58 | 6.94 | 5.90 | 5.33 | 4.76 |
| $s = 2$ | 5.30 | 4.44 | 3.97 | 3.50 | 7.74 | 6.53 | 5.87 | 5.22 | 13.92 | 11.82 | 10.69 | 9.56 |

| $n = 4$ | $s = 0$ | 2.51 | 2.06 | 1.81 | 1.56 | 3.66 | 3.01 | 2.66 | 2.31 | 6.60 | 5.45 | 4.83 | 4.21 |
| $s = 1$ | 2.55 | 2.10 | 1.86 | 1.62 | 3.79 | 3.17 | 2.84 | 2.51 | 6.91 | 5.85 | 5.27 | 4.70 |

**Table 3.** Critical Values of $\lambda^*$ ($s, k, z, n$)

<table>
<thead>
<tr>
<th>$\kappa$</th>
<th>0.860</th>
<th>0.600</th>
<th>0.338</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z = 0.7$</td>
<td>0.818</td>
<td>0.9</td>
<td>1.0</td>
</tr>
</tbody>
</table>

| $n = 3$ | $s = 0$ | 2.58 | 2.14 | 1.90 | 1.67 | 3.74 | 3.11 | 2.77 | 2.43 | 6.71 | 5.60 | 5.00 | 4.40 |
| $s = 1$ | 2.61 | 2.18 | 1.94 | 1.71 | 3.84 | 3.23 | 2.91 | 2.58 | 6.94 | 5.90 | 5.33 | 4.76 |
| $s = 2$ | 5.30 | 4.44 | 3.97 | 3.50 | 7.74 | 6.53 | 5.87 | 5.22 | 13.92 | 11.82 | 10.69 | 9.56 |

| $n = 4$ | $s = 0$ | 2.51 | 2.06 | 1.81 | 1.56 | 3.66 | 3.01 | 2.66 | 2.31 | 6.60 | 5.45 | 4.83 | 4.21 |
| $s = 1$ | 2.55 | 2.10 | 1.86 | 1.62 | 3.79 | 3.17 | 2.84 | 2.51 | 6.91 | 5.85 | 5.27 | 4.70 |
The results are summarized in table 4, which provides a compact summary of all three benchmarks that relate to the values of the parameters calibrated above. From 1990 onward, a representative family under pooling comprises two surviving adults and 3.49 children in the absence of AIDS and two surviving adults and 8.87 children in its presence.

BENCHMARK 2: POOLING, AIDS, AND NO INTERVENTION. If, following the full onset of the AIDS epidemic, premature adult mortality remains at the level that

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda$</th>
<th>$e$</th>
<th>$y(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>2.62</td>
<td>0.50</td>
<td>19,500</td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.64</td>
<td>22,340</td>
</tr>
<tr>
<td>2020</td>
<td>4.32</td>
<td>0.97</td>
<td>29,590</td>
</tr>
<tr>
<td>2050</td>
<td>7.86</td>
<td>1.00</td>
<td>53,720</td>
</tr>
<tr>
<td>2080</td>
<td>13.85</td>
<td>1.00</td>
<td>94,720</td>
</tr>
</tbody>
</table>

Note: All results are based on $30\%_{20}$.

$\kappa_{120} = \kappa_{90}$: in 1990, households formed expectations about adult mortality in 2020, when their children will have reached adulthood, that reflected the actual course of the epidemic over the period 1990–2020, as set out in table 3; $\kappa_{150} = \kappa_{90}$ is analogously defined when such expectations are revised starting only in 2020.
yields the steady-state probabilities in table 2, the consequences of doing nothing will be nothing short of disastrous, as seen in figure 2. Within a few generations, the epidemic sets in train a complete collapse of both the economy and, almost surely, the social institution of pooling. The extremely high level of premature mortality among adults leaves the community relatively impoverished from the start and with an intolerable burden of dependency: each surviving couple has to care for almost two adopted children for each one of their own. Education is correspondingly neglected, with unrelieved child labor \((e = 0)\) for the generation born starting in 2020. The descent into backwardness \((\hat{\lambda} = 1)\) is complete by 2050, when family income is a little less than two-thirds its level in 1960, and there are almost twice as many children for each couple to care for. The results are summarized in table 4.

It might be argued that both variants with AIDS in table 4 constitute unduly pessimistic estimates of the conditions prevailing in 1990–2020 and beyond in terms of the level of mortality and the growth of long-term productivity. The sensitivity analysis in section II covers this possibility, but those findings are expanded on here.

Suppose, for example, that from 1990 onward, \(\kappa\) were to fall, not to 0.338 as above, but less precipitously, to 0.6, say. Since \(\lambda^* (0, 0.6) < \lambda_{90}\) (table 3), no collapse follows; but the system teters on the brink, with virtual stagnation thereafter (figure 2). Turning to the growth of productivity over the long run, prudent economic management and social integration after 1990 ought to yield an improvement over \(z = 0.818\). Suppose, then, that \(z = 1\), which corresponds to a doubling of \(\hat{\lambda}\) every generation (or 2.31 percent a year) under full-time schooling. If \(\kappa\) continues at 0.338, however, the collapse that ensues is scarcely less dramatic than that when \(z = 0.818\) (figure 2).

**BENCHMARK 3: POOLING, AIDS, AND DELAYED EXPECTATIONS.** The second variant in table 4 reflects the possibility that households will take some time to revise their expectations. Suppose this revision does not occur until the very start of the next generation, when the childhood experience of parental death will be vivid in the minds of the next cohort of young adults: their firm expectations are \(\kappa_{150} = \kappa_{90}\). Suppose, further, that these expectations are realized and that this scale of mortality persists into the future. The happy—but false—expectations about future mortality that are formed in 1990, coupled with what is assumed to be the generous altruism of full pooling, induce adults to invest heavily in the children's education, despite the sharp reductions in available resources caused by the outbreak of the epidemic. Yet, although the adults in the generation starting out in 2020 are every bit as well endowed with human capital as they would have been in the absence of the epidemic, their expectations concerning their children's future are so bleak as to induce them to roll back investment in schooling to levels not seen since the mid-20th century. The result is to send the entire system into a progressive decline.
As reported in table 4, income per capita in benchmark 3 peaks in the period starting in 2020, and two generations later, the fresh cohort of adults will be scarcely more productive than their forebears in 1960. Only a revival of optimism about the future and the resumption of low levels of premature adult mortality to confirm it will stave off a complete collapse.\(^5\) Note that a collapse is possible even when the mortality shock affects only one 30-year generation, depending upon how and when expectations are formed.

### III. Policy Options

All policies are assumed to be financed by lump-sum taxes. Furthermore, the government chooses the level of public expenditure not to optimize a classically specified intertemporal welfare function over an infinite horizon—a problem that is almost impossible to solve in the framework—but to restore steady-state growth and then maintain it. The policy program takes the form of a sequence of taxes and expenditures that achieves this objective, if it is at all feasible.

*Health Policy*

Health policy takes the form of spending on measures to combat the disease. For some diseases, treatment may result in a complete cure. There is no such prospect for the victims of AIDS, but the treatment of opportunistic infections in the later stages and the use of antiretroviral therapies can prolong life and maintain productivity. In the present overlapping generations setting, therefore, treatment may be thought of as reducing premature adult mortality in the probabilistic sense.

It remains to establish the relationship between the state probabilities and spending on combating the disease. This is accomplished by choosing a functional form for the relationship between the probability of premature death among adults, \(q\), and the level of expenditures on combating the disease, \(\eta\), and then making the simplifying assumption that the incidence of the disease is independently and identically distributed. For simplicity, and erring on the side of optimism, it is also assumed that such aggregate expenditures produce a pure public good, so that

\[
q(D = 1) = q(\eta; D = 1)
\]

where \(q(\eta; D = 1)\) is to be interpreted as the efficiency frontier of the set of all measures that can be undertaken to reduce \(q\) in the presence of the disease.

Very little is known about the exact shape of the function \(q(\cdot)\), but \(q(0; D = 1)\) should yield the estimates in table 2. A second, plausible, condition

\(^5\) The fact that false expectations can be helpful in overcoming shocks raises delicate questions about the value of transparency in public policy in this context. They are avoided here.
is that arbitrarily large spending on combating the epidemic should lead to the restoration of the status quo ante, that is, \( q(\infty; D = 1) = q(D = 0) \). For reasons that will become clear shortly, it is desirable to choose a functional form that not only possesses an asymptote but also allows sufficient curvature over some relevant interval of \( \eta \), so that the natural choice falls on the logistic:

\[
q(\eta; D = 1) = d - \frac{1}{a + ce^{-b\eta}}
\]

Hence,

\[
q(0; D = 1) = d - \frac{1}{a + c}
\]

and

\[
q(\infty; D = 1) = d - \frac{1}{a} = q(D = 0).
\]

The full estimation of the function \( q(\cdot) \) is described in appendix A2. The procedure yields the values of the parameters \( a, b, c, \) and \( d \) for men, women, and both combined, which are set out in table A2-1 for two values of the cost of saving a disability-adjusted life year. The associated functions \( q(\eta; D = 1) \) are convex to the origin and have relatively strong curvature over the interval \( \eta \in (300, 700) \) (figure A2-1). The said values depend on the annual cost \( (K) \) of a course of generic drugs. Marseille, Hoffmann, and Kahn (2002) set \( K \) at $395. Early in 2006, however, the annual cost of a course of generic drugs was about $200, so some might regard the first estimate as too conservative in terms of the cost-effectiveness of treatment, as opposed to prevention—even though it bears stating that neither estimate makes any allowance for the other components of highly active antiretroviral therapy and the threat that drug-resistant strains will proliferate when the full regime is not rigorously followed. The subsections that follow begin with the results based on the calibrated values of the parameters and the health-cost factor \( K = 395 \). The robustness of these findings to changes in all these parameters are then examined.

Policy Option 1: Spending on Health Under Pooling

The results of spending on health under pooling are qualitatively striking (table 5). The optimal level of spending on combating the epidemic immediately upon outbreak in 1990 \(( t = 0 )\) is $963, which is about 4.5 percent of GDP, rising to $1,029, or 3.6 percent of GDP, in 2020, when productivity is 30 percent higher. Fiscally speaking, this is a tall order and a very substantial long-run burden, especially in view of the fact that the additional taxes are assumed to be raised in lump-sum form. If this program is politically feasible, it will eventually yield

---

6. Under the distortionary tax systems that rule in practice, the marginal cost of a unit of public revenue can range between 1.3 and 1.7 or higher still.
steady-state growth, with full and universal education attained in 2050. With (optimal) spending at this level, premature mortality among adults would be scarcely higher than in the complete absence of the disease.

A comparison with benchmark 1 reveals that the costs of dealing with AIDS in terms of lost output are modest at first but become quite large by 2080, when productivity is about 88 percent of its benchmark level, even with the optimal package of interventions under the favorable conditions of the case considered here (table 4). The long-run rate of growth is unaffected by AIDS under this policy program; for once full-time schooling is reached, the growth rate depends only on \( z(0, \kappa) \), which is assumed to be constant at \( z(0, 0.86) = 0.818 \). Taking a somewhat broader view, therefore, the outcome is encouraging, in that the general character of benchmark 1 is still attainable (figure 3), including a relatively low level of premature adult mortality. Thus, the maintained assumption that pooling will survive the shock is arguably validated.

Given this rather encouraging qualitative finding, it is still natural to ask whether lower costs of generic drugs will yield significant quantitative gains. When \( K = 200 \), the optimal values of \( \eta \) in 1990 and 2020 and thereafter are substantially lower at $714 and $755, respectively, but the corresponding values of \( \kappa \) still rise, to 0.854 and 0.855, respectively. With a much lower fiscal burden and a slight improvement in premature adult mortality, \( \lambda \) increases a little more rapidly than when \( K = 395 \), so that its level in 2080 is almost 4 percent higher.

**Policy Option 2: Nuclear Families, Lump-Sum Subsidies**

The results under policy option 1 are predicated on the assumption that the government acts at once to nip the epidemic in the bud. In fact, the epidemic had assumed alarming proportions by 2000, with many children already left as

---

**Table 5. Policy Option 1: Spending on Combating the Disease under Pooling**

<table>
<thead>
<tr>
<th>Year (1960)</th>
<th>( \lambda ) (2.62)</th>
<th>( e ) (0.50)</th>
<th>( \eta ) (0)</th>
<th>( \kappa ) (0.860)</th>
<th>( n/\kappa ) (4.65)</th>
<th>( y(0) ) (19,503)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K = 395 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.60</td>
<td>963</td>
<td>0.849</td>
<td>3.53</td>
<td>22,445</td>
</tr>
<tr>
<td>2020</td>
<td>4.10</td>
<td>0.87</td>
<td>1,029</td>
<td>0.852</td>
<td>3.52</td>
<td>28,365</td>
</tr>
<tr>
<td>2050</td>
<td>6.83</td>
<td>1.00</td>
<td>1,029</td>
<td>0.852</td>
<td>3.52</td>
<td>46,725</td>
</tr>
<tr>
<td>2080</td>
<td>12.18</td>
<td>1.00</td>
<td>1,029</td>
<td>0.852</td>
<td>3.52</td>
<td>83,269</td>
</tr>
<tr>
<td>( K = 200 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.62</td>
<td>714</td>
<td>0.854</td>
<td>3.51</td>
<td>22,412</td>
</tr>
<tr>
<td>2020</td>
<td>4.16</td>
<td>0.90</td>
<td>755</td>
<td>0.855</td>
<td>3.51</td>
<td>28,718</td>
</tr>
<tr>
<td>2050</td>
<td>7.12</td>
<td>1.00</td>
<td>755</td>
<td>0.855</td>
<td>3.51</td>
<td>48,713</td>
</tr>
<tr>
<td>2080</td>
<td>12.65</td>
<td>1.00</td>
<td>755</td>
<td>0.855</td>
<td>3.51</td>
<td>86,521</td>
</tr>
</tbody>
</table>

\( \eta \) is the per household level of spending on combating the disease; \( y(0) \) is the level of income accruing to each pair of surviving adults and the children in their care.
orphans and even more destined to become orphans, thus calling into question the whole system of pooling. If this social institution does break down, leaving tightly defined nuclear families to emerge instead, then the government will face the challenging task not only of averting a collapse, but also of preserving equality within each generation. To make both possible, additional assumptions are needed about the formation of human capital when children are left as half or full orphans. Under the assumption that $z(1) = z(2)/2$, that is, single parents can do just as well as couples in raising their children if they have the income?it is possible to preserve equality of educational outcomes among all children with at least one living parent by subsidizing one-parent families so as to induce them to choose the same level of education that two-parent families choose. By hypothesis, no family takes in full orphans, so that these children must be cared for in orphanages. It is assumed that these institutions, when properly staffed and run, substitute perfectly for parents, at least where the formation of human capital is concerned. The operating rule is that each full orphan also enjoys the same level of consumption as a child in a single-parent household.

When the family structure is nuclear, a good policy program to overcome the shock caused by AIDS must ensure a substantial tax base, not only in the present but also in the next generation. The instruments available for this purpose are taxes on two-parent households, spending on combating the disease, the size of the subsidy to single-parent households, and the proportions of half and full orphans to be supported. They are chosen subject to the above restrictions.
designed to preserve equality, if at all possible, and to the government's budget constraint.

Given the complexity of using full-scale forward induction, a somewhat simpler approach is chosen. The aim is to maximize the expected size of the tax base in the next period, where all parties hold the firm expectation that there will be a continuation of the level of premature adult mortality (and hence of $\eta$) prevailing in the present. That is, stationary expectations are assumed, which permit the maximization problem to be written so that it effectively contains no variables or parameters pertaining to the future. In particular, families' decisions about education depend on $\kappa_{t+1}$ but under stationary expectations, $\kappa_{t+1} = \kappa_t$. The (bounded) rationality of these expectations is secured by imposing the condition that $\eta$ does not fall from one period to the next, for this will rule out policy programs under which the value of investments in education will be reduced ex post by failures to take adequate measures against the disease in the next period. It should be emphasized that if it is possible to stave off a collapse of the economy through a policy program derived on the basis of stationary expectations so formulated, then it certainly will be possible to do even better by using the full apparatus of forward induction. Since all adults possess at least one unit of human capital, the tax base is defined, for present purposes, as the excess of the aggregate level of human capital over the aggregate level when all adults have but one unit.

The optimum sequence\(^7\) yields a continuation of growth with complete equality—all orphans receive the support needed to bring them up to par with the children of two-parent households in each and every period (table 6). Growth is distinctly sluggish, however, which points to a collapse that would be somewhat narrowly averted. The (uniform) years of schooling rise noticeably more slowly across succeeding generations than under pooling, with full-time schooling achieved only in 2080, when the level of productivity is only slightly more than double its value in 1990. Spending on combating the disease is also higher in absolute terms throughout and, combined with the transfers required to support needy children, $g(2)$, this yields a much heavier fiscal burden than under pooling. Two-parent households pay a little over 20 percent of their income in the form of a lump-sum tax ($\tau$) to finance this program in 1990 and receive very little relief until rapid growth begins from 2080 onward, when one-parent families need less support.

The differences between policy options 1 and 2 call for some explanation. Under pooling, which ensures equality, the objective is to maximize the (uniform) level of productive efficiency ($\bar{z}$) in the next generation, whereas with nuclear families, it is the size of the future tax base that matters when the government has to undertake the task of replacing the institution of pooling with subsidies and orphanages. In the latter arrangement, it may be worthwhile to trade off educational attainment to secure more surviving adults at

---

7. The optimization problem is set out in full in Bell, Devarajan, and Gersbach (2003).
TABLE 6. Policy Option 2: Nuclear Families, Lump-Sum Subsidies

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda$</th>
<th>$\epsilon(1)$</th>
<th>$\epsilon(2)$</th>
<th>$\eta$</th>
<th>$g(2)$</th>
<th>$\delta(2)$</th>
<th>$\delta(4)$</th>
<th>$\kappa$</th>
<th>$y(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 395$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.49</td>
<td>0.49</td>
<td>1,101</td>
<td>2,174</td>
<td>4,223</td>
<td>1.0</td>
<td>1.0</td>
<td>0.854</td>
</tr>
<tr>
<td>2020</td>
<td>3.51</td>
<td>0.58</td>
<td>0.58</td>
<td>1,127</td>
<td>2,470</td>
<td>4,607</td>
<td>1.0</td>
<td>1.0</td>
<td>0.854</td>
</tr>
<tr>
<td>2050</td>
<td>4.36</td>
<td>0.79</td>
<td>0.79</td>
<td>1,179</td>
<td>3,143</td>
<td>5,466</td>
<td>1.0</td>
<td>1.0</td>
<td>0.854</td>
</tr>
<tr>
<td>2080</td>
<td>6.60</td>
<td>1.00</td>
<td>1.00</td>
<td>1,179</td>
<td>2,214</td>
<td>4,707</td>
<td>1.0</td>
<td>1.0</td>
<td>0.854</td>
</tr>
<tr>
<td>$K = 200$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.50</td>
<td>0.50</td>
<td>796</td>
<td>2,294</td>
<td>3,862</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>2020</td>
<td>3.59</td>
<td>0.62</td>
<td>0.62</td>
<td>814</td>
<td>2,654</td>
<td>4,309</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>2050</td>
<td>4.62</td>
<td>0.86</td>
<td>0.86</td>
<td>850</td>
<td>3,491</td>
<td>5,335</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
</tr>
<tr>
<td>2080</td>
<td>7.53</td>
<td>1.00</td>
<td>1.00</td>
<td>850</td>
<td>1,030</td>
<td>3,027</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Note: $g(2)$ is the income transfer to each one-parent family that receives such support; $\tau$ is the level of the special lump-sum tax on each two-parent family; $\delta(s)$, $s = 2,4$ is the fraction of all children in family state $s$ receiving public support; $y(1)$ is the level of gross income accruing to a two-parent family.

The later date. That is exactly what has happened here: the absolute level of $\eta$ is 14 percent higher than under pooling in both 1990 and 2050, despite the fact that productivity under pooling is 57 percent higher in the latter period. The other contributing factor arises from the fact that raising children in orphanages draws some adults out of the production of the aggregate private good—a cost that does not arise (by assumption) under pooling. The upshot is that families have less disposable income than under pooling, so that their children receive fewer years of schooling and growth is much slower. As under pooling, the long-run rate of growth is unaffected by AIDS in this fairly good sequence; but the traverse to steady-state growth is a painfully long one.

How much less painful would this trek be when $K = 200$? As under pooling, the optimal levels of $\eta$ are just over 25 percent lower than when $K = 395$, and $\kappa$ edges up further, almost to what its level would be in the absence of the epidemic. The absolute tax burden on two-parent families is also somewhat lighter: 8.6 percent lower in 1990, 7.5 percent in 2020, 2.6 percent in 2050, and almost 36 percent lower in 2080, when single-parent families need much less income support to be induced to choose full-time schooling. The effects on the accumulation of human capital are small at first, but by 2080, $\lambda$ is 14 percent higher than when $K = 395$. Since full equality in terms of human capital within each generation emerges as part of the optimal program, this faster pace requires that one-parent families need more generous support up to 2080, and $g(2)$ is correspondingly more generous—5.5 percent higher in 1990, 7.5 percent in 2020, and 11 percent in 2050.
Policy Option 3: Nuclear Families, School Attendance Subsidies

The results for this option are qualitatively similar to those under policy option 2, but growth is considerably more rapid (table 7; figures 3 and 4). Given the efficiency of school attendance subsidies relative to lump-sum transfers (and hence the lower taxes on two-parent households), one would expect a swifter attainment of full-time schooling in this variant, and this is indeed the case here. The precise reasoning runs as follows. Choose the optimal levels of taxes on two-parent households and spending on health under policy option 2. This program will yield the same demographic structure, the same level of education among such families, and the same total tax revenues. The outlays under policy option 3 needed to induce the same level of education among the children of one-parent households, however, will be smaller than under policy option 2. These children will also have a lower level of consumption, a standard to which full orphans are tethered. It follows that there will be an excess of total revenue over expenditures. Let $\eta$ be held constant, so as to keep the demographic structure unchanged, and let the taxes on two-parent households be reduced slightly, which will induce a small rise in $e(1)$. By continuity, there will still be enough funds to finance the additional subsidies to half and full orphans that will be needed to preserve equality in education, and hence in human capital in the next generation of adults. It follows that policy option 3 strictly dominates option 2 in all periods from $t = 0$ onward.

Full education is reached in 2050, as is the case under pooling, although productivity is 12 percent lower, because of the accumulated effects of lower attainments in the two preceding generations. Spending on measures to combat the disease is a little higher than under pooling at first but a little less from 2020 onward. It is about 13 percent lower than its counterpart under policy option 2 throughout, so that more premature deaths are implicitly accepted, although the differences in $K_t$ are small. A measure of the comparative efficiency of conditional educational subsidies is that satisfactory growth is

### Table 7. Policy Option 3: Nuclear Families, School Attendance Subsidies

<table>
<thead>
<tr>
<th>Year</th>
<th>$\lambda$</th>
<th>$e(1)$</th>
<th>$e(2)$</th>
<th>$\eta$</th>
<th>$g(2)$</th>
<th>$\tau$</th>
<th>$\delta(2)$</th>
<th>$\delta(4)$</th>
<th>$\kappa$</th>
<th>$y(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.57</td>
<td>0.57</td>
<td>973</td>
<td>280</td>
<td>1,886</td>
<td>1.0</td>
<td>1.0</td>
<td>0.850</td>
<td>22,337</td>
</tr>
<tr>
<td>2020</td>
<td>3.91</td>
<td>0.78</td>
<td>0.78</td>
<td>1,022</td>
<td>289</td>
<td>2,101</td>
<td>1.0</td>
<td>1.0</td>
<td>0.852</td>
<td>27,220</td>
</tr>
<tr>
<td>2050</td>
<td>5.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1,022</td>
<td>197</td>
<td>2,310</td>
<td>1.0</td>
<td>1.0</td>
<td>0.852</td>
<td>40,960</td>
</tr>
<tr>
<td>2080</td>
<td>10.80</td>
<td>1.00</td>
<td>1.00</td>
<td>1,022</td>
<td>0</td>
<td>2,642</td>
<td>1.0</td>
<td>1.0</td>
<td>0.852</td>
<td>73,839</td>
</tr>
<tr>
<td>$K = 395$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.58</td>
<td>0.58</td>
<td>716</td>
<td>287</td>
<td>1,502</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
<td>22,342</td>
</tr>
<tr>
<td>2020</td>
<td>4.00</td>
<td>0.82</td>
<td>0.82</td>
<td>748</td>
<td>295</td>
<td>1,713</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
<td>27,723</td>
</tr>
<tr>
<td>2050</td>
<td>6.36</td>
<td>1.00</td>
<td>1.00</td>
<td>748</td>
<td>0</td>
<td>1,702</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
<td>43,515</td>
</tr>
<tr>
<td>2080</td>
<td>11.41</td>
<td>1.00</td>
<td>1.00</td>
<td>748</td>
<td>0</td>
<td>2,256</td>
<td>1.0</td>
<td>1.0</td>
<td>0.86</td>
<td>78,019</td>
</tr>
</tbody>
</table>
achieved with amounts paid to one-parent households that are barely a tenth of the lump-sum transfers made under policy option 2.

The tax burden on two-parent households is correspondingly lighter: the absolute payment per household is a little less than one-half of that under policy option 2 in 1990, rising to 56 percent in 2080. The difference in productivities is very large in 2080: namely $2^\gamma(1)$ to one, or 1.636, which implies a much lower relative tax burden. The latter falls from about 8.6 percent of income in 1990 to 3.6 percent in 2080 under policy option 3, and from 19.3 percent to 10.4 percent under policy option 2. Observe that although the payment of school attendance subsidies ends from 2080 onward, $\tau$ is higher than in 2050. The reason is that the raising and caring for full orphans require the time and effort of adults specifically employed for this purpose, the costs of which rise with $\lambda$.

An Optimistic Variant

This analysis of alternative policies concludes with brighter assumptions about the long-term rate of growth of productivity and the costs of antiretroviral drugs. The results for $\gamma = 1$ and $K = 200$, perhaps the most plausible of optimistic constellations, are reported in tables 8–10. Table 8 sets out the first benchmark ($D = 0$). In the absence of the epidemic, the increase in $\gamma$ from 0.818 to 1 makes for large gains indeed over three generations. Full schooling is achieved in 2020, and $\lambda$ in 2080 is almost 75 percent higher than in the reference case. Turning to policy interventions in the face of the epidemic, the optimal value of $\eta$ is not
Table 8. An Optimistic Variant (z = 1, K = 200): The First Benchmark (D = 0)

<table>
<thead>
<tr>
<th>Year</th>
<th>( \lambda )</th>
<th>( e )</th>
<th>( \kappa )</th>
<th>( n/\kappa )</th>
<th>( y(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>2.62</td>
<td>0.50</td>
<td>0.86</td>
<td>4.65</td>
<td>19,503</td>
</tr>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.68</td>
<td>0.86</td>
<td>3.49</td>
<td>22,257</td>
</tr>
<tr>
<td>2020</td>
<td>5.26</td>
<td>1.00</td>
<td>0.86</td>
<td>3.49</td>
<td>35,967</td>
</tr>
<tr>
<td>2050</td>
<td>11.52</td>
<td>1.00</td>
<td>0.86</td>
<td>3.49</td>
<td>78,771</td>
</tr>
<tr>
<td>2080</td>
<td>24.04</td>
<td>1.00</td>
<td>0.86</td>
<td>3.49</td>
<td>164,380</td>
</tr>
</tbody>
</table>

Table 9. Policy Option 1: Spending on Combating the Disease under Pooling (z = 1, K = 200)

<table>
<thead>
<tr>
<th>Year</th>
<th>( \lambda )</th>
<th>( e )</th>
<th>( \eta )</th>
<th>( \kappa )</th>
<th>( n/\kappa )</th>
<th>( y(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>3.14</td>
<td>0.65</td>
<td>714</td>
<td>0.85</td>
<td>3.51</td>
<td>22,334</td>
</tr>
<tr>
<td>2020</td>
<td>5.07</td>
<td>1.00</td>
<td>714</td>
<td>0.85</td>
<td>3.51</td>
<td>34,690</td>
</tr>
<tr>
<td>2050</td>
<td>11.15</td>
<td>1.00</td>
<td>714</td>
<td>0.85</td>
<td>3.51</td>
<td>76,218</td>
</tr>
<tr>
<td>2080</td>
<td>23.29</td>
<td>1.00</td>
<td>714</td>
<td>0.85</td>
<td>3.51</td>
<td>159,274</td>
</tr>
</tbody>
</table>

Table 10. Policy Options 2 and 3: Nuclear Families, Lump-Sum and School Attendance Subsidies (z = 1, K = 200)

| Year | \( \lambda \) | \( e(1) \) | \( e(2) \) | \( \eta \) | \( g(2) \) | \( \tau \) | \( \delta(2) \) | \( \delta(4) \) | \( \kappa \) | \( y(1) \) |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Policy option 2 | | | | | | | | | | | |
| 1990 | 3.14 | 0.54 | 0.54 | 796 | 2,294 | 3,863 | 1.0 | 1.0 | 0.86 | 22,439 |
| 2020 | 4.37 | 0.83 | 0.83 | 842 | 3,287 | 5,087 | 1.0 | 1.0 | 0.86 | 30,248 |
| 2050 | 8.25 | 1.00 | 1.00 | 842 | 3 | 2,033 | 1.0 | 1.0 | 0.86 | 56,400 |
| 2080 | 11.50 | 1.00 | 1.00 | 842 | 0 | 2,979 | 1.0 | 1.0 | 0.86 | 119,637 |
| Policy option 3 | | | | | | | | | | | |
| 1990 | 3.14 | 0.62 | 0.62 | 716 | 287 | 1,513 | 1 | 1 | 0.85 | 22,276 |
| 2020 | 4.87 | 1.00 | 1.00 | 716 | 0 | 1,507 | 1 | 1 | 0.85 | 33,319 |
| 2050 | 10.75 | 1.00 | 1.00 | 716 | 0 | 2,164 | 1 | 1 | 0.85 | 73,476 |
| 2080 | 22.49 | 1.00 | 1.00 | 716 | 0 | 3,477 | 1 | 1 | 0.85 | 153,790 |

affected under pooling, and \( \lambda \) in 2080 is a mere 3 percent smaller than the corresponding benchmark value. With nuclear families, the pace is also distinctly quicker, even under policy option 2. A more efficient educational technology will do much to ease the task of maintaining growth and welfare in the face of the epidemic, but it will not necessarily stave off a collapse in the absence of any other intervention (table 3).
The AIDS epidemic will peak far in advance of the economic damage it will ultimately cause. In southern Africa, where prevalence rates among people aged 15-49 years are already 20 percent and higher, the worst is still to come. The scale of that damage, in terms of accumulated losses in GDP per capita, will also be large even if the measures designed to combat the disease and to ensure the education of orphans are well chosen, and the fiscal means employed to finance them are highly efficient. Without such measures, and given a continuation of high levels of mortality, economic collapse is a very real danger.

The main reason for these gloomy findings lies in the peculiarly insidious and selective character of the disease. By killing mostly young adults, AIDS does more than destroy the human capital embodied in them; it deprives their children of the very things they need to become economically productive adults—their parents’ loving care, knowledge, and capacity to finance education. This weakening of the mechanism through which human capital is transmitted and accumulated across generations becomes apparent only after a long lag, and it is progressively cumulative in its effects. Therein lies the source of the difference between the findings in this article and those of many previous studies, which have focused either on the role of quasi-fixed factors over the medium run or on the historical record to date.

What are the lessons for public policy? Where the prevalence rate is still low, as in much of Asia, Eastern Europe, the Middle East, and Latin America, it is of the utmost importance to contain the disease at once: for the economic system as well as for individuals, an ounce of prevention is worth more than a pound of cure.

Where the epidemic is more advanced, combating the disease and its economic effects successfully will require a large and determined fiscal effort, the correct design of which is a complicated matter. Intuitively, the question is: What combination of measures should be adopted to promote the formation of human capital and good health when the threat of a collapse looms? These measures are partly complementary. Maintaining good health means that the human capital embodied in individuals during childhood and training will survive and pay off into old age, not only for them but also for their children. When public funds are very scarce, however, some tradeoffs will be unavoidable, requiring the concentration of resources on some programs or groups at the expense of others. The hope here is that the knowledge about what works in the fields of child rearing, education, the care of orphans, health, and so forth can reveal how to formulate combined programs of interventions that will ward off the threat of an economic collapse. The true social rate of return to such programs can be extremely high, whereas that derived from calculations based on standard (local) cost–benefit analysis may be quite modest. Fiscal policy in general, and policy in the social sectors in particular, must be formulated with a clear eye on its contribution to solving the
long-run economic problem posed by AIDS. For in the event of a collapse of productivity, little else will matter.

These points are vividly illustrated by the results for South Africa. In the absence of the epidemic, there would have been the prospect of modest, but accelerating growth of per capita income. An unabated continuation of the epidemic could bring about a progressive collapse. With the right interventions, this fate can be averted, although the costs are high, even under favorable social arrangements for the care of orphans. If those arrangements break down, growth is likely to be rather sluggish. These conclusions must be regarded as preliminary, and various aspects of the analysis need further work and refinement. That much conceded, the sensitivity analysis nevertheless suggests that these findings are robust to changes in a variety of key assumptions and parameter values. And it would be unconscionable to err on the side of optimism.

APPENDIX A1. MICROFOUNDERATION OF THE MODEL

This appendix presents the microfoundation of the dynamical system. It does not attempt to model individuals' sexual behavior, with all its gratifications and risks. While the said probabilities are therefore exogenous, their values in the application to South Africa are based on the epidemiological work of Dorrington and others (2001).

The first topic is the formation of human capital. Consider a family at the start of period \( t \). Let \( \lambda^f_t \) be the father's endowment of human capital, \( \lambda^m_t \) the mother's, and \( A_t(s_t) \) their total human capital when the family is revealed to be in state \( s_t \). Then,

\[
(A1.1) \quad A_t(1) = \lambda^f_t + \lambda^m_t, A_t(2) = \lambda^m_t, A_t(3) = \lambda^f_t, A_t(4) = 0.
\]

Assuming that there is assortative mating, \( \lambda^f_t = \lambda^m_t \), equation (A1.1) specializes to

\[
(A1.2) \quad A_t(1) = 2\lambda_t, A_t(2) = A_t(3) = \lambda_t, A_t(4) = 0,
\]

where the superscripts \( f \) and \( m \) may now be dropped without introducing ambiguity.

Human capital is assumed to be formed by a process of child-rearing combined with formal education. In the course of rearing their children, parents give them a certain capacity to build human capital for adulthood, a capacity that is itself increasing in the parents' own human capital. This gift will be of little use, however, unless it is complemented by at least some formal education, in the course of which the basic skills of reading, writing, and calculating can be learned. Let the proportion of childhood devoted to education be denoted by \( e_t \in [0, 1] \), the residual being allocated to work, and for simplicity, let all the
children in a family be treated in the same way. The human capital attained by each of the children on reaching adulthood is given by

\[
\text{(A1.3)} \quad \lambda_{t+1} = \begin{cases} 
z(s_t)f(e_t)A_t(s_t) + 1, & s_t = 1, 2, 3 \\
\xi, & s_t = 4
\end{cases}
\]

Beginning with the upper branch of equation (A1.3), the term \(z_t(s_t)\) represents the strength with which capacity is transmitted across generations. For simplicity, the father's and mother's contributions are assumed to be perfect substitutes: \(z(2) = z(3)\). It is also assumed that where transmitting this capacity is concerned, two parents can rear a child at least as well as one, but, in view of perfect substitutability, no better than twice as well as one. Hence, recalling equation (A1.2),

\[
\text{(A1.4)} \quad z(2) = z(3) \geq z(1) \geq z(2)/2 = z(3)/2.
\]

Thus, the upper branch of equation (A1.3) can be rewritten as

\[
\text{(A1.5)} \quad \lambda_{t+1} = (3 - s_t)z(s_t)f(e_t)\lambda_t + 1, \quad s_t = 1, 2
\]

with both types of single-parent families being identical in this respect. The function \(f(\cdot)\) represents educational technology — translating time spent on education into learning. It is assumed to be strictly increasing and differentiable, with \(f(0) = 0\). Observe that equation (A1.3) and \(f(0) = 0\) imply that children who do not attend school at all attain, as adults, only some basic level of human capital, which has been normalized to unity.

According to the lower branch of equation (A1.3), there is a miserable outcome for full orphans who do not enjoy the good fortune to be adopted or placed in (good) institutional care. Deprived of love and care, and left to their own devices, they go through childhood uneducated, to attain human capital \(\xi < (5\gamma)\) in adulthood.

The next step is to relate human capital to current output, which takes the form of an aggregate consumption good. Output is assumed to be proportional to inputs of labor measured in efficiency units. A natural normalization is that an adult who possesses human capital in the amount \(\lambda_t\) is endowed with \(\lambda_t\) efficiency units of labor, which he or she supplies completely inelastically. A child's efficiency will be somewhat lower than the parents', all other things being equal, on the grounds of age alone. To reflect these considerations, let a child supply \((1 - e_t)\gamma\) efficiency units of labor when the child works \(1 - e_t\) units of time. It is

8. This analysis skips the fact that girls often receive less education than boys. The ensuing inequality in human capital introduces analytical and empirical difficulties whose importance, for the purpose of this article, does not seem to warrant specific treatment.
plausible to assume that $\gamma \in (0, \xi)$, that is, a full-time working child is at most as productive as an adult who happened to be an uneducated orphan. A family with $n_t$ children therefore has a total income in state $s_t (s_t = 1, 2, 3)$ of

(A1.6) \[ y_t(s_t) = \alpha[A_t(s_t) + n_t(1 - e_t)\gamma] \]

where the scalar $\alpha (>0)$ denotes the productivity of human capital, measured in units of output per efficiency unit of labor input.

Household Behavior

All allocative decisions are assumed to lie in the parents' hands, as long as they are alive. Any bequests at death are ruled out, so that the whole of current income, as given by equation (A1.6), is consumed. Within the family, let the husband and wife enjoy equality as partners, and let each child obtain a fraction $\beta \in (0, 1)$ of an adult's consumption if at least one adult survives. Full orphans ($s_t = 4$) do not attend school and consume what they produce as child laborers.

Without any taxes or subsidies, the household's budget constraint may therefore be written as

(A1.7) \[ [(3 - s_t) + n_t\beta]c_t + \alpha n_t\gamma e_t \leq \alpha[(3 - s_t)\lambda_t + n_t\gamma], \quad s_t = 1, 2 \]

where $c_t$ is the level of each adult's consumption. The expression on the left hand side represents the costs of consumption and the opportunity costs of the children's schooling. The expression on the right hand side is the family's so-called full income in state $s_t = 1, 2, 3$. Observe that single-parent households not only have lower levels of full income than their otherwise identical two-parent counterparts, but they also face a higher relative price of education, defined as $\alpha n_t\gamma/[(3 - s_t) + n_t\beta]$.

Couples have children while they are young until some exogenously fixed number have survived infancy, a target that may vary from period to period. With $n_t$ thus fixed, the adults wait until the state of the family becomes known, and the survivor then chooses some feasible pair $(c, e)$ subject to condition (A1.7).

Parents are assumed to have preferences over their own current consumption and the human capital attained by their children in adulthood, taking into account the fact that investment in a child's education will be wholly wasted if that child dies prematurely in adulthood. Let mothers and fathers have identical preferences, and for two-parent households, let there be no joint aspect to the consumption of the pair $(c, e)$: each surviving adult derives (expected) utility

9. A household's full income is the scalar product of its endowment vector and the vector of market prices. Here, output is taken as the numéraire.

10. Although there is much evidence in favor of at least some replacement fertility, this is evidently a strong assumption. In the numerical application, however, variations in $n_t$ turn out to have only weak effects on the system’s dynamics (see table 4).
from the pair so chosen, and these utilities are then added up within the family. In effect, whereas \( c_t \) is a private good, the human capital of the children in adulthood is a public good within the marriage.

Since all the children attain \( \lambda_{t+1} \), the only form of uncertainty is that surrounding the number who will not die prematurely as adults, which is denoted by the random variable \( a_{t+1} \). Let preferences be separable, with representation

\[
EU_t(s_t) = (3 - s_t)[u(c_t) + (E_t a_{t+1}) v(\lambda_{t+1})], \quad s_t = 1, 2
\]

where the contribution \( v(\lambda_{t+1}) \) counts only when death does not come early, \( E_t \) is the expectation operator, and \( E_t a_{t+1} \) is the expected number of children surviving into old age. The subutility functions \( u(\cdot) \) and \( v(\cdot) \) are assumed to be increasing, continuous, concave, and twice-differentiable. Denoting by \( \kappa_t^e \), the parents' subjective probability that a child will survive to old age and recalling assumption \( I \) and that all children are treated identically yield

\[
E_t a_{t+1} v(\lambda_{t+1}) = n_t \kappa_t^e v(\lambda_{t+1})
\]

where \( \lambda_{t+1} \) is given by equation (A1.3). A reduction in \( \kappa_{t+1}^e \) therefore effectively entails a weaker taste for children's education. It will be convenient in what follows to rewrite equation (A1.8) as

\[
(A1.10) \quad EU_t(s_t) = (3 - s_t)[u(c_t) + n_t \kappa_t^e \cdot v[z(s_t) f(e_t) \lambda_t(s_t) + 1]], \quad s_t = 1, 2
\]

since both types of single-parent families are identical. Hence, it suffices to examine the states \( s_t = 1, 2 \). A family in state \( s_t = 1, 2 \) in period \( t \) solves the following problem:

\[
(A1.11) \quad \max_{[c_t(s_t), e_t(s_t)]} \quad EU_t(s_t) \quad \text{s.t.} \quad (A1.7), c_t \geq 0, e_t \in [0, 1].
\]

Let \([c^0_t(s_t), e^0_t(s_t)]\) solve problem (111), whose parameters are \((\alpha, \beta, \gamma, \kappa_t^e, \lambda_t, n_t)\). Using the envelope theorem yields

\[
\frac{\partial EU_t(s_t)}{\partial \lambda_t} > 0, \quad \frac{\partial EU_t(s_t)}{\partial \kappa_t^e} > 0.
\]

Since current consumption is maximized by choosing \( e_t = 0 \), it follows that the parents' altruism toward their children must be sufficiently strong if they are to choose \( e_t > 0 \).
If both goods are noninferior, it follows at once that
\[
\frac{\partial e^0_t(s_t)}{\partial \Lambda_t(s_t)} \geq 0, \quad \frac{\partial c^0_t(s_t)}{\partial \Lambda_t(s_t)} \geq 0.
\]

**Dynamics**

There are no insurance arrangements in the above account, so that premature adult mortality in period \(t\) will affect not only the level but also the distribution of human capital in period \(t + 1\). As noted above, full orphans will suffer low productivity in adulthood, as expressed in the lower branch of equation (A1.3). Such mortality also affects the distribution of families across states 1, 2, and 3 in period \(t\) and will thus affect the level and distribution of human capital in period \(t + 1\) if \(e^0_t(s_t)\) varies across states and with the severity of premature adult mortality, as it normally will when \(\lambda_t\) is not too large. These repercussions will then make themselves felt in future periods, even if premature adult mortality vanishes after period \(t\).

To state all this formally, recall that the family chooses \(e^0_t(s_t; \cdot)\) in light of its resources and expectations so as to solve problem (A1.11). Hence, equation (A1.3) may be written so as to make these influences explicit:

\[
(A1.12) \quad \lambda_{t+1} = \begin{cases} 
\left( z(s_t) f\left(e^0_t\left(\Lambda_t(s_t), s_t, \kappa^e_{t+1}\right)\right) \Lambda_t(s_t) + 1, \quad s_t = 1, 2, 3 \\
\xi, \quad s_t = 4
\end{cases}
\]

Equation (A1.12) describes a random dynamical system—random in the sense that each child in any given family state \(s_t\) can wind up in any of the states \(s_{t+1} \in \{1, 2, 3, 4\}\) after reaching adulthood and forming a family in period \(t + 1\).

**APPENDIX A2. THE FUNCTION q(·)**

With four parameters to be estimated, two additional independent conditions beyond those in the text are required. One way of proceeding is to pose the question: What is the marginal effect of efficient spending on \(q\) in high- and low-prevalence environments? That is to say, estimates are needed of the derivatives of \(q(\eta; D = 1)\) at \(\eta = 0\) and some value of \(\eta\) that corresponds to heavy spending, when the scope for exploiting cheap interventions has been exhausted. To obtain such estimates, we used the estimated costs of preventing a case of AIDS or saving a disability-adjusted life year by various methods, as reported by Marseille, Hofmann, and Kahn (2002).

When the prevalence rate is high, the authors argue, the most cost-efficient form of intervention is to target prostitutes for the specific purpose of controlling sexually transmitted diseases and promoting the use of condoms. The associated cost per AIDS case averted in Kenya is given as $8–12. It seems reasonable to infer that this cost recurs annually. Other preventive measures are less cost-effective by a factor of up to 10 or more. Marseille, Hofmann, and Kahn (2002) put the
average cost per disability-adjusted life year of a diverse bundle of such measures at $12.50. For these measures, the assumption that \( \eta \) produces a pure public good is not far off the mark. Now, a reduction in \( q \) of 0.01 over a span of 30 years yields 0.3 disability-adjusted life years. Allowing for the fact that there is substitution among diseases, that is, if one does not succumb to AIDS, there is always the threat of something else, the expenditure of another $12.50 when \( \eta \) is small will yield a net reduction in \( q(D = 1) \) of about 
\[
(0.01) \cdot (1 - q(D = 0)) = 0.028.
\]
Recalling that \( \eta \) is defined with reference to a population of adults whose measure has been normalized to unity and rounding up to $15, we have

\[
(A2.1) \quad \frac{dq(0; D = 1)}{d\eta} = - \frac{cb}{(a + c)^2} = - \frac{0.028}{15}.
\]

Following the purposive and determined implementation of the full battery of preventive measures, the remaining intervention is to treat the infected. There is now neither a cure nor the prospect of one for perhaps decades to come. Opportunistic infections can, of course, be treated in the later stages of the disease, and the onset of full-blown AIDS can be delayed for some years through the controlled use of antiretroviral therapies. Such measures will do little to reduce \( q \) as strictly defined, but by keeping infected individuals healthier and extending life a bit, they will raise lifetime income and improve the parental care enjoyed by children in affected families.

In the context of the model, therefore, it seems perfectly defensible to interpret these gains as equivalent to a reduction in \( q \). Marseille, Hofmann, and Kahn (2002) put the cost of saving a disability-adjusted life year by such means at $395, assuming that the drugs take the form of low-cost generics and explicitly neglecting the costs of the technical and human infrastructure needed to support an effective, so-called highly active antiretroviral therapy regimen of this kind. It is assumed here that the highly active antiretroviral therapy regimen is the efficient, marginal form of intervention when a low prevalence rate has resulted from a determined, extensive, and continuing effort at prevention. To complete the specification of this case, the level of aggregate spending at which highly active antiretroviral therapy becomes the best choice at the margin must be determined. Note that in the absence of diminishing returns to preventive measures, it would be possible to attain the status quo ante \( (D = 0) \) by spending

\[
(A2.2) \quad [q(0; D = 1) - q(D = 0)] \cdot (15/0.028) = 278.
\]

In fact, diminishing returns will set in as the prevalence rate falls. Where preventing mother-to-child transmission is concerned, for example, a drop in the prevalence rate from 30 percent to 15 percent will almost double the cost of
saving a disability-adjusted life year (Marseille, Hofmann, and Kahn 2002). Since 15 percent hardly counts as a low level of prevalence, it seems fairly safe to assume that highly active antiretroviral therapy will not become cost-efficient until spending on preventive measures, and the treatment of opportunistic infections is at least triple the above estimate. Thus the required fourth condition is:

(A2.3) \[ \frac{q'(0; D = 1)}{q'(835; D = 1)} = \frac{395}{15} \]

The four conditions (8), (9), (A2.1), and (A2.3) may be solved to yield the values of the parameters \(a, b, c,\) and \(d\) for men, women, and both combined, as set out in table A2-1. Premature adult mortality is defined precisely as death before age 50, conditional on surviving to age 20, the corresponding probability being denoted as \(q_{20}\).

By way of sensitivity analysis, to \(K\), the function \(q(\cdot)\) is respecified as follows:

(A2.4) \[ q(\eta; D = 1) = d - \frac{1}{a + ce^{-b\eta(K/395)}} \]

where \(K\) denotes the annual cost of a course of generic drugs. Equations (A2.1) and (A2.3) are then modified to read

(A2.5) \[ \frac{dq(0; D = 1)}{d\eta} = -\frac{cb(K/395)}{(a + c)^2} = \frac{0.028}{15} \]

and

(A2.6) \[ q(278 + (835 - 278)(K/395), K, D = 1) = q(835, K = 395, D = 1) \]

respectively. Setting \(K = 200\) yields the parameters reported in the lower half of Table A2-1 and the associated function in figure A2-1. The shift of \(q(\cdot)\) represents the favorable effects of the said reduction in costs on the government's possibilities of reducing premature adult mortality.

<table>
<thead>
<tr>
<th>Table A2-1. Parameters of the (q(\cdot)) Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K = 395)</td>
</tr>
<tr>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>(K = 200)</td>
</tr>
<tr>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>Average</td>
</tr>
</tbody>
</table>
**Figure A2-1.** Premature Adult Mortality $30q_{20}$ as a function of $\eta$ and $K$

**References**


Robust Multidimensional Spatial Poverty Comparisons in Ghana, Madagascar, and Uganda

Jean-Yves Duclos, David Sahn, and Stephen D. Younger

Spatial poverty comparisons are investigated in three African countries using multidimensional indicators of well-being. The work is analogous to the univariate stochastic dominance literature in that it seeks poverty orderings that are robust to the choice of multidimensional poverty lines and indices. In addition, the study seeks to ensure that the comparisons are robust to aggregation procedures for multiple welfare variables. In contrast to earlier work, the methodology applies equally well to what can be defined as "union," "intersection," and "intermediate" approaches to dealing with multidimensional indicators of well-being. Furthermore, unlike much of the stochastic dominance literature, this work computes the sampling distributions of the poverty estimators to perform statistical tests of the difference in poverty measures. The methods are applied to two measures of well-being, the log of household expenditures per capita and children's height-for-age z scores, using data from the 1988 Ghana Living Standards Study survey, the 1993 National Household Survey in Madagascar, and the 1999 National Household Survey in Uganda. Bivariate poverty comparisons are at odds with univariate comparisons in several interesting ways. Most important, it cannot always be concluded that poverty is lower in urban areas in one region compared with that in rural areas in another, even though univariate comparisons based on household expenditures per capita almost always lead to that conclusion.

It is common to assert that poverty is a multidimensional phenomenon, yet most empirical work on poverty, including spatial poverty, uses a unidimensional yardstick to judge a person's well-being, usually household expenditures or income per capita or per adult equivalent. When studies use more than one indicator of well-being, poverty comparisons are either made independently for
each indicator\(^1\) or made using an arbitrarily defined aggregation of the multiple indicators into a single index.\(^2\) In either case, aggregation across multiple welfare indicators and across the welfare statuses of individuals or households requires specific aggregation rules that are necessarily arbitrary.\(^3\) Multidimensional poverty comparisons also require the estimation of multidimensional poverty lines, a procedure that is problematic even in a unidimensional setting.

Taking as a starting point the conviction that multidimensional poverty comparisons are ethically and theoretically attractive, the purpose here is to apply quite general methods for multidimensional poverty comparisons to the particular question of spatial poverty in three African countries—Ghana, Madagascar, and Uganda. The relevant welfare theory and accompanying statistics are developed elsewhere (Duclos, Sahn, and Younger 2003). The purpose here is to give an intuitive explanation of the methods and to show that they are both tractable and useful when applied to spatial poverty in Africa.

The poverty comparisons use the dominance approach initially developed by Atkinson (1987) and Foster and Shorrocks (1988a, 1988b, 1988c) in a unidimensional context.\(^4\) In a review of this literature, Zheng (2000) distinguishes between poverty comparisons that are robust to the choice of a poverty line and those that are robust to the choice of a poverty measure or index. Both are attractive features of the dominance approach because they enable the analyst to avoid relying on ethically arbitrary choices of a poverty line and a poverty measure. The poverty comparisons used here are robust to the selection of both a poverty line and a poverty measure. In the multidimensional context, this includes robustness over the manner in which multiple indicators interact to generate overall individual well-being.

Section I briefly presents the data and provides an intuitive discussion of multidimensional poverty comparisons. In addition to the stochastic dominance conditions that are familiar from the univariate literature, it discusses two concepts that arise only in a multivariate context. First, it distinguishes between intersection and union definitions of poverty.\(^5\) By the well-known focus axiom used in poverty measurement (see, for instance, Foster 1984), these definitions

---

1. This would involve, say, comparing incomes across regions and then comparing mortality rates across regions and so on.
2. The best-known example is the human development index of the United Nations Development Programme (UNDP 1990), which uses a weighted average of life expectancy, literacy, and GDP per capita across the population.
3. Such rules have been the focus of some of the recent literature. See, for instance, Tsui (2002) and Bourguignon and Chakravarty (2003). Bourguignon and Chakravarty (2002) also give several interesting examples in which poverty orderings vary with the choice of aggregation rules.
5. For further recent discussion, see Bourguignon and Chakravarty (2002, 2003), Atkinson (2003), and Tsui (2002).
identify the individual poverty statuses to be aggregated to obtain poverty indices. If well-being is measured in the dimensions of income and height, say, then a person whose income falls below an income poverty line or whose height falls below a height poverty line could be considered poor. This is a union definition of multidimensional poverty. By an intersection definition, however, a person would have to fall below both poverty lines to be considered poor. In contrast to earlier work, the tests used here are valid for both definitions—or for any choice of intermediate definitions for which the poverty line in one dimension is a function of well-being measured in the other dimension.

A second key concept that arises only in the context of multivariate poverty comparisons is that, roughly speaking, the correlation between individual measures of well-being matters. If two populations have the same univariate distributions for two measures of well-being, but one has a higher correlation between these measures, then it should not have lower poverty. This is because a person’s deprivation in one dimension of well-being should matter more if the person is also poorer in the other dimension. The dimensions of well-being are substitutes in the poverty measure. While this is apparently intuitive, counterexamples are also presented, although the poverty comparisons are valid only for the case in which the dimensions are substitutes.

Section I concludes with examples of why the poverty comparisons here are more general than comparisons of indices such as the United Nations Development Programme’s human development index (UNDP 1990) and comparisons that consider each dimension of well-being independently.

Section II applies these methods to spatial poverty comparisons in Ghana, Madagascar, and Uganda, comparing poverty across regions and areas (urban and rural) in the dimensions of household expenditures per capita and nutritional status for children under the age of 5. Univariate comparisons based on expenditures or nutritional status alone almost always show greater poverty in rural areas in any one region than in urban areas in any other region. Bivariate comparisons, however, are less likely to draw this conclusion for a variety of reasons. For this particular application, all of the interesting deviations from the generally accepted conclusion that poverty is higher in rural areas result from the fact that the correlation between these two dimensions of well-being is often higher in urban areas.

Previous work on multidimensional poverty comparisons has ignored sampling variability, yet this is fundamental if the study of multidimensional poverty comparisons is to have any practical application. The poverty comparisons here are all statistical, using consistent, distribution-free estimators of the sampling distributions of the statistics of each poverty comparison.

6. Bourguignon and Chakravarty (2003, p. 31) refer to this as a "correlation increasing switch" and discuss it in detail. It is closely related to Tsui’s (1999) concept of correlation increasing majorization.
I. Methods to Compare Poverty with Multiple Indicators of Well-Being

This section discusses the data and provides an intuitive presentation of multidimensional poverty comparisons.

Data

The data for this study come from the 1988 Ghana Living Standards Survey, the 1993 National Household Survey (Enquête Permanente auprès des Ménages) in Madagascar, and the 1999 National Household Survey in Uganda. All are nationally representative multipurpose household surveys.

The first measure of well-being is household expenditures per capita, the standard variable for empirical poverty analysis in developing economies. The second is children's height-for-age $z$ score (HAZ) that measures how a child's height compares with the median of the World Health Organization reference sample of healthy children (WHO 1983). In particular, the $z$ scores standardize a child's height by age and gender as \( (x_i - x_{\text{median}}) / \sigma_x \), where \( x_i \) is a child's height, \( x_{\text{median}} \) the median height of children in a healthy and well-nourished reference population of the same age and gender, and \( \sigma_x \) the standard deviation from the mean of the reference population. Thus, the $z$-score measures the number of standard deviations that a child's height is above or below the median for a reference population of healthy children of the same age and gender.

The nutrition literature includes a wealth of studies showing that in poor countries children's height is a particularly good summary measure of children's general health status (Cole and Parkin 1977; Mosley and Chen 1984; WHO 1995). As summarized by Beaton and others (1990, p. 2), growth failure is "the best general proxy for constraints to human welfare of the poorest, including dietary inadequacy, infectious diseases and other environmental health risks." They go on to point out that the usefulness of stature is that it captures the "multiple dimensions of individual health and development and their socioeconomic and environmental determinants." In addition, HAZ is an interesting variable to consider with expenditures per capita because the two are, surprisingly, not highly correlated, so that they capture different dimensions of well-being (Haddad and others 2003).

Univariate Poverty Dominance Methods

The theoretical and statistical bases for the methods used here are developed in Duclos, Sahn, and Younger (2003). This section provides only an intuitive presentation; the formal argument is presented in the appendix. Even though the goal is to make multidimensional poverty comparisons, it is easier to grasp the intuition with a unidimensional example.

Consider the question: Is poverty greater in urban or rural areas? The dominance approach to poverty analysis addresses this question by making poverty comparisons that are valid for a wide range of poverty lines and a broad class of poverty measures. Figure 1 displays the cumulative density functions—or distribution functions—for real household expenditures per capita in urban and rural areas of Uganda in 1999. If the values on the x axis are thought of as potential poverty lines—the amount that a household has to spend per capita in order not to be poor—then the corresponding value on the y axis would be the headcount poverty rate—the share of people whose expenditure is below that particular poverty line. Note that this particular cumulative density function is sometimes called a poverty incidence curve. The graph makes clear that no matter which poverty line one chooses, the headcount poverty index (the share of the sample that is poor) will always be lower for urban areas than for rural. Thus, this sort of poverty comparison is robust to the choice of a poverty line.

What is less obvious is that this type of comparison also permits drawing conclusions about poverty according to a very broad class of poverty measures. In particular, if the poverty incidence curve for one sample is everywhere below the poverty incidence curve for another sample over a bottom range of poverty lines, then poverty will be lower in the first sample for all those poverty lines and for all additive poverty measures that obey two conditions: they are nondecreasing and anonymous. Nondecreasing means that if any one person’s income increases, the poverty measure cannot increase as well. Anonymous means that

**Figure 1. Poverty Incidence Curves, Urban and Rural Areas of Uganda 1999**

*Source:* Authors’ analysis based on data from the Uganda 1999 National Household Survey.
it does not matter which person occupies which position or rank in the income distribution. It is helpful to denote as $\Pi^1$ the class of all poverty measures that have these characteristics. $\Pi^1$ includes virtually every standard poverty measure. It should be clear that the nondecreasing and anonymous characteristics of the class $\Pi^1$ are entirely unobjectionable. Additivity is perhaps less benign, but it is a standard feature of the poverty measures because it allows subgroup decomposition (Foster, Greer, and Thorbecke 1984).

Comparing cumulative density curves as in figure 1 thus enables making a very general statement about poverty in urban and rural Uganda: for any reasonable poverty line and for the class of poverty measures $\Pi^1$, poverty is lower in urban areas than in rural areas. This is called first-order poverty dominance. The generality of such conclusions makes poverty dominance methods attractive. However, such generality comes at a cost. If the cumulative density functions cross one or more times, there is no clear ordering—it cannot be said whether poverty is lower in one group or the other.

There are two ways to deal with this problem, both reasonably general. First, it is possible to conclude that poverty is lower in one sample than in another for the same large class of poverty measures, but only for poverty lines up to the first point at which the cumulative density functions cross (for a recent treatment of this, see Duclos and Makdissi 2005). If reasonable people agree that this crossing point is at a level of well-being safely beyond any sensible poverty line, this conclusion may be sufficient. Second, it is possible to make comparisons over a smaller class of poverty measures. For example, if the condition is added that the poverty measure respects the Pigou–Dalton transfer principle, it turns out that the areas under the crossing poverty incidence curves can be compared. If the area under one curve is less than the area under another for a bottom range of reasonable poverty lines, poverty will be lower for the first sample for all additive poverty measures that are nondecreasing, are anonymous, and obey the Pigou–Dalton transfer principle. This is called second-order poverty dominance, and the associated class of poverty measures is called $\Pi^2$. While not as general as first-order dominance, it is still a quite general conclusion.

**Bivariate Poverty Dominance Methods**

Bivariate poverty dominance comparisons extend the univariate methods discussed above. If there are two measures of well-being rather than one, figure 1 becomes a three-dimensional graph, with one measure of well-being on the $x$ axis, a second on the $y$ axis, and the bivariate cumulative density function

---

8. The Pigou–Dalton transfer principle says that a marginal transfer from a richer person to a poorer person should decrease (or not increase) the poverty measure. Again, this seems entirely sensible, but note that it does not work for the headcount whenever a richer person located initially just above the poverty line falls below the poverty line because of the transfer to the poorer person.

9. If second-order poverty dominance cannot be established, it is possible to integrate once again and check for poverty dominance for a still smaller class of poverty indices and so on. See Zheng (2000) and Davidson and Duclos (2000) for more detailed discussions.
The bivariate cumulative density function is now a surface rather than a line, and one cumulative density function surface is compared with another, just as in figure 1. If one such surface is everywhere below another, poverty in the first sample is lower than poverty in the second sample for a broad class of poverty measures, just as in the univariate case. It is also useful to note that univariate poverty incidence curves are the marginal cumulative densities in the picture found at the extreme edges of the bivariate surface.

That class, now called \( \Pi^{1,1} \) to indicate that it is first order in both dimensions of well-being, has characteristics analogous to those of the univariate case—additive, nondecreasing in each dimension, and anonymous—and one more: the two dimensions of well-being must be substitutes (or more precisely, must not be complements) in the poverty measure. Roughly, this means that an increase of well-being in one dimension should have a greater effect on poverty the lower the level of well-being in the other dimension. In most cases, this restriction is sensible: if we are able to improve a child's health, for example, it seems ethically right that this should reduce overall poverty the most when the child is very poor in the income dimension. But there are some plausible exceptions. For example, suppose that only healthy children can learn in school. Then, it might reduce poverty more to concentrate health improvements on children who are in school (better-off in the education dimension) because of the complementarity of health and education.
Practically, it is not easy to plot two surfaces such as the one in figure 2 on the same graph and to see the differences between them, but the differences can be plotted directly. If this difference always has the same sign, one or the other of the samples has lower poverty for a large class \( \Pi_{1,1} \) of poverty measures. If the surfaces cross, the distributions can be compared at higher orders of dominance, just as in the univariate case. This can be done in one or both dimensions of well-being, and the restrictions on the applicable classes of poverty measures are similar to the univariate case.

**Intersection, Union, and Intermediate Poverty Definitions.** In addition to the extra conditions on the class of poverty indices, multivariate dominance comparisons require distinguishing among union, intersection, and intermediate poverty measures. This can be done with the help of figure 3 that shows the domain of dominance surfaces—the \((x,y)\) plane. The function \( \lambda_1(x,y) \) defines an "intersection" poverty index: someone can be considered poor only when poor in both dimensions \( x \) and \( y \) and therefore when lying within the dashed rectangle of figure 3. The function \( \lambda_2(x,y) \) (the L-shaped dotted line) defines a union poverty index: someone can be considered poor when poor in either of the two

**Figure 3.** Intersection, Union, and Intermediate Dominance Test Domains
dimensions and therefore when lying below or to the right of the dotted line. Finally, \( \lambda_3(x,y) \) provides an intermediate approach. Someone can be considered poor even with a \( y \) value greater than the poverty line in the \( y \) dimension if the \( x \) value is low enough to lie to the left of \( \lambda_3(x,y) \).

For one sample to have less intersection poverty than another for any poverty line up to \( z \) and \( z_x \), its dominance surface must be below the second sample’s everywhere within an area such as the one defined by \( \lambda_1(x,y) \). To have less union poverty, its surface must be below the second sample’s everywhere within an area such as the one defined by \( \lambda_2(x,y) \) and, similarly, for intermediate definitions and \( \lambda_3(x,y) \). The \( \lambda(x,y) \) function delimits the domain over which dominance tests are compared. As such, it is comparable to the maximal poverty line in a univariate comparison.

**Multivariate and Human Development Index Poverty Comparisons.** Figure 3 is also helpful for understanding the difference between the general multivariate poverty comparisons used here and comparisons that rely on indices created with multiple indicators of well-being, the best known of which is the human development index (UNDP 1990). An individual-level index of the \( x \) and \( y \) measures of well-being in figure 3 might be written as

\[
I = a_x x + a_y y
\]

where \( a_x \) and \( a_y \) are weights assigned to each variable. This index is now a univariate measure of well-being and could be used for poverty comparisons such as those in figure 1. The domain of this test for such an index would follow a ray starting at the origin and extending into the \( (x,y) \) plane at an angle that depends on the relative size of the weights \( a_x \) and \( a_y \). Testing for dominance at these points only is clearly less general than testing over the entire area defined by a \( \lambda(x,y) \) function in figure 3.

A comparison of poverty in rural Toliara and urban Mahajanga/Antsiranana in Madagascar shows why this generalization of human development index-type univariate indices is important. Table 1 summarizes the value of the \( t \) statistic for a test of the difference in the two areas’ poverty surfaces at a 10 x 10 grid of test points in the domain of figure 3—the \( (x,y) \) plane of that figure. The origin (the poorest people) is in the lower left corner, and the grid of test points is set at each decile of the marginal distributions. The significantly negative differences are highlighted in light gray and the significantly positive differences in dark gray. In

---

10. The human development index is actually cruder than this, as it first aggregates across individuals each dimension of well-being to generate a single scalar measure and then constructs a weighted average of those scalars to generate the index, which is also a scalar. Dutta, Pattanaik, and Xu (2003) discuss the severe restrictions needed on a social welfare function to justify such an index.

11. In theory, differences in the surfaces should be tested for everywhere, but this is computationally expensive. In practice, because the surfaces are smoothly increasing functions, it is usually sufficient to test at a grid of points, as is done here.
Table 1: Dominance Tests for Rural and Urban Areas in Toliara, Madagascar, 1993 (Differences Between Rural and Urban Dominance Surfaces)

<table>
<thead>
<tr>
<th></th>
<th>Rural</th>
<th>Urban</th>
<th>Rural</th>
<th>Urban</th>
<th>Rural</th>
<th>Urban</th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.51</td>
<td>-8.84</td>
<td>-16.32</td>
<td>-16.39</td>
<td>-11.43</td>
<td>-8.08</td>
<td>-6.65</td>
<td>-4.17</td>
<td>-2.21</td>
</tr>
<tr>
<td>13.19</td>
<td>-9.28</td>
<td>-16.70</td>
<td>-16.09</td>
<td>-10.87</td>
<td>-8.71</td>
<td>-6.26</td>
<td>-4.82</td>
<td>-2.80</td>
</tr>
<tr>
<td>12.84</td>
<td>-9.84</td>
<td>-15.69</td>
<td>-15.93</td>
<td>-10.72</td>
<td>-7.05</td>
<td>-5.25</td>
<td>-3.53</td>
<td>-1.93</td>
</tr>
<tr>
<td>12.60</td>
<td>-10.07</td>
<td>-16.96</td>
<td>-11.74</td>
<td>-7.34</td>
<td>-5.63</td>
<td>-4.30</td>
<td>-2.77</td>
<td>-1.38</td>
</tr>
<tr>
<td>12.44</td>
<td>-10.39</td>
<td>-17.50</td>
<td>-12.66</td>
<td>-7.66</td>
<td>-5.81</td>
<td>-4.57</td>
<td>-3.05</td>
<td>-1.67</td>
</tr>
<tr>
<td>12.29</td>
<td>-10.64</td>
<td>-18.04</td>
<td>-13.33</td>
<td>-7.81</td>
<td>-6.04</td>
<td>-5.27</td>
<td>-3.50</td>
<td>-2.09</td>
</tr>
<tr>
<td>12.16</td>
<td>-0.63</td>
<td>-3.07</td>
<td>-2.15</td>
<td>-1.42</td>
<td>-1.63</td>
<td>-1.06</td>
<td>-0.49</td>
<td>-0.08</td>
</tr>
<tr>
<td>12.00</td>
<td>-0.78</td>
<td>-3.38</td>
<td>-2.21</td>
<td>-1.48</td>
<td>-1.81</td>
<td>-1.23</td>
<td>-0.66</td>
<td>-0.12</td>
</tr>
<tr>
<td>11.82</td>
<td>-1.13</td>
<td>-3.61</td>
<td>-2.31</td>
<td>-1.54</td>
<td>-2.15</td>
<td>-1.56</td>
<td>-0.99</td>
<td>-0.18</td>
</tr>
<tr>
<td>11.48</td>
<td>-1.48</td>
<td>-3.84</td>
<td>-2.45</td>
<td>-1.62</td>
<td>-2.45</td>
<td>-1.97</td>
<td>-1.02</td>
<td>-0.27</td>
</tr>
<tr>
<td>0.00</td>
<td>-0.40</td>
<td>-3.33</td>
<td>-2.84</td>
<td>-1.98</td>
<td>-1.63</td>
<td>-1.21</td>
<td>-0.71</td>
<td>0.12</td>
</tr>
</tbody>
</table>

**Note:** The significantly negative differences are highlighted in light gray and the significantly positive differences in dark gray. Weights \(a_1\) and \(a_2\) are chosen, so that a human development index-type index of these two dimensions of well-being traces the diagonal, here highlighted in bold.

**Source:** Authors' analysis based on data from the Madagascar 1993 National Household Survey (Enquête Permanente auprès des Ménages).

choosing the weights \(a_1\) and \(a_2\) so that a human development index-type index of these two dimensions of well-being traces out the diagonal of table 1, it can be concluded that poverty is higher in rural Toliara for a wide range of poverty lines—up to the 70th percentile—and all poverty measures in the \(\Pi_1\) class. However, another choice of \(a_1\) and \(a_2\) that gives more weight to household expenditures would yield test points on a steeper ray from the origin and thus imply a significant crossing of the index's poverty incidence curves, yielding no dominance result. Testing over the entire two-dimensional domain rather than a single ray within that domain avoids this problem.

**Multivariate and Multiple Univariate Poverty Comparisons.** Suppose that a univariate comparison of expenditures per capita in two samples, as in figure 1, and children's heights in two samples finds that for both variables, one sample shows lower poverty for all poverty lines and a large class of poverty measures. Is that not sufficient to conclude that poverty differs in the two samples? Unfortunately, no.

The complication comes from the "hump" in the middle of the dominance surface shown in figure 2. How sharply the hump rises depends on the correlation between the two measures of well-being. If they are highly correlated, the surface rises rapidly in the center and vice versa. Thus, it is possible for one surface to be lower than another at both extremes (the edges of the surface farthest from the origin) and yet higher in the middle if the correlation between the welfare variables is higher. (The far edges of each surface integrate out one variable, and so are the univariate cumulative density functions depicted in figure 1.) Thus, in this case one surface would have lower univariate
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HAZ</td>
<td>-3.100</td>
<td>-2.450</td>
<td>-1.970</td>
<td>-1.580</td>
<td>-1.220</td>
<td>-0.880</td>
<td>-0.500</td>
<td>-0.010</td>
<td>0.690</td>
<td>5.820</td>
</tr>
</tbody>
</table>

**Note:** The significantly negative differences are highlighted in light gray and the significantly positive differences in dark gray.

**Source:** Authors’ analysis based on data from the Uganda 1999 National Household Survey.

Cumulative density functions, and thus lower poverty, for both measures of well-being independently, but it would not have lower bivariate poverty. Intuitively, samples with higher correlation of deprivation in multiple dimensions have higher poverty than samples with lower correlation because lower well-being in one dimension contributes more to poverty if well-being is also low in the other dimension.12

Consider this example. Univariate poverty is unambiguously higher in the rural Central region of Uganda than in the urban Eastern region in both dimensions—the difference between the dominance surfaces at the extreme top and right edges of Table 2 is always positive—yet bivariate poverty is not unambiguously higher because of the statistically significant reversal of the dominance surfaces in the interior. Similar comparisons up to third order in each dimension also find that the dominance surfaces cross for these two areas.

It is also possible that two samples with different correlations between measures of well-being have univariate comparisons that are inconclusive—they cross at the extreme edges of the dominance surfaces—but have bivariate surfaces that are different for a large part of the interior of the dominance surface. (The sample with lower correlation would have a lower dominance surface.) This would establish different intersection multivariate poverty even though either one or both of the univariate comparisons are inconclusive. It could not,

---

12. "Correlation" is actually overly strict. For instance, a recent literature has emerged on copulas, namely, functions that link two univariate distributions in ways that are more general than simple linear correlations but less flexible than the nonparametric distributions here. If these copulas differ for two groups, even if their correlations between dimensions of well-being are the same, it is still the case that one-at-a-time univariate dominance results could be reversed with a multivariate comparison.
however, establish union poverty dominance, since that requires difference in the surfaces at the extremes as well as in the middle.

Consider the example for rural Central and urban Northern Uganda (table 3). There is no statistically significant univariate dominance in the height-for-age dimension of well-being, and only a limited range of poverty lines for which poverty differs in the expenditure dimension, but there is a sizable domain—up to the ninth decile in each dimension—over which poverty is lower in the rural Central region than in the urban Northern region for all intersection poverty indices in the $\Pi^{2,2}$ class. Thus, for many intersection and intermediate poverty measures, it can be concluded that the rural Central region in Uganda is less poor than the urban Northern region, even though neither univariate comparison is conclusive.

### II. Bivariate Spatial Poverty Comparisons in Africa

This section applies bivariate dominance tests to spatial poverty comparisons in Ghana, Madagascar, and Uganda. Poverty, measured by household expenditures per capita and children's HAZ, is compared in urban and rural areas, nationally and by region. The tests produce a large amount of output in the

13. The regions used in Ghana are its standard ecological zones of Coast, Forest, and Savannah. In Uganda, the four political regions are used: Central, Eastern, Western, and Northern. In Madagascar, political regions are also used, but because of small sample sizes Fianarantsoa and Toamasina are combined into one region, as are Mahajanga and Antsiranana. This choice is based on similar agro-ecological characteristics. In all countries, rural and urban areas in these regions are considered.
form of tables, such as table 1.\textsuperscript{14} Only summaries of the dominance results are reported here.\textsuperscript{15}

Table 4 gives descriptive statistics for HAZ and the log of household expenditures per capita, \( \ln(y) \). As expected, poverty measured by expenditures per capita and also stunting\textsuperscript{16} is higher in rural areas than in urban areas in each country. The same is true within each region of each country, except for the Toliara region in Madagascar, where stunting is higher in urban areas than in rural areas. In fact, with a few exceptions in Madagascar, both expenditure and height poverty are lower in urban areas in any region of each country than in rural areas in any other region in the same country.

In addition to the means and poverty rates, table 4 reports the correlation between the log of expenditures per capita and HAZ. Note that in Madagascar and Uganda, expenditures and heights are more highly correlated in urban areas than in rural areas, whereas both expenditures and heights tend to be higher in urban areas. As noted, this combination can cause bivariate poverty comparisons to differ from univariate comparisons carried out separately in each dimension of well-being.\textsuperscript{17}

The dominance results for tests across urban and rural areas in Ghana, Madagascar, and Uganda show that for each country as a whole, poverty is higher in rural areas than in urban areas for each univariate poverty comparison and for both intersection and union bivariate comparisons. These results are entirely consistent with virtually every known poverty comparison based on incomes or expenditures alone—poverty is lower in urban areas.

In the regional comparisons, however, a significant number of exceptions to this widely held belief emerge, especially for the bivariate comparisons. Ghana has the fewest of these exceptions, with two of nine urban–rural comparisons being statistically insignificant for both intersection and union bivariate comparisons. These results are entirely consistent with virtually every known poverty comparison based on incomes or expenditures alone—poverty is lower in urban areas.

14. The results are relegated to appendixes, which are available from the authors.
15. The relevant statistics and their asymptotic standard errors can be readily computed using the software DAD (version 4.4 and higher) that is freely available at www.mimap.ecn.ulaval.ca. The authors can also provide a GAUSS program that does the same.
16. Stunting is usually defined as an HAZ of less than \(-2\).
17. It is difficult to find universal explanations for the empirical correlations between indicators. The reasons are clearly context specific. As an example, expenditures and heights may be more highly correlated in urban than in rural areas because in urban areas the use of food markets may be prevalent. Purchasing power would then be better correlated with nutrient intake. In rural areas, nutrient intake is plausibly less correlated with purchasing power and more correlated with the proximity of food producers.
18. In each country, rural areas in each region are compared with urban areas in each region. Since there are three regions in Ghana, this yields nine comparisons. For Uganda and Madagascar, with four regions, this yields 16 comparisons.
TABLE 4. Descriptive Statistics for Poverty and Stunting for Ghana, Madagascar, and Uganda

<table>
<thead>
<tr>
<th>Region</th>
<th>Mean</th>
<th>Percent Stunted</th>
<th>Poor</th>
<th>N</th>
<th>Correlation ln(y), HAZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghana 1988</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coast</td>
<td>-0.98</td>
<td>11.90</td>
<td>0.22</td>
<td>0.41</td>
<td>911</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.12</td>
<td>11.76</td>
<td>0.27</td>
<td>0.51</td>
<td>488</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.82</td>
<td>12.06</td>
<td>0.16</td>
<td>0.30</td>
<td>423</td>
</tr>
<tr>
<td>Forest</td>
<td>-1.38</td>
<td>11.81</td>
<td>0.32</td>
<td>0.46</td>
<td>1,074</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.48</td>
<td>11.79</td>
<td>0.35</td>
<td>0.48</td>
<td>793</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.10</td>
<td>11.88</td>
<td>0.24</td>
<td>0.39</td>
<td>281</td>
</tr>
<tr>
<td>Savannah</td>
<td>-1.30</td>
<td>11.66</td>
<td>0.32</td>
<td>0.55</td>
<td>683</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.37</td>
<td>11.63</td>
<td>0.33</td>
<td>0.56</td>
<td>591</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.86</td>
<td>11.85</td>
<td>0.23</td>
<td>0.48</td>
<td>92</td>
</tr>
<tr>
<td>National</td>
<td>-1.22</td>
<td>11.80</td>
<td>0.28</td>
<td>0.47</td>
<td>2,668</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.35</td>
<td>11.73</td>
<td>0.32</td>
<td>0.51</td>
<td>1,872</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.92</td>
<td>11.97</td>
<td>0.19</td>
<td>0.35</td>
<td>796</td>
</tr>
<tr>
<td>Madagascar 1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tana</td>
<td>-2.24</td>
<td>12.32</td>
<td>0.57</td>
<td>0.73</td>
<td>928</td>
</tr>
<tr>
<td>Rural</td>
<td>-2.33</td>
<td>12.26</td>
<td>0.60</td>
<td>0.78</td>
<td>534</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.80</td>
<td>12.65</td>
<td>0.40</td>
<td>0.48</td>
<td>394</td>
</tr>
<tr>
<td>Fian/T'ils</td>
<td>-2.15</td>
<td>12.26</td>
<td>0.53</td>
<td>0.77</td>
<td>975</td>
</tr>
<tr>
<td>Rural</td>
<td>-2.19</td>
<td>12.22</td>
<td>0.54</td>
<td>0.80</td>
<td>705</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.74</td>
<td>12.56</td>
<td>0.48</td>
<td>0.56</td>
<td>270</td>
</tr>
<tr>
<td>Mahajanga/Antsiranana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>-1.32</td>
<td>12.61</td>
<td>0.34</td>
<td>0.56</td>
<td>346</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.44</td>
<td>12.71</td>
<td>0.34</td>
<td>0.50</td>
<td>215</td>
</tr>
<tr>
<td>Toliara</td>
<td>-1.91</td>
<td>12.06</td>
<td>0.48</td>
<td>0.78</td>
<td>457</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.82</td>
<td>11.98</td>
<td>0.45</td>
<td>0.82</td>
<td>302</td>
</tr>
<tr>
<td>Urban</td>
<td>-2.36</td>
<td>12.46</td>
<td>0.60</td>
<td>0.57</td>
<td>155</td>
</tr>
<tr>
<td>National</td>
<td>-1.97</td>
<td>12.33</td>
<td>0.50</td>
<td>0.71</td>
<td>2,921</td>
</tr>
<tr>
<td>Rural</td>
<td>-2.01</td>
<td>12.27</td>
<td>0.51</td>
<td>0.75</td>
<td>1,887</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.79</td>
<td>12.61</td>
<td>0.44</td>
<td>0.52</td>
<td>1,034</td>
</tr>
<tr>
<td>Uganda 1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>-1.00</td>
<td>8.80</td>
<td>0.25</td>
<td>0.19</td>
<td>1,806</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.08</td>
<td>8.65</td>
<td>0.27</td>
<td>0.23</td>
<td>1,390</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.77</td>
<td>9.22</td>
<td>0.18</td>
<td>0.08</td>
<td>416</td>
</tr>
<tr>
<td>Eastern</td>
<td>-1.22</td>
<td>8.48</td>
<td>0.28</td>
<td>0.38</td>
<td>2,349</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.25</td>
<td>8.45</td>
<td>0.28</td>
<td>0.39</td>
<td>2,010</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.75</td>
<td>8.99</td>
<td>0.21</td>
<td>0.14</td>
<td>339</td>
</tr>
<tr>
<td>Western</td>
<td>-1.42</td>
<td>8.63</td>
<td>0.34</td>
<td>0.28</td>
<td>2,996</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.46</td>
<td>8.60</td>
<td>0.35</td>
<td>0.29</td>
<td>1,860</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.59</td>
<td>9.35</td>
<td>0.15</td>
<td>0.06</td>
<td>236</td>
</tr>
<tr>
<td>Northern</td>
<td>-1.24</td>
<td>8.16</td>
<td>0.30</td>
<td>0.60</td>
<td>1,230</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.24</td>
<td>8.13</td>
<td>0.30</td>
<td>0.62</td>
<td>1,008</td>
</tr>
<tr>
<td>Urban</td>
<td>-1.23</td>
<td>8.72</td>
<td>0.26</td>
<td>0.19</td>
<td>222</td>
</tr>
<tr>
<td>National</td>
<td>-1.22</td>
<td>8.54</td>
<td>0.29</td>
<td>0.35</td>
<td>7,481</td>
</tr>
<tr>
<td>Rural</td>
<td>-1.27</td>
<td>8.47</td>
<td>0.30</td>
<td>0.37</td>
<td>6,268</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.79</td>
<td>9.15</td>
<td>0.19</td>
<td>0.10</td>
<td>1,213</td>
</tr>
</tbody>
</table>

Source: Authors' analysis based on data from the 1988 Ghana Living Standards Study survey, the Madagascar 1993 National Household Survey (Enquête Permanente auprès des Ménages), and the Uganda 1999 National Household Survey.
Madagascar, for 7 of 16 intersection comparisons and 10 of 16 union comparisons, the null hypothesis that bivariate poverty is the same in urban and rural areas cannot be rejected, though none of these reject the null in favor of rural areas. While it is true that in only a minority of cases are urban areas not found to have significantly lower poverty, the fact that there are any such cases is surprising, given the overwhelming number of studies that find lower univariate poverty in urban areas in all developing economies.

One immediate concern with these results is that the interesting cases are the ones in which the null hypothesis of nondominance is not rejected, so the results may be driven by a lack of power in the statistical tests. This concern is reinforced by the relatively few observations available in some urban areas. Review of the appendix tables shows, however, that in most cases in which bivariate dominance is not found, the dominance surfaces actually cross significantly.19 That is, there are points in the test domain where the rural surface is significantly above the urban surface and vice versa. Thus, the lack of bivariate dominance is typically not due to a lack of power.

To gain a better understanding of how bivariate and univariate dominance methods can differ, we classified the results into five types. Type 1 has dominance (usually first order) for both univariate comparisons and for intersection and union bivariate comparisons. This is the most common result, accounting for 25 of the 41 comparisons. This is also the least interesting type of result for the methods applied here. Why bother with the more complicated bivariate comparisons if, in the end, they produce the same results as simpler univariate dominance tests or even scalar comparisons?

Type 2 occurs when neither the univariate nor the bivariate method finds dominance. This is equally uninteresting for the methods used here. There is only one such case for urban and rural Mahajanga/Antsiranana region in Madagascar.

Type 3 is a case in which urban areas dominate rural areas for both univariate comparisons but not for the bivariate comparisons. There are six of these cases. There is also one case, rural Mahajanga/Antsiranana compared with urban Toliara, in which the rural area dominates on both univariate comparisons but not on the bivariate comparisons. For cases in which the bivariate comparisons are inconsistent with the univariate comparisons, a type 3 result is the most common. The bivariate comparisons are more demanding than univariate comparisons, so it makes sense that they reject the null hypothesis of nondominance less often and this happens in five of the seven cases. In two cases, both involving urban areas in the Northern region of Uganda, the dominance result is actually reversed for intersection poverty measures over a limited domain. This is surprising, but understandable considering the very high correlation (0.36) between expenditures and heights in urban Northern region compared with rural Western region (0.07) and Eastern region (0.06).

19. The results are relegated to appendixes, which are available from the authors.
Type 4 occurs when the univariate results are contradictory in the sense that univariate dominance is found in one dimension but not in the other. There are six such occurrences, and in all but one the urban area dominates in one dimension, usually expenditures, although in one case, rural Central compared with urban Northern region in Uganda, the rural area dominates, albeit only for the $I^3$ class. Of these six cases, intersection dominance is found for four bivariate tests. That is, the bivariate tests are able to "resolve" the conflicting univariate results for at least some classes of poverty measures and areas of poverty lines.

Type 5 is similar to type 4 except that the contradictory univariate results are statistically significant in each univariate comparison. There are only two of these cases, rural compared with urban Toliara and rural Coast compared with urban Forest in Ghana. Unlike the type 4 results, in neither case are any of the bivariate poverty comparisons statistically significant, so the bivariate comparisons cannot resolve the univariate conflict.

Overall, sufficient evidence has not been amassed to overturn the standard presumption that poverty is lower in urban than in rural areas, but enough of the results are at odds with this idea to introduce doubt. Furthermore, the reasons that this is not found for bivariate poverty comparisons vary. For the type 4 and 5 cases, no univariate dominance is found in one dimension or another and the bivariate results follow from that. But this is relatively rare, and in about half of the cases the bivariate tests for intersection poverty measures do find that poverty is lower in urban areas despite the contradictory univariate results. Most of the differences, though, come from the fact that the two measures of well-being are often more highly correlated in urban areas than in rural areas. As noted, this correlation causes the poverty incidence surface to rise more rapidly near the origin of the distribution, raising it above the rural surface in the center even though it is below it at the extremes, where the univariate poverty incidence curves lie. In most cases, this gives results in which an urban area dominates a rural area in each dimension individually, but not jointly, because multiple deprivation is more common in urban areas. There are two cases, however, in which the dominance is reversed, so that for some intersection poverty measures the rural area dominates the urban area.

III. Conclusions

This article used bivariate stochastic dominance techniques to compare poverty in urban and rural areas in three African countries, measuring poverty in terms of expenditure per capita and children's standardized heights, a good measure of children's health status. The comparisons are shown to be more general than either a comparison of a human development index-type index or one-at-a-time

---

20. As noted in the methods discussion, bivariate dominance for union poverty measures requires univariate dominance in each dimension, so it is impossible for this type of result.
comparisons of multiple measures of well-being. More important, the article finds that its more general methods are at odds with simpler univariate poverty comparisons in a nontrivial number of cases.

Expenditure-based urban–rural poverty comparisons almost always find that rural areas are poorer than urban areas. The results are consistent with that finding whether univariate or bivariate comparisons are used. However, differences emerge when urban areas in one region of a country are compared with rural areas in another region. In several cases, univariate poverty is lower in urban areas in both dimensions, but bivariate poverty is not. This happens because the correlation between expenditures per capita and children's heights is higher in the urban areas, so that urban residents who are expenditure poor are also more likely to be health poor. This correlation yields a higher density of observations in the poorest part of the bivariate welfare domain for urban areas, even though there are fewer observations for urban residents at the lower end of the density for each individual measure of well-being. Taking such a correlation into account is important for welfare comparisons because the social cost of poverty in one dimension, say health, is higher if the person affected is also poor in the other dimension (expenditures in this case).

It is interesting to note that the share of cases in which urban areas do not dominate rural is much higher in the bivariate comparisons than in the expenditure- or income-based comparisons in the literature, where poverty is almost always found to be lower in urban areas. With two exceptions in Madagascar, however, the urban area in the region where the capital city is located always has lower poverty than every rural area in both univariate and bivariate comparisons. Thus, the doubts raised here apply only to other urban areas in these countries.

There are other instances in which the bivariate comparisons are at odds with univariate comparisons. Perhaps the most interesting are cases in which univariate results are inconclusive because one or the other univariate comparison is inconclusive, yet the bivariate results find dominance for a large domain of intersection poverty indices. This arises in about 10 percent of the examples and occurs again when the correlation between expenditures per capita and children's heights differs significantly across areas. These results are interesting because they show that bivariate comparisons may provide statistically significant results when univariate comparisons do not.

The finding that bivariate results often differ from the standard perception of greater rural poverty typically occurs not because children are taller in rural areas, but rather because the correlation between expenditures and heights is lower there than in urban areas. This, however, is based on only three countries. Pursuing similar research in other countries will yield insight into whether these results are anomalous. Why this should be is also an interesting question for future research. But a clear implication of these results for researchers and policymakers interested in multiple dimensions of poverty is that, at a minimum, one should check the correlations
between measures of well-being in the groups of interest. Large differences in these correlations may lead to unexpected multivariate dominance comparisons.

APPENDIX

The following is based on the companion paper, Duclos, Sahn, and Younger (2003).

Making Poverty Comparisons with Multiple Indicators of Well-Being

For expositional simplicity, the focus is on the case of two dimensions of individual well-being. Let $x$ and $y$ be two such indicators. Assuming differentiability, denote by

$$\lambda(x, y) : \mathbb{R}^2 \to \mathbb{R}, \quad \frac{\partial \lambda(x, y)}{\partial x} \geq 0, \quad \frac{\partial \lambda(x, y)}{\partial y} \geq 0$$

a summary indicator of individual well-being, analogous to but not necessarily the same as a utility function. Note that the derivative conditions in equation (A.1) simply mean that different indicators can each contribute to overall well-being. Assume that an unknown poverty frontier separates the poor from the rich, defined implicitly by a locus of the form $\lambda(x, y) = 0$ and analogous to the usual downward sloping indifference curves on the $(x, y)$ space. The set of the poor is then obtained as:

$$\Lambda(\lambda) = \{(x, y) | (\lambda(x, y) \leq 0)\}.$$

Letting the joint distribution of $x$ and $y$ be denoted by $F(x, y)$, assume for simplicity that the multidimensional poverty indices are additive across individuals and define such indices by $P(\lambda)$:

$$P(\lambda) = \int_{\Lambda(\lambda)} \pi(x, y; \lambda) dF(x, y)$$

where $\pi(x, y; \lambda)$ is the contribution to poverty of an individual with well-being indicators $x$ and $y$:

$$\pi(x, y; \lambda) \begin{cases} \geq 0 & \text{if } \lambda(x, y) \leq 0 \\ = 0 & \text{otherwise.} \end{cases}$$

Here, $\pi$ is the weight that the poverty measure attaches to someone inside the poverty frontier. By the focus axiom, it has to be 0 for those outside the poverty frontier. A bidimensional stochastic dominance surface can then be defined as:

$$P^{x, y}(z_x, z_y) = \int_0^{z_y} \int_0^{z_x} (z_x - x)^{\alpha_x}(z_y - y)^{\alpha_y} dF(x, y).$$

This function looks like a two-dimensional generalization of the Foster–Greer–Thorbecke index and can also be interpreted as such. The poverty comparisons
here make use of orders of dominance, $s_x$ in the $x$ and $s_y$ in the $y$ dimensions, which correspond respectively to $s_x = \alpha_x + 1$ and $s_y = \alpha_y + 1$.

Assume that the general poverty index in equation (A.3) is left differentiable with respect to $x$ and $y$ over the set $A(\lambda)$, up to the relevant orders of dominance, $s_x$ for derivatives with respect to $x$ and $s_y$ for derivatives with respect to $y$. Denote by $\pi x$ the first derivative of $\pi(x,y;A)$ with respect to $x$, by $\pi y$ the first derivative of $\pi(x,y;A)$ with respect to $y$, and by $\pi xy$ the derivative of $\pi(x,y;A)$ with respect to $x$ and to $y$. The following class $\Pi_{1,1}^{*}(A^*)$ of bidimensional poverty indices can be defined as:

\[
\Pi_{1,1}^{*}(A^*) = \begin{cases} 
\Delta(\lambda) \subset A(\lambda^*) \\
\pi(x,y;A) = 0 & \text{whenever } \lambda(x,y) = 0 \\
\pi x \leq 0 \text{ and } \pi y \leq 0 & \forall x,y \\
\pi xy \geq 0 & \forall x,y 
\end{cases}
\]

The first line on the right of equation (A.6) defines the largest poverty set to which the poor must belong: the poverty set covered by the $P(\lambda)$ indices should lie within the maximal set $\Delta(\lambda^*)$. The second line assumes that the poverty indices are continuous along the poverty frontier. The third line says that indices that are members of $\Pi_{1,1}^{*}$ are weakly decreasing in $x$ and in $y$. The last line assumes that the marginal poverty benefit of an increase in either $x$ or $y$ decreases with the value of the other variable.

Denote by $A F = F_A - F_B$ the difference between a function $F$ for $A$ and for $B$. The class of indices defined in equation (A.6) then gives rise to the following theorem:

**Theorem 1**

\[
\Delta P(\lambda) > 0, \quad \forall P(\lambda) \sqsubseteq \Pi_{1,1}^{*}(A^*), \\
\text{if } \Delta P^{0,0}(x,y) > 0, \quad \forall(x,y) \sqsubseteq \Delta(\lambda^*).
\]

Proof: Denote the points on the outer poverty frontier $\lambda^*(x,y) = 0$ as $z_x(y)$ for a point above $y$ and $z_y(x)$ for a point above $x$. The derivative conditions in equation (A.1) imply that $z_x^{(1)}(y) \leq 0$ and $z_y^{(1)}(x) \leq 0$, where the superscript 1 indicates the first-order derivative of the function with respect to its argument. Note that the inverse of $z_x(y)$ is simply $z_y(x) = x \equiv z_y(z_x(x))$. Next, equation (A.3) is integrated by parts with respect to $x$, over an interval of $y$ ranging from 0 to $z_y$. This gives:

\[
P(z_x(y),z_y) = \int_{0}^{z_y} [\pi(x,y;\lambda^*)F(x|y)]|z_x^{(1)}(y)f(y)dy - \int_{0}^{z_y} \int_{0}^{z_x(y)} \pi x(x,y;\lambda^*)F(x|y)f(y)dx\,dy.
\]

(A.8)

To integrate by parts with respect to $y$ the second term, define a general function $K(y) = \int_{0}^{y} h(x,y)l(x,y)dx$ and note that:
\[
\frac{dK(y)}{dy} = g^{(1)}(y)h(g(y), y)l(g(y), y) + g(y) \frac{\partial h(x, y)}{\partial y} l(x, y)dx + \int_{0}^{c} g(y) h(x, y) \frac{\partial l(x, y)}{\partial y} dx.
\]

(A.9)

Reordering equation (A.9) and integrating it from 0 to \(c\) yields:

\[
-\int_{0}^{c} \int_{0}^{g(y)} h(x, y) \frac{\partial l(x, y)}{\partial y} dx dy = -K(c) + K(0) + \int_{0}^{c} g^{(1)}(y)h(g(y), y)l(g(y), y)dy + \int_{0}^{c} g(y) h(x, y) \frac{\partial l(x, y)}{\partial y} l(x, y)dx dy.
\]

(A.10)

Now replace in equation (A.10) \(c\) by \(z_2\), \(g(y)\) by \(z_2(y)\), \(h(x, y)\) by \(\pi^x(x, y; A^*)\), \(l(x, y)\) by \(F(x, y)\), and \(K(y)\) by its definition \(K(y) = \int_{0}^{g(y)} h(x, y)l(x, y)dx\). This gives:

\[
P(z_2(y), z_2) = -\int_{0}^{z_2(y)} \pi^x(x, z_2; A^*) P^{0,0}(x, z_2) dx + \int_{0}^{z_2(y)} z_2^{(1)}(y) \pi^x(z_2(y), y; \lambda^*) P^{0,0}(z_2(y), y) dy + \int_{0}^{z_2(y)} \pi^y(x, y; A^*) P^{0,0}(x, y) dx dy.
\]

(A.11)

For the sufficiency of equation (A.7), recall that \(z_2^{(1)}(y) \leq 0\), \(\pi^x \leq 0\), and \(\pi^y \geq 0\), with strict inequalities for either of these inequalities over at least some inner ranges of \(x\) and \(y\). Hence, if \(\Delta P^{0,0}(x, y) > 0\) for all \(y \in [0, z_2]\) and for all \(x \in [0, z_2(y)]\) (that is, for all \((x, y) \in \Lambda(\lambda^*)\)), it must be that \(\Delta P(\lambda^*) > 0\) for all of the indices that use the poverty set \(\Lambda(\lambda^*)\) and that obey the first two lines of conditions in equation (A.6). But note that for other poverty sets \(\Lambda(\lambda) \subset \Lambda(\lambda^*)\), the relevant sufficient conditions are only a subset of those for \(\Lambda(\lambda^*)\). The sufficiency part of theorem 1 thus follows.

For the necessity part, assume that \(\Delta P^{0,0}(x, y) \leq 0\) for an area defined over \(x \in [c_2^-, c_2^+]\) and \(y \in [c_y^-, c_y^+]\), with \(c_x^+ \leq z_2\) and \(c_x^- \leq z_2(y)\). Then any of the poverty indices in \(\Pi^{1,1}(A^*)\) for which \(\pi^y < 0\) over that area, \(\pi^x = 0\) outside that area, and for which \(\pi^x(x, z_2; A^*) = \pi^x(z_2(y), y; A^*) = 0\), will indicate that \(\Delta P < 0\). Equation (A.7) is thus also a necessary condition for the ordering specified in theorem 1.
Note that similar proofs are possible for dominance comparisons at higher orders (Duclos, Sahn, and Younger 2003).

**Estimation and Inference**

Suppose a random sample of N independently and identically distributed observations drawn from the joint distribution of x and y. These observations of \( x^L \) and \( y^L \), drawn from a population L, can be written as \( (x^L_i, y^L_i) \), \( i = 1, \ldots, N \). A natural estimator of the bidimensional dominance surfaces \( P^{\alpha_x, \alpha_y}(z_x, z_y) \) is then:

\[
\hat{P}_L^{\alpha_x, \alpha_y}(z_x, z_y) = \int_0^{z_y} \int_0^{z_x} (z_y - y)^{\alpha_y}(z_x - x)^{\alpha_x} d \hat{F}_L(x, y)
\]

(A.12)

\[
= \frac{1}{N} \sum_{i=1}^{N} (z_y - y^L_i)^{\alpha_y} (z_x - x^L_i)^{\alpha_x} I(y^L_i \leq z_y) I(x^L_i \leq z_x)
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} (z_y - y^L_i)^{\alpha_y} (z_x - x^L_i)^{\alpha_x}
\]

where \( \hat{F} \) denotes the empirical joint distribution function, \( I(\cdot) \) is an indicator function equal to 1 when its argument is true and 0 otherwise, and \( x_+ = \max(0, x) \). This gives rise to theorem 2:

**Theorem 2**

Let the joint population moments of order 2 of \((z_y - y^A)^{\alpha_y}(z_x - x^A)^{\alpha_x}\) and \((z_y - y^B)^{\alpha_y}(z_x - x^B)^{\alpha_x}\) be finite. Then \( N^{1/2}\left(\hat{P}_A^{\alpha_x, \alpha_y}(z_x, z_y) - P_A^{\alpha_x, \alpha_y}(z_x, z_y)\right) \) and \( N^{1/2}\left(\hat{P}_B^{\alpha_x, \alpha_y}(z_x, z_y) - P_B^{\alpha_x, \alpha_y}(z_x, z_y)\right) \) are asymptotically normal with mean 0, with asymptotic covariance structure given by \((L, M = A, B)\):

\[
\lim_{N \to \infty} N \text{cov}\left(\hat{P}_L^{\alpha_x, \alpha_y}(z_x, z_y), \hat{P}_M^{\alpha_x, \alpha_y}(z_x, z_y)\right) = E\left((-y - y^L)^{\alpha_y}(z_x - x^L)^{\alpha_x} (z_y - y^M)^{\alpha_y}(z_x - x^M)^{-\alpha_x - 1}\right) - P_L^{\alpha_x, \alpha_y}(z_x, z_y) P_M^{\alpha_x, \alpha_y}(z_x, z_y).
\]

(A.13)

**Proof:** For each distribution, the existence of the appropriate population moments of order 1 permits application of the law of large numbers to equation (A.12), thus showing that \( \hat{P}_L^{\alpha_x, \alpha_y}(z_x, z_y) \) is a consistent estimator of \( P_L^{\alpha_x, \alpha_y}(z_x, z_y) \). Given also the existence of the population moments of order 2, the central limit theorem shows that the estimator in equation (A.12) is \( \text{root-N} \) consistent and asymptotically normal with asymptotic covariance matrix given by equation (A.13). When the samples are dependent, the covariance between the estimator for A and for B is also provided by equation (A.13).
REFERENCES


The Impact of Regulatory Governance and Privatization on Electricity Industry Generation Capacity in Developing Economies

*John Cubbin* and *Jon Stern*

This article assesses whether a regulatory law and higher quality regulatory governance are associated with superior outcomes in the electricity industry. The analysis, for 28 developing economies over 1980–2001, draws on theoretical and empirical work on the impact of telecommunications regulators in developing economies. Controlling for privatization and competition and allowing for country-specific fixed effects, both regulatory law and higher quality regulatory governance are positively and significantly associated with higher per capita generation capacity. This positive impact increases for more than 10 years, as experience develops and regulatory reputation grows. The results are robust to estimating alternative dynamic specifications (including error correction models), to inclusion of economy governance political risk indicators, and to controlling for possible endogeneity biases. The article concludes with a short discussion of causality in panel data modeling of governance models and the policy implications for regulatory reform.

Over the past 10–15 years, much attention has gone to the role of institutions in economic growth—in large part to determine the economic policy priorities for developing functioning market economies in Central and Eastern Europe and the former Soviet Union and fostering economic growth in lagging regions, such as sub-Saharan Africa.

In parallel, and partly in response, there have been major explorations of the role of institutions in market economies. There has also been a substantial empirical literature on the relative roles of institutions, policy, geography, and trade openness on growth across economies. This literature places considerable weight on institutional quality as a major determinant of variations in long-term growth.1 In particular, Rodrik (2003, p. 25) argues that there is a requirement

---

for a "cumulative process of institution building to ensure that growth does not run out of steam and that the economy remains resilient to shocks."

These arguments apply with extra force to utility service industries—not just because they are highly capital intensive but also because most of their assets are very long lived and (in economic terms) sunk. So, an effective institutional framework is essential to sustain growth in output, efficiency, and capacity for commercialized utility service industries, such as electricity, telecommunications, and water, particularly if they have significant private investment (physical or financial).

The standard institutional solution to handle these infrastructure industry issues is to introduce an independent regulatory agency, operating within a clearly defined legal framework.2 The agency is intended to provide the high-quality institution that permits and fosters sustained growth in capacity and efficiency in the utility service industries, particularly the network elements. So, whether country X has a high- or a low-quality institution is determined primarily by the quality of governance of the regulatory agency (conditional on the quality of governance for the economy as a whole). Developing economies with high-quality regulatory agencies (as measured by regulatory governance) should attract more sustained investment into their utility service industries and at a lower cost of capital. The regulated utilities should also have higher efficiency and growth rates.

The perspective just outlined is at the heart of the recent literature on regulatory governance for utility service industries, particularly focusing on developing and transition economies. It is set out in Levy and Spiller (1994)—which draws explicitly on North (1990)—and in subsequent papers (see, among others, Smith 1997; Stern and Holder 1999; Noll 2000). But there have also been apparently disappointing outcomes of regulatory (and electricity) reforms. The many case studies, though illuminating, do not allow reliable generalizations. Only recently has there been formal statistical or econometric testing of the view—and policy—that better regulatory governance increases investment and efficiency in the electricity industry.

This article provides an econometric analysis of the relationship between the quality of regulatory governance and the level of generation capacity per capita for electricity supply industries in 28 African, Asian, Caribbean, and Latin American countries over 1980–2001, controlling for privatization and competition.

2. An independent regulatory agency is not the only way of providing the necessary institutional support either in theory or in practice (see Domah, Pollitt, and Stern 2002). In addition, an independent regulator may be combined with a high or low degree of reliance on contracts and courts. A major issue is whether low-income developing economies have the human and other resources to sustain independent regulatory agencies, particularly regulatory agencies with a significant degree of discretion. Nevertheless, an independent regulatory agency has become the standard recommended solution to the private investment problem for utilities in the same way as an independent central bank has become the standard recommended solution to handle commitment and time inconsistency problems in monetary policy (see Stern and Cubbin 2004).
This article first outlines the underlying economic issues and the main institutional design considerations and related recent research. It then outlines the modeling approach. The main econometric results from static and dynamic models are examined, followed by a discussion of endogeneity and causality issues.  

I. UNDERLYING ECONOMIC ISSUES, INSTITUTIONAL DESIGN, AND IMPLICATIONS FOR EMPIRICAL ANALYSIS

The main focus here is the inability of governments to make credible and binding commitments about utility pricing to sustain private investment while retaining decisionmaking powers. The discussion concentrates on commercialized utilities facing genuine budget constraints, particularly where private investment and private finance are important, with a focus on regulatory governance (autonomy and accountability) rather than regulatory content (prices and investments). The underlying economic issue for utility regulation—as for monetary policy—is that governments, particularly at certain times, have a strong incentive to behave in a shortsighted and populist manner that reduces welfare over a medium- to long-term period.

Output Measures for Utility Regulatory Agencies

For utility service industries, there are two main output measures for utility regulation:

- The level and rate of growth of technical efficiency and productivity (and the quality of service).
- The level of capacity.

This article focuses on capacity in developing economies, testing the key policy objective of the World Bank and many countries in the sample: that significantly higher investment (and private investment) is the most important reason for promoting independent regulatory agencies in electricity and similar utility service industries (World Bank 1994, chapter 3). According to this view, following the establishment of a regulatory agency, there should be:

- Sizable increases in investment flows (domestic and foreign) to developing economy electricity industries.
- Larger increases with higher quality regulatory governance.
- Larger impacts as the regulatory agency gains experience and reputation.

3. See Cubbin and Stern (2004) for a much fuller version of the study, particularly the data and tables of results.

4. The problems for developing economy governments in making credible commitments to support new investment in the presence of major fixed costs arise in other contexts beside utility regulation. A good example is export taxes for exportable cash crops (see McMillan 2001). We are grateful to the editor for this observation.
Previous Literature

The empirical work here adopts and extends the fixed-effects panel data modeling used in the literature on the impact of regulation on telecommunications outcomes (see, for instance, Fink, Mattoo, and Rathindran 2002; Wallsten 2002; Gutierrez 2003).

The approach of Gutierrez (2003) is particularly relevant. He constructs a regulatory governance index for his sample of 22 Latin American and Caribbean countries. A seven-element index (derived from the Stern–Holder typology) is calculated from examining each country's telecommunications laws and changes in the laws. (This article's model for electricity outcomes adopts a similar approach and uses a "snapshot" four-element index as one regulatory variable.) Gutierrez (2003) finds statistically significant positive direct effects of his regulatory index on both tele-density and efficiency, in static and dynamic models and after testing for the endogeneity of regulation. The estimated effect of a 1 percentage point increase in the index on mainlines per 100 inhabitants varies somewhat depending on the precise model specification but is, in general, on the order of 20 percent.

For electricity, there are only a few (very preliminary) empirical studies of the impact of regulation, such as Zhang, Kirkpatrick, and Parker (2002) and part of Pargal (2003). They concentrate on generation capacity, as does this article, but find only weak effects (if any) of regulation. Their studies also have major problems in disentangling the effects of regulation from those of privatization and liberalization. Drawing on a 2001 study by Domah (see Domah, Pollitt, and Stern 2002 for details), this study had access to better data on regulatory governance and its variation across countries, but again data constraints confined the estimates to capacity models for generation rather than transmission, distribution, sales, or commercial losses.6

Regulatory issues are, of course, only one aspect of electricity industry reform. For a comprehensive discussion of electricity reform in developing economies, see Jamasb and others (2005).

II. Economic Rationale, Model Specification, and Modeling Issues

The modeling work reported here is concerned with whether better regulatory governance in developing economies increases the rated generation capacity per capita.

Economic Rationale

For the electricity industry in developing economies, explicit regulation focuses on providing sufficient supplies, and that typically means increasing investment

---

5. The authors are grateful to Preetum Domah for permission to use the information from his survey here.
and capacity. In some cases, this has been done by harnessing private ownership and domestic or foreign private investment. In others, a workable financial framework is provided for the electricity industry to develop by loosening the ties with government. A country might enact an electricity law giving various powers and duties to a ministry or independent regulator. Such changes can also increase public investment in infrastructure—say, by requiring state-owned electricity companies to operate more commercially and allowing them to use private debt finance for investment at a reasonable cost of capital.

Investment is encouraged once effective regulation supports a workable financial framework. If the electricity industry is privately owned, the owners have the prospect of earning a reasonable return on their investment. If publicly owned, the industry can become independent of tax revenue or continually increasing loans. In addition, an effective regulatory framework can encourage the growth of private investment and private finance within state systems, as in India and China in recent years.

In an unconstrained market economy, per capita generation capacity adjusts to demand, which depends on per capita income, the price of electricity, and such environmental factors as climate. The price of electricity will be determined in part by the efficiency of the sector, with efficiency depending on regulatory factors and energy sources, such as coal, gas, hydro, and oil. Many developing economies with a traditional, vertically integrated, and state-owned electricity sector will be constrained not so much by market demand but by the availability of a continuing subsidy. In electricity markets with implicit or explicit subsidies, capacity constraints arise because of inadequate government revenues for electricity investment or subsidy payments or insufficient revenue flows to support viable private investment or commercial debt obligations—or both.

Electricity generation models for unconstrained markets typically find per capita GDP to be the major determinant of electricity demand (and thus of generation capacity). Therefore, per capita GDP is included in the model here as well as other control variables found to be statistically significant in previous studies of developing economy infrastructure industry, such as the share of industry in value added, country debt levels, and country economywide governance indicators. Variables for electricity privatization and competition are also included.

An effective regulatory framework can be expected to reduce the constraint on the operation of the market, increasing supply and moving the outcome closer to the market equilibrium. The better the governance of the regulator, the greater the expected increases in capacity and electricity supply.

**Model Specification**

For developing economies, with supply-constrained electricity, improved regulatory governance is expected to raise the equilibrium generation capacity levels. But the adjustment to the new equilibrium is very likely to take some time to achieve.
This suggests a long-run static model of the following form, specified below in panel data format:

\[
\log(ELCAPPCT)_{it} = (a_0 + a_i) + a_1 \log(GDPPC)_{it} + a_2 \text{RegVar}_{it} + a_3 X_{it} + \nu_{it}.
\]

\(\log(ELCAPPCT)\) is the log of per capita electricity generation capacity in gigawatts, \(a_0\) is a constant term, \(a_i\) is a time-invariant country-specific fixed effect, GDPPC is real per capita national income in 1995 U.S. dollars, \(\text{RegVar}\) comprises one or more of the regulatory governance variables, \(X\) is a vector of other possibly relevant sectoral and country-level control variables, and \(\nu_{it}\) is an error term.

In all cases, the variables are defined for \(i = 1, \ldots, I\) countries over \(t = 1, \ldots, T\) time periods.

The \(X\) vector of control variables for this equation might well include domestic fuel/hydro source availability and a variety of other country-specific economic or institutional variables. But both of these are expected to be captured largely by the country-specific fixed effects. Similar arguments apply to institutional and country governance effects since country rankings on these indicators tend to be fairly stable over 10–20 year periods. Also explored is whether either privatization or competition affects generation capacity growth; both direct and indirect effects are investigated (such as interactions between these variables and the regulatory variables). And on the basis of previous studies of electricity demand, \(a_1\) would be expected to be close to but probably less than 1.8

Equation (1) is a static representation of the model, which provides evidence on long-run equilibrium effects. Some dynamic error correction models are also considered, which provide evidence on the adjustment time path and separate short-run adjustment effects from long-run equilibrium effects.

To ensure that the modeling yields estimates of supply responses, the sample is confined to countries with unsatisfied demand for electricity throughout 1980–2001—that is, to developing economies in Africa, Asia, and Latin America. Developed economies and European transition economies are excluded because both have significant planning margins or unused capacity for some if not all years of the period.

Data

The sample is of 28 developing economies, for which there is complete (or near complete) generation capacity data for 1980–2001. This gives a longer panel than is usually available for such studies, greatly reducing the econometric problems associated with short panels. But because of some missing observations, the panel is unbalanced.

Of the 28 countries in the sample, 15 are in Latin America, six in the Caribbean, five in Africa, and four in Asia. They include large countries (Brazil

---

7. GDP is on an exchange rate basis rather than a purchasing power parity basis.
Cubbin and Stern

and India), small countries (Jamaica), middle-income countries (Chile and Mexico), and poor countries (Ethiopia and Sudan).

The full list of countries and summary regulatory characteristics is in appendix table A.1. Of the 26 regulatory reforms listed, only six are before 1995.

The dependent variable in the regressions is per capita generation capacity by country and year, derived from U.S. Energy Information Agency data on generation capacity by country for 1980–2001 (appendix figure A.1). There are some significant decreases (Nigeria and Nicaragua 1990–95) as well as large increases (Paraguay 1985–88). Generation capacity changes tend to be lumpy, so that the dependent variable does not obviously exhibit common or stable trends. (Note that the series does not distinguish between publicly and privately owned generation capacity.)

The key independent variables for this study are the regulatory variables available — data for each country on the existence (or absence) of:

- An electricity (or energy) regulatory law.
- An autonomous or ministry regulator.
- License fee or government budget regulatory funding.
- Free or mandatory civil service pay scales for regulatory staff.

Each is measured by a 0 or 1 dummy variable. The dating of the switch from 0 to 1 on the appropriate variables (subsequently maintained at a constant level) is derived from the date of enactment of a primary electricity reform or regulatory law (except where other information provided a known, superior alternative). So, the effect of age of the regulatory agency as well as its existence can be investigated to estimate alternative measures of the impact of regulation. Given the time needed to establish a functioning regulatory entity, the start date for regulation is taken as the year following the enactment of the law.

The regulatory variables in the index are all measures of formal attributes of regulation. No comparable data are available on the informal, practical qualities of electricity regulation, such as the transparency and quality of regulatory processes. The necessary omission of data on these characteristics may lead to biased estimates and standard errors. And unlike Gutierrez (2003), there is no time dimension on changes in formal governance attributes after the enactment of the primary electricity/energy regulatory law.

The Domah data set, suitable for a preliminary investigation of the impact of regulation, is not ideal. It suffers from an absence of data on the informal, practical aspects of regulation, such as the length of tenure of regulatory agency heads or commissioners.

9. See Stern and Cubbin (2004), pp. 30–32, where preliminary simulation results based on the Stern–Holder data set suggest that omitting data on the informal, practical aspects of regulation can lead to coefficients being underestimated by about 5–10 percent and a similar underestimate of t-values.

10. For a fuller description of the data and a range of descriptive statistics, see Cubbin and Stern (2004), section IV, p. 19.
Although much of the regulatory activity took place in the second half of the period, the earlier period is important for establishing benchmark pre-reform levels of generation capacity and for reducing some of the biases that can arise in the use of short panels. Of the total number of country-sample years, 21 percent were with an autonomous regulator and 31 percent with an electricity or energy regulatory law. By the end of the period, only two of the countries had not enacted an electricity law and there were nine with a ministry regulator operating under a law.

A key feature of the regulatory data is that the correlation between the four regulatory variables is (not surprisingly) very high (table 1). In addition, all countries with an autonomous regulator had an electricity law, as did all the countries with license fee funding.12 This high collinearity between the regulatory variables presents estimation problems, discussed in the next section.

For privatization and competition, the Henisz–Zellner–Guilen (HZG; Henisz, Zellner, and Guillen 2004) electricity data are used.13 The data on privatization provide information on the year all countries introduced one of the following: minority privatization of their electricity industries, majority privatization of their electricity industries, and total privatization of their electricity industries.

The HZG data on competition include a variable for the year private firms were legally allowed to generate electricity for resale. But this does not necessarily mean that such electricity sales were important or even took place, and half the countries in the sample had this attribute over the whole period. More seriously, this variable provides no information on the market structure of generation or wholesale electricity purchasing. But it is the only consistently available "competition" variable for developing economies over the period.

The other main data source was World Bank data, including the World Bank World Development Indicators (as for per capita GDP and population) and the Kaufmann governance indicators.

| Table 1. Correlation Matrix of Regulatory Variables |
|-----------------|-----------------|-----------------|-----------------|
|                  | Electricity Law | License Fee     | Autonomous      | Non-Civil Service |
|                  |                 | Funding         | Regulator       | Pay Scales        |
| Electricity law  | 1.000           |                 |                 |                  |
| License fee funding | 0.849         | 1.000           |                 |                  |
| Autonomous regulator | 0.783         | 0.703           | 1.000           |                  |
| Non-civil service pay scales | 0.783       | 0.551           | 0.443           | 1.000            |

Source: Authors' analysis based on data described in the text.

11. For a fuller description of the data and a range of descriptive statistics, see Cubbin and Stern 2004, Section 4, p 19.
12. Uruguay was a partial exception, introducing license fee funding, 3 years before its law came into force.
13. We are grateful to Professor Henisz for permission to use these data.
III. Econometric Results

This section covers results from a static, long-run model, dynamic models, and endogeneity and causality issues.

Econometric Results for Models of Generation Capacity: Static Model

An ordinary least squares (OLS) equation was estimated as a baseline. However, all the equations reported here are modeled using a fixed-effects estimator. Moving from OLS to a fixed-effects model reduced the standard error of the regression by more than half. Given the nature of the underlying model, a fixed-effects model would be expected to be more appropriate than a random-effects model. For some of the equations, this assumption was tested using the Hausman test and the random-effects model was consistently rejected in favor of a fixed-effects model.

The individual regulatory variables are sizable and with high t-values, though the coefficient on civil service pay has the opposite sign to the one predicted (table 2). Debt and industry control variables had consistently low t-values and are dropped in subsequent regressions.

<table>
<thead>
<tr>
<th>Table 2. Basic Static Generation Capacity Model Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory Variables</strong></td>
</tr>
<tr>
<td>Real GDP per capita (log)</td>
</tr>
<tr>
<td>Electricity law</td>
</tr>
<tr>
<td>Autonomous regulator</td>
</tr>
<tr>
<td>License funding of regulator</td>
</tr>
<tr>
<td>Civil service pay scales nonmandatory</td>
</tr>
<tr>
<td>Debt payments as a proportion of national income</td>
</tr>
<tr>
<td>Industry value added as proportion of GDP</td>
</tr>
<tr>
<td>Estimation method</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
</tr>
<tr>
<td>Standard error of regression</td>
</tr>
<tr>
<td>F-statistic</td>
</tr>
<tr>
<td>Durbin-Watson</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

Dependent variable = log(electricity generation capacity per capita).

Note: Numbers in parentheses are t-statistics.

Source: Authors' analysis based on data described in the text.
A sample average country is estimated to increase per capita generation capacity in the long run by 18 percent through enacting an electricity law. In this equation, as elsewhere, the long-run elasticity of per capita electricity generation capacity to per capita GDP is estimated at about 0.70–0.85. But the equation clearly fits the data well and provides powerful initial support for the importance of good regulation for generation investment.

The problem with the results in table 2 is that the high level of collinearity between the regulatory variables implies that the coefficient estimates on the individual effects are likely to be biased upward when taken in isolation. This conjecture is confirmed when all four regulatory variables are included in a single regression: the coefficient on the electricity law variable rises to 0.27, and all the other variables become insignificant. Omitting the law variable led to the funding variable becoming significant but with less than a 1 percent reduction in the standard error of the regression.

These results provide strong evidence that the high multicollinearity between the regulatory variables significantly affects the coefficient estimates when included in combination. The standard statistical solution to this problem is to estimate a model using principal components—to better identify the effects of the individual governance elements. The results showed that only the coefficient estimate of the first principal component (accounting for 76 percent of the total index variance) was statistically significant at the 5 percent level, with a t-value of 3.8. The loadings of the individual components in the first principal component were broadly similar to one another, and the loading on the electricity law element was the highest.

The problem with principal components is that the results do not necessarily have any economic rationale. So, the preferred solution is to assemble the four regulatory variables into a regulatory index and to use that index as an explanatory variable. This procedure, used in the Gutierrez (2003) telecommunications regulatory study, has been used extensively in the literature on the economic impact of independent central banks (see Geraats 2002 for a recent survey).

The standard procedure, adopted here, is to use a simple additive index. The index takes the values 0, 1, ..., 4 for each country in each year depending on whether the country scores 1 or 0 on each of the four regulatory variables. However, as pointed out by an anonymous referee, this procedure imposes the restriction that each of the variables included in the index has the same proportionate impact on the dependent variable. This is a strong and highly debatable assumption, but at least our index is derived from direct observation rather than from impressionistic indicators. In view of this and other concerns, the summary results are reported estimating alternative current and lagged versions of various regulatory variables.

The fixed-effects equations were all estimated with per capita GDP and the HZG privatization variables as controls (table 3). In all the regressions, the coefficient

14. Estimation with a Guttman hierarchical index produced very similar results to those using a simple additive index.
Table 3. Static Generation Capacity Model with Alternative Regulatory Variables

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Lagged 3-Year Regulatory</th>
<th>3-Year Plus Electric Act</th>
<th>Lagged Electric Act in Age of Regulator</th>
<th>Quadratic in Age of Regulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulatory index(t)</td>
<td>0.022 (1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulatory index(t-3)</td>
<td>0.041 (2.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent or ministry regulator in place 3 years or more</td>
<td>0.164 (2.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricity act(t)</td>
<td></td>
<td>0.116 (2.6)</td>
<td></td>
<td>0.044 (3.6)</td>
</tr>
<tr>
<td>Electricity act(t-3)</td>
<td></td>
<td></td>
<td>0.143 (2.8)</td>
<td>-0.0018 (-2.8)</td>
</tr>
<tr>
<td>Age of regulator squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Estimation method

<table>
<thead>
<tr>
<th>Adjusted R²</th>
<th>Fixed effects</th>
<th>0.95</th>
<th>0.97</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard error of regression</td>
<td>Fixed effects</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>Fixed effects</td>
<td>0.17</td>
<td>0.29</td>
</tr>
<tr>
<td>Number of observations</td>
<td>Fixed effects</td>
<td>557</td>
<td>476</td>
</tr>
</tbody>
</table>

Dependent variable = log(electricity generation capacity per capita).

Note: Numbers in parentheses are t-statistics.

Source: Authors’ analysis based on data described in the text.

on per capita GDP was about 0.7, with a t-value of 8 or more. Here, attention is concentrated on the results concerning the alternative measures of regulatory governance.

Apart from the unlagged index, all the regulatory variables are positive and significantly different from 0 at the 5 percent level or better. The lagged variables (including the 3-year plus dummy) are all larger than the contemporaneous indicators and have higher t-values. The implication that it takes time to build up the effect on regulation is supported in the age quadratic model, where the maximum regulatory impact is estimated to be at about 14 years.

The dynamic modeling concentrates on the 3-year lagged index in column 2 of table 3 and the 3-year plus regulator dummy in column 3. The lagged regulatory law has the best overall fit, but given its collinearity with the other regulatory variables, it is less satisfactory overall as a descriptor of the regulatory framework.

The lagged index variable implies a maximum impact on per capita generation capacity of 16 percent—the same as for the 3-year regulator dummy. Note that the latter includes ministry regulators as well as autonomous ones. But particularly toward the end of the period, many ministry regulators were operating with powers and duties specified in a regulatory law.
One concern about the results in tables 2 and 3 is the low value of the Durbin–Watson statistic. In the static form of the model, this would not be expected to lead to biased coefficient estimates, but it may lead to overestimated t-statistics. As a preliminary test, the column 3 model was estimated incorporating a first-order autoregressive process. The coefficient on the lagged residuals was 0.79, with a t-value of 42.8. The estimated coefficient on the lagged regulatory index was both positive (0.02) and statistically significant at the 5 percent level (with a t-value of 2.1) and was also positive and statistically significant for per capita GDP. The estimated Durbin–Watson was 1.74.

The result suggests that autocorrelation in the static model does not significantly affect either the coefficient estimates or their statistical significance. This is explored more fully when explicitly considering the results from dynamic models.

Econometric Results for Models of Generation Capacity: Static Model Privatization and Competition

Incorporating the Domah data on privatization and competition into the model did not produce any significant effects, but those data had major weaknesses. This section reports estimates, using the better HZG data, of the relevant coefficients in fixed-effects regressions with the 3-year plus regulatory dummy variable as the measure of regulation. In all cases, the estimated overall fit of the equation and the coefficients on per capita GDP and the 3-year plus regulatory dummy variable was within 1 percent of those in table 3. Results using the regulatory index were similar.

On competition, the results reported above and others show consistent and significant long-run effects on generation capacity levels of about 10–15 percent. But the competition variable as defined provides no information on the amount of private company electricity generated for sale—let alone whether it was from an independent power producer selling to a single buyer or more from liberalized wholesale markets. So, the results are more likely to indicate a degree of country commitment to electricity reform rather than any genuine economic impact of competition in generation markets.15 Sixty-one percent of the observations scored 1 on this variable.

On privatization, the results varied. Unsurprisingly, there was no evidence that minority privatization had any significant effect on generation capacity levels. But neither did full privatization—though it applied to only 2.3 percent of observations. There was some evidence, weak in statistical significance, that majority privatization had a long-run positive effect on generation capacity levels of about 8–10 percent.

In tests for interaction effects between the regulatory variables and both the privatization and “competition” variables, none was significant at the 10 percent level or better.

---

15. In many countries, including the U.K., the legal right for new entrants to generate for resale was the first step in electricity reform—but achieved little or nothing in itself.
Econometric Results for Models of Generation Capacity:
Static Model — Country Governance Effects

On country governance, the Kaufmann indexes for rule of law and corruption by country in 1998 were first included as explanatory variables. The corruption index was never significant in the fixed-effect regressions at the 5 percent level or better, either as a separate variable or when interacting with regulatory variables. Estimated coefficients on the Kaufmann rule of law index were never statistically significant in their own right, but they sometimes approached significance when interacting with the regulatory variables (Kaufmann, Kraay, and Mastruzzi 2005).

The Kaufmann rule of law index was, however, highly significant in an OLS equation — leading to the nonsignificance of the electricity regulatory variable. This last result (together with the relative constancy of the cross-country rankings of general country governance indicators over long periods) is a major reason why the estimated fixed effects may well capture a large part of the countrywide institutional differences. On this last point:

- No statistically significant correlation was found between the fixed effects and the Kaufmann rule of law index.
- But a sizable and statistically significant interaction term between the regulatory index and the Kaufmann rule of law index was found in a random-effects specification (a coefficient of 0.07, with a t-value of 2.3).

These results provide interesting pointers to the role of governance effects in the model but are clearly far from conclusive.

A further test included in the static model the values of the World Bank CHECKS index, a time-varying index of political risk. The index "counts the number of veto players in a political system, adjusting for whether these veto players are independent of each other, as determined by the level of electoral competitiveness in a system, their respective party affiliations, and the electoral rules” (Beck and others 2000, p. 28). The index yields a minimum score in the absence of an effective legislature. The index score then increases linearly with the addition of subsequent veto points. The index is available for all sample countries for almost all years 1980–2000.

Including this index in the regressions rather than the single-year Kaufmann index is a much stronger test of whether the estimates of our electricity regulatory governance effects are biased because of the absence of explicit country governance measures. If country governance measures vary over time, as appears to be the case, the impact is not captured either by the country-specific fixed effects or by inclusion of the Kaufmann index for 1998.

Including the CHECKS index in the equations confirms the robustness of the estimates reported in table 3. The estimated coefficient on the CHECKS index is

16. For further details on the definition of the index, see Beck and others (2000). We are grateful to an anonymous referee not only for the suggestion that we include the index in our modeling but also for providing the data for the countries and time periods in our sample.
correctly signed (positive) and about 0.015–0.020, with t-values of about 2.1 (an increase of 1 point on the CHECKS index increases expected per capita generation capacity in the long run by 1.5–2.0 percent). Lagging the CHECKS index variable has a very small impact on the value on the coefficients for itself or any other variable.

The coefficient estimate on the (3-year lagged) electricity act was slightly reduced by adding the CHECKS index (from 0.20 to 0.16), but its significance remained high with an estimated t-value of 4.1. For the 3-year lagged regulatory index, the coefficient on the regulatory variable was reduced from 0.058 to 0.049, but the t-value again remained high at 3.5. In addition, the coefficients on privatization and "competition in generation" were virtually unchanged from the results in table 4. This leads to the conclusion that the impact of the electricity regulatory governance variables is genuine and not just a proxy for variations in country governance.

Interaction effects between the CHECKS index and the electricity regulatory governance variables were also tested for. The estimated coefficients were both small and not significantly different from 0 (with t-values of about 0.6).

Last, an equation was estimated for per capita generation capacity including the CHECKS index and per capita GDP but omitting any electricity regulatory variable. The resulting coefficient estimate for the CHECKS variable increased only slightly (to about 0.025), indicating that the collinearity between the electricity regulatory variables and the CHECKS index was very small.

These results show that both sectoral regulatory governance and country governance significantly affect the level of investment in per capita generation "competition in generation".

### Table 4. Static Generation Capacity Model: Alternative Privatization and Competition Variables

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>All privatization and Competition Variables</th>
<th>Minority and Majority Privatization Variables Only</th>
<th>50% or More Privatization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minority privatization</td>
<td>0.05 (0.57)</td>
<td>0.04 (0.49)</td>
<td></td>
</tr>
<tr>
<td>Majority privatization</td>
<td>0.09 (1.47)</td>
<td>0.12 (2.13)</td>
<td></td>
</tr>
<tr>
<td>100% privatization</td>
<td>0.09 (0.07)</td>
<td>0.12 (2.13)</td>
<td></td>
</tr>
<tr>
<td>Majority or full privatization</td>
<td></td>
<td></td>
<td>0.07 (1.24)</td>
</tr>
<tr>
<td>Competition (legal right to generate electricity for resale)</td>
<td>0.14 (3.05)</td>
<td>0.14 (3.03)</td>
<td>0.15 (3.41)</td>
</tr>
</tbody>
</table>

Dependent Variable = \( \log(\text{electricity generation capacity per capita}) \).

**Note:** Numbers in parentheses are t-statistics. Other independent variables in the regression are per capita GDP and existence of 3-year plus regulator.

**Source:** Authors' analysis based on data described in the text.
capacity, but that the impact of the sectoral variables is several orders of magnitude larger. Furthermore, the effects appear to be empirically separable at least for electricity, an issue returned to in the conclusion.

**Dynamic Models and Autocorrelation**

Given the nature of the investment planning and construction, lags would be expected to be quite long. Two main concerns were to establish whether the results, particularly for the regulatory variables, were genuine causal processes or merely spurious regression and to consider autocorrelation explicitly within a dynamic modeling framework rather than as a statistical autocorrelation “correction.” To test whether the estimated long-run static fixed-effects levels equations are genuine rather than spurious, check whether there appears to be a plausible adjustment process. The levels equation can be written as:

\[(2) \quad Y_{it} = \varphi_i + \beta G_{it} + \gamma R_{it} + \nu_{it}\]

which can be estimated as:

\[(3) \quad Y_{it} = f_i + bG_{it} + cR_{it} + \nu_{it}\]

where \(Y_{it} = \log(\text{electricity generation capacity per capita})\), \(G_{it} = \log(\text{GDP per capita})\), \(R_{it}\) is a regulatory governance variable, and \(f_i\) is the fixed effect for country \(i\).

Equation (3) calculates the implied steady-state equilibrium, or long-term value of \(Y_{it}\), which can be written as:

\[(4) \quad Y_{it}^* = \varphi_i + \beta G_{it} + \gamma R_{it}\]

Now, postulate a partial adjustment error correction mechanism, under which the actual value of capacity changes by a constant proportion of last year's deviation from the long-term value:

\[(5) \quad \Delta Y_{it} = Y_{it} - Y_{it-1} = -\lambda (Y_{it} - Y_{it-1}^*)\]

where \((Y_{it-1} - Y_{it-1}^*)\) is last year's deviation from equilibrium.

Equation (5) can be estimated by taking the residuals \(u_{it}\) from equation (3) and estimating:

\[(6) \quad \Delta Y_{it} = -\lambda u_{it-1} + \epsilon_{it}\]

An alternative procedure is to estimate directly a differenced version of the long-run relationship, including country-specific fixed effects.
Since the particular interest is in the size and significance of the regulatory variable, \( R \) impose the estimate of from the long-term levels equation (3) and estimate:

\[
\Delta Y_{it} = -\lambda (Y_{it-1} - \phi_i - \beta G_{it-1} - \gamma R_{it-1}) + \varepsilon_{it} \\
= \lambda \varphi - \lambda Y_{it-1} + \lambda \beta G_{it-1} + \lambda \gamma R_{it-1} + \varepsilon_{it}.
\]

Equations (6) and (8) yield alternative estimates of \( \lambda \), the speed of adjustment, which can be compared. In addition, alternative estimates of \( \gamma \), the impact of regulation, come from the levels equation (2) and from the associated differenced equation (8).

The validity of this procedure depends on the stationarity of the data generation process. The Pesharan–Shin W-statistic was used to test for stationarity. By applying this test to the differenced equation (8), with the regulatory index as a measure of \( R_{it} \), the test clearly rejects the presence of a unit root in the residuals with a \( t \)-value of -8.05.\(^{17}\)

The key results:

- The estimates of \( \lambda \), the speed of adjustment, were low at 0.12 but very similar between the levels and differenced equations, both with \( t \)-values of 8.9.
- The estimates of the impact of regulation in the differenced equation (8) were positive and significant, with \( t \)-values of 2.0 for the 3-year plus regulator and 3.2 for the regulatory index.
- The estimated long-run impact on per capita electricity generation capacity in the differenced equation was 24 percent for the 3-year plus regulator and almost 40 percent for the regulatory index.
- There was no evidence of serial correlation in the differenced equations (Durbin–Watson of 1.78). The overall fit of the differenced equations was good, with an adjusted \( R^2 \) of about 0.15 and F-statistics of 4.5 or higher.\(^{18}\)

Further estimated versions of these equations were produced using the 3-year lagged regulatory index but also including the HZG privatization and competition variables. These indicate a faster overall speed of adjustment, with an estimated error correction term of 0.24–0.27. But this still imply more than 5 years before half the regulatory effect on capacity appears (table 5).

The equations in table 5 demonstrate positive and statistically significant coefficients in both level and differenced equations not just for the (3-year lagged) regulatory index but also for majority privatization and competition.

17. Even in the corresponding levels equation, the Pesharan–Shin W-statistic does not appear to suggest nonstationarity in the residuals, implying that our generation capacity variable, GDP, and our regulatory variables are co-integrated. Very similar results were obtained on the unit root test with alternative definitions of the regulatory variable.

TABLE 5. Generation Capacity — Error Correction Models

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Levels 1</th>
<th>Differences 2</th>
<th>Differences 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log real GDP per capita</td>
<td>0.751 (10.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of regulatory governance</td>
<td>0.041 (4.454)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of regulatory governance (t−3)</td>
<td></td>
<td>0.0126 (1.97)</td>
<td></td>
</tr>
<tr>
<td>Lagged residuals from 1</td>
<td></td>
<td>−0.270 (15.44)</td>
<td></td>
</tr>
<tr>
<td>Error correction term</td>
<td></td>
<td></td>
<td>0.267 (15.40)</td>
</tr>
<tr>
<td>Majority privatization</td>
<td>0.125 (3.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majority privatization (t−1)</td>
<td></td>
<td>0.054 (2.95)</td>
<td></td>
</tr>
<tr>
<td>Legal right of independent power producer sales</td>
<td>0.120 (2.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legal right of independent power producer sales (t−1)</td>
<td></td>
<td>0.037 (2.82)</td>
<td></td>
</tr>
<tr>
<td>Estimation method</td>
<td>Fixed effects</td>
<td>Fixed effects</td>
<td>Fixed effects</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.974</td>
<td>0.413</td>
<td>0.416</td>
</tr>
<tr>
<td>Standard error of regression</td>
<td>0.205</td>
<td>0.076</td>
<td>0.076</td>
</tr>
<tr>
<td>F-statistic</td>
<td>581.6</td>
<td>11.34</td>
<td>10.33</td>
</tr>
<tr>
<td>Durbin–Watson</td>
<td>0.294</td>
<td>1.80</td>
<td>1.82</td>
</tr>
<tr>
<td>Number of observations</td>
<td>488</td>
<td>481</td>
<td>481</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are t-statistics.
Source: Authors’ analysis based on data described in the text.

The estimated long-run impact on per capita generation capacity from the privatization and competition variables, derived from the differenced equation in column 3, is 19 percent for the maximum regulatory index score, 20 percent for a majority privatization, and 14 percent for the legal right for competition.

These results are very similar to those from the static models in tables 3 and 4. They provide strong support for the hypothesis that the impact of regulation and privatization on generation capacity in developing economies is positive and sizable but that it takes some years to build up.

IV. Endogeneity and Causality in Generation Capacity Models

This section examines the endogeneity and causality issues that arise in the context of the model.19

19. We are grateful to Richard Gilbert and Jean-Michel Glachant for helpful discussions on these issues.
Endogeneity

Much of the literature on regulatory effectiveness expresses concerns about the endogeneity of countries choosing to have an independent autonomous regulatory agency and of the quality of that agency’s governance (see, for instance, Fink, Mattoo, and Rathindran 2002; Gutierrez 2003). Countries with better (unobservable) governance have better functioning regulatory agencies because, for example, they have socioeconomic characteristics that better support the rule of law, contracts, and commercialization. The problem is that it is very difficult to find good instruments— instruments that are both correlated with the suspected endogenous variable and uncorrelated with the error term—so that they can be treated as exogenous. The alternative is to try to model explicitly the decision to adopt regulatory reform. But this is difficult, and so far the results of such modeling have been interesting but not very successful (see, for example, Gual and Trillas 2004).

It is possible, however, to use a rank-based instrument to test for endogeneity and to derive an instrumental variable (rv) estimator to control for it.20 With this procedure, the coefficient on the residuals of the equation with the rank-based index in the basic static equation for per capita generation capacity has a t-value of 1.7, implying marginal evidence of endogeneity of the Cubbin–Stern regulatory index. But instrumenting the Cubbin–Stern index by using its predicted value in place of the actual value produces virtually identical results—an estimated coefficient of 0.047 with a t-value of 4.3 in the instrumented case as opposed to an estimate of 0.049 and a t-value of 4.0 in the noninstrumented case.

Like in Edwards and Waverman (2004) and Gutierrez (2003), the results show some weak evidence of endogeneity of regulatory governance quality but very little change in coefficient estimates after correcting for it.

Causality

Do the regulatory governance coefficient estimates reported here have any causal interpretation? That the core results are maintained even with 3-year lags on the regulatory index and with sophisticated dynamic modeling strongly suggests an underlying causal relationship.

Even if they are not statistical artifacts arising from failures to address dynamics or endogeneity adequately, they may still be merely descriptions of a past set of events that cannot be applied to future electricity regulatory governance changes in sample countries—let alone to the introduction or development of electricity regulation in other developing economies.21


21. For the reasons stated in sections I and II, we would not wish to claim that they are applicable to countries with an excess supply of generation capacity at any time after 1980. This would exclude the Central and East European countries, the Commonwealth of Independent States, and almost all Organization for Economic Co-operation and Development countries.
One reason this issue arises is that the regulatory literature derived from Levy and Spiller (1994) emphasizes country-specific constitutional, legal, economic, and political differences as crucial for the success or failure of utility regulation. So, a highly reduced form that abstracts from all those issues may well fail to reflect the local issues that seem to be so important in practice.

The answer lies in country-specific fixed effects. With 28 countries each having up to 21 years of data, estimates can be obtained of the fixed effects that should capture most if not all of the factors identified by Levy and Spiller and the subsequent literature. The estimated impact of, say, enacting a regulatory law with an autonomous regulator in Chile or Sudan (both countries in the sample) will be very different. That impact is the combination of the predicted effect of the relevant regulatory variables and each country’s predicted fixed effect. The fixed effect for Chile is strongly positive, relative to the sample average, that for Sudan is strongly negative.

In other words, the coefficients reported here are "highest common factor" estimates of the impact of regulatory governance indicators, where the fixed effects not just control for but effectively "wash out" all the Levy and Spiller and similar factors, including non-time-varying cross-country differences in country governance. This means that the regulatory governance effects reported here are not just average cross-country sample effects—they refer to a country with average scores on country-specific fixed effects, including country-governance fixed effects. Moreover, they are the impacts that one might expect, looking forward, for a country:

- With an average country-specific fixed effect.
- Implementing an average-quality law.
- Establishing an average-quality autonomous regulator—and so on.

For such a country, one might expect that implementing a best quality electricity regulator would increase per capita generation capacity in the long run by about 15–25 percent. But countries that reduce their political risk scores also increase their expected per capita generation capacity levels over and above the impact of sectoral regulatory governance impacts (and vice versa).

Two policy implications emerge. First, the quality of overall country governance matters considerably for the impact of regulation on outcomes (as in the rule of law). Second, countries cannot expect to achieve the gains estimated here by enacting low-quality regulatory laws or introducing autonomous regulatory agencies with very low staffing levels. The corollary is that the potential gains from introducing an electricity regulator could be significantly higher than the average for countries with good overall governance that deliberately try to introduce best-practice regulatory agencies and practices.

The higher potential gains from good regulatory institutions follow not just from the logic of the fixed-effect modeling—they are confirmed by the significance (and orthogonality) of the political risks index in our model.\footnote{This view is also supported by the strong impact of the Kaufmann rule of law index variables in an \textit{OLS} equation and the highly significant, positive coefficient of an interactive governance-regulation variable in a random-effects equation.}

V. Discussion of Results and Concluding Comments

The results of this study seem to provide a broadly consistent picture that a regulatory agency with good governance characteristics can not only improve regulatory outcomes in principle but that it seems actually to do so in practice. For electricity supply industries in 28 developing economies over 1980–2001, an index of regulatory governance is a consistently positive and statistically significant determinant of per capita generation. The results, using fixed-effects estimation methods, are similar to those found for telecommunications in developing economies (Gutierrez 2003).

The main findings for per capita generation capacity in developing economies:

- The effects of enacting a regulatory law, having an autonomous regulator, and using license fees to fund the regulatory agency were each positive and statistically significant at the 1 percent level.
- Averaging over developing economies regulatory agencies, the estimated long-run impact of the preferred measures of regulation is on the order of 15–25 percent, all other things equal, after controlling for country-specific fixed effects.
- The effects on per capita generation capacity are robust to modeling with a dynamic error correction model and to instrumental variable modeling to allow for potential possible endogeneity biases.

On privatization and competition, there was some evidence of the effects on generation capacity of majority privatization and of competition (its legal introduction). But the effects of competition are almost certainly more a reflection of a country's commitment to electricity reform than of a genuine market effect. A positive and well-determined impact of majority privatization was found in the dynamic modeling.

The strength and robustness of the results were surprising. And since the regulatory changes typically took place in the mid- to late-1990s and the regulatory variables are most significant when included with a 3-year lag, the results are clearly not capturing just the Asian independent power producer boom or the Latin American privatization boom. But the recentness of the regulatory changes may account for the surprisingly high significance of the impact of passing a regulatory law relative to that of having an autonomous regulator. It will be interesting to see whether the estimated effect of having an autonomous regulator is higher in 5–10 years.
This article concentrated on the role of regulatory quality for capacity and investment in the electricity industry. The results are very similar in type to those previously found for telecommunications, and similar approaches could likely be used to examine the institutional underpinnings for investment in other infrastructure. The approach here might also be useful for exploring the institutional contribution to capacity expansion for other industries. That, however, would require long data sets to allow robust estimates of country-specific fixed effects—and country governance measures, including the estimation of sophisticated dynamic models, such as error correction models.

This article presented evidence suggesting that good regulatory governance does have a positive and statistically significant effect on some electricity industry outcomes in developing economies—notably per capita generation capacity levels. But it did not examine why this is so.

Examining why and how regulation operates to improve outcomes is not a task that obviously recommends itself to econometric analysis. At this stage, that is better pursued by case studies, with econometric work concentrated on whether the results reported here are confirmed in subsequent analysis with superior data, particularly for regulatory practice, privatization, and competition.

The results here are consistent with the literature on the role of institutions in economic growth and with good country governance. Indeed, the evidence suggests that good country governance and specific regulatory effectiveness are mutually reinforcing. Both the quality of the electricity regulatory framework and the quality of country governance (as measured by a political risk indicator) are strongly associated with higher capacity, but as one might expect, the sectoral variables have a markedly larger impact.

The key points are that regulatory agencies with better governance are:

- Less likely to make mistakes.
- More likely to correct mistakes speedily.
- Less likely to repeat mistakes.
- More likely to develop procedures and methodologies that involve participants and to develop good practice.
- More likely to copy and implement best practice from other countries.

All these attributes reduce uncertainties for commercially operating companies, particularly private and foreign companies. This is especially important to sustain and encourage long-lived, sunk investments in highly capital-intensive industries. Regulatory agencies that maintain good governance thus provide an effective underpinning for the operation of contracts and for the sound regulation of monopoly elements.

These conclusions are also likely to be relevant for the institutional underpinnings of other sectors where raising sustained levels of long-lived or sunk investments is important.
APPENDIX: COUNTRIES IN SAMPLE AND PER CAPITA GENERATION CAPACITY DATA BY COUNTRY


Source: US Energy Information Agency and World Bank Database.
Figure A-1. Continued
### Table A-1. Countries in Sample

<table>
<thead>
<tr>
<th>Country</th>
<th>Year of Regulatory Start</th>
<th>Autonomous Regulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>1993</td>
<td>Yes</td>
</tr>
<tr>
<td>Barbados</td>
<td>n.a.</td>
<td>No</td>
</tr>
<tr>
<td>Bolivia</td>
<td>1995</td>
<td>Yes</td>
</tr>
<tr>
<td>Brazil</td>
<td>1997</td>
<td>Yes</td>
</tr>
<tr>
<td>Chile</td>
<td>Pre-1980</td>
<td>No</td>
</tr>
<tr>
<td>Colombia</td>
<td>1993</td>
<td>No</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Pre-1980</td>
<td>Yes</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>1999</td>
<td>Yes</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1997</td>
<td>Yes</td>
</tr>
<tr>
<td>El Salvador</td>
<td>1997</td>
<td>Yes</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>2000</td>
<td>No</td>
</tr>
<tr>
<td>Grenada</td>
<td>1995</td>
<td>No</td>
</tr>
<tr>
<td>India</td>
<td>1999</td>
<td>No</td>
</tr>
<tr>
<td>Indonesia</td>
<td>n.a.</td>
<td>No</td>
</tr>
<tr>
<td>Jamaica</td>
<td>1996</td>
<td>Yes</td>
</tr>
<tr>
<td>Kenya</td>
<td>2000</td>
<td>Yes</td>
</tr>
<tr>
<td>Malaysia</td>
<td>1991</td>
<td>No</td>
</tr>
<tr>
<td>Mexico</td>
<td>1996</td>
<td>Yes</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>1996</td>
<td>Yes</td>
</tr>
<tr>
<td>Nigeria</td>
<td>2001</td>
<td>No</td>
</tr>
<tr>
<td>Paraguay</td>
<td>n.a.</td>
<td>No</td>
</tr>
<tr>
<td>Peru</td>
<td>1994</td>
<td>Yes</td>
</tr>
<tr>
<td>Philippines</td>
<td>1988</td>
<td>Yes</td>
</tr>
<tr>
<td>Sudan</td>
<td>2001</td>
<td>No</td>
</tr>
<tr>
<td>Trinidad</td>
<td>Pre-1980</td>
<td>Yes</td>
</tr>
<tr>
<td>Uganda</td>
<td>2000</td>
<td>Yes</td>
</tr>
<tr>
<td>Uruguay</td>
<td>1998</td>
<td>Yes</td>
</tr>
<tr>
<td>Venezuela</td>
<td>1999</td>
<td>No</td>
</tr>
</tbody>
</table>

*Note:* Given the time necessary to establish a functioning regulatory entity, the year of the start of regulation was typically taken as the year after the enactment of the relevant law.

*Source:* Authors’ compilation based on data described in the text.

### References


An Empirical Analysis of State and Private-Sector Provision of Water Services in Africa

Colin Kirkpatrick, David Parker, and Yin-Fang Zhang

Under pressure from donor agencies and international financial institutions such as the World Bank, some developing countries have experimented with the privatization of water services. This article reviews the econometric evidence on the effects of water privatization in developing economies and presents new results using statistical data envelopment analysis and stochastic cost frontier techniques and data from Africa. The analysis fails to show evidence of better performance by private utilities than by state-owned utilities. Among the reasons why water privatization could prove problematic in lower-income economies are the technology of water provision and the nature of the product, transaction costs, and regulatory weaknesses.

The provision of safe and affordable water services is a priority for most developing economies. According to the World Bank (2003, p. 1), more than 1 billion people in the developing world lack access to clean water and nearly 1.2 billion lack access to adequate sanitation services. An estimated 12.2 million people die each year of diseases directly related to drinking contaminated water. The inclusion of a water access target in the Millennium Development Goals—to halve the proportion of people without access to safe drinking water by 2015—is a recognition of the importance of safe water supply in reducing poverty in the developing world (Calderon and Serven 2004).

A major cause of poor access to water services in developing countries is the inefficiencies of water utilities, which serve mainly urban areas. In many systems, as much as a third of production is lost, revenues are insufficient to cover operating costs, and the quality of the water is poor (World Bank 2004b, p. 220). Faced with the deterioration in water sector performance, and with
most water utilities under public ownership, donor agencies have advocated privatization to promote more efficient operation, increase investment and service coverage, and reduce the financial burden on government budgets (World Bank 1995).

A range of services including water supply have now been opened up to private capital (Harris 2003; World Bank 2003). Private participation has been less common in water systems, however, than in other infrastructure sectors, and the pace of reform has been slower and harder to sustain politically (World Bank 2004b, p. 220). Although privatization appears to have the potential to improve water services and meet the needs of the poor, these goals may be difficult to achieve. The technology of water provision (high fixed costs and location specificity) severely restricts prospects for competition, the transaction costs of organizing long-term concession agreements are considerable, and regulatory weaknesses suggest the need for caution. There is also the difficulty of balancing adequate returns to investors and ensuring that water services remain affordable to the poor.

The challenge for public policy is to meet both efficiency and social welfare objectives and to determine whether or to what extent privatization is critical to achieving the Millennium Development Goal for safe, accessible, and affordable water services. This article explores these issues by examining the impact of privatization of water services in Africa. It reviews the econometric evidence on the impact of water privatization and then, for a data set for African water utilities, uses statistical data envelope analysis and stochastic cost frontier measures to triangulate the evidence and assess consistency across results.

While data availability restricted the number of dimensions of performance that could be estimated, the results for cost efficiency and service quality fail to show that privatized water utilities perform better than state-run utilities. The data deficiency may explain the failure to identify better performance under private operation. However, special difficulties that face privatization and regulation in water services, also examined, likely play a role.

I. Evidence to Date

Private water suppliers have long been active as water vendors at the street level in all developing countries, but there was little privatization of piped water services before 1990 (Snell 1998; Collignon and Vézina 2000). Privatized services could be found in only a few countries, generally French-speaking former colonies such as Côte d’Ivoire that had inherited a reliance on private firms for water services, as is the practice in France. Between 1984 and 1990, only eight contracts for water and sewerage projects were awarded to the private sector.

1. As Bauer et al. (1998) emphasize, there can be greater confidence in comparative analysis if different measurements produce reasonably consistent conclusions.
worldwide, and cumulative new capital expenditure in private water services totaled less than $1 billion.

During the 1990s, however, there was a significant increase in water privatization, stimulated by donor agency pressures, and in 1997 private investment had risen to $25 billion (World Bank 2003). By the end of 2000, at least 93 countries had privatized some of their piped water services, including Argentina, Chile, China, Colombia, the Philippines, South Africa, the transition economies of Central Europe, and, among industrial countries, Australia and the UK (Brubaker 2001). Based on the World Bank Private Participation in Infrastructure (PPI) Database for the period 1990–2002, there were 106 such projects in Latin America and the Caribbean and 73 in East Asia and Pacific, but only seven projects in the Middle East and North Africa, and 14 in Sub-Saharan Africa. Latin America and the Caribbean and East Asia and Pacific together accounted for more than 95 percent of total investment. During 1990–2002, a small number of countries accounted for most of the privatization of water services, and within these countries, the totals were dominated by a few large contracts (table 1). In Argentina, one project accounted for $4.9 billion, or 20 percent, of all private investment in water services in Latin America, and in the Philippines five contracts accounted for 38 percent of the investments in East Asia.

Studies of privatization have found that competition is generally more important than ownership itself in explaining improvements in performance in developing countries (Zhang, Kirkpatrick, and Parker 2003; Parker and Kirkpatrick 2005). But whereas competition is feasible in telecommunications and parts of energy supply, such as generation, it is usually cost inefficient in the market for water services. While there is scope for introducing some competition into billing and metering and construction, replacement, and repair work within water services, competition in the provision of water supplies is normally ruled out by the scale of the investment in network assets that is needed to deliver the

<table>
<thead>
<tr>
<th>Country</th>
<th>Value (US$ Billions)</th>
<th>Number of Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>7.23</td>
<td>10</td>
</tr>
<tr>
<td>Philippines</td>
<td>5.87</td>
<td>5</td>
</tr>
<tr>
<td>Chile</td>
<td>3.95</td>
<td>13</td>
</tr>
<tr>
<td>Brazil</td>
<td>3.17</td>
<td>33</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2.75</td>
<td>6</td>
</tr>
<tr>
<td>China</td>
<td>1.93</td>
<td>44</td>
</tr>
<tr>
<td>Romania</td>
<td>1.04</td>
<td>3</td>
</tr>
<tr>
<td>Turkey</td>
<td>0.94</td>
<td>2</td>
</tr>
<tr>
<td>Indonesia</td>
<td>0.92</td>
<td>8</td>
</tr>
</tbody>
</table>

product. Moreover, even where competition for consumers might seem feasible, for example, at the boundaries between different water utilities, the costs of moving water down pipes is far higher than the costs of transmitting telephone calls and distributing electricity, placing a serious limitation on competition. Also, mixing water from different sources can affect water quality, an important consideration for domestic consumers and especially for water-using industries, such as brewing and food processing. In other words, the technology of water supply and the nature of the product severely restrict prospects for competition in the market and therefore the efficiency gains that can result from encouraging competition following privatization. This leaves rivalry under privatization taking the form mainly of competition for the market or competition to win the contract or concession agreement.

Evidence suggests that privatization in noncompetitive markets produces ambiguous results in terms of improving economic performance (Megginson and Netter 2001), highlighting the need for effective regulation of privatized utilities. The institutional requirements to ensure that privatized monopolies perform well—an effective system of state regulation and supporting governance structures—are likely to be missing in many developing countries (Parker and Kirkpatrick 2005). This represents a further difficulty to significantly improving performance in the short term through water privatization.

Privatized water services contracts can be set up as service contracts for specialized services (such as billing), management contracts and leases for existing facilities (operating existing facilities without new private-sector investment), concessions (requiring private-sector investment in facilities), divestitures (sale by the state of some or all of the equity in state enterprises), and greenfield investments (including build-operate-transfer schemes) (Johnstone and Wood 2001; World Bank 2004b). The most common are contracts under which private firms provide the services, but the government remains the ultimate owner of the water system and may remain responsible for some investment (OECD 2003). Of 233 water and sewerage contracts with the private sector during 1990–2002 included in the World Bank’s PPI Database, 40 percent involved concession contracts, accounting for 64 percent of total investment. Greenfield projects, less common, have often involved the building and operation of new water treatment plants, as in China, and build-operate-transfer schemes for water supplies have been used in Latin America and the Caribbean. Sales of state-owned water businesses to the private sector have been rare, accounting for only 15.6 percent of water projects and 8 percent of the total funds invested.

Although privatization of water services has occurred, it is important not to exaggerate its importance. Little more than 5 percent of the world’s population receives drinking water through private operators (OECD 2003), and since the Asian economic crisis of 1997/98, there has been a marked slowdown in infrastructure privatization in lower-income economies (Harris 2003). Moreover, the main forms that water privatization take raise concerns about the
transfer of risk from the public to the private sector, an issue discussed later in this article.

The case study evidence on water privatization presents a mixed picture, with some cases showing improvements in labor productivity, operating costs, reliability and quality of services, and share of the population served (Crampes and Estache 1996; Estache, Gomez-Lobo, and Leipziger 2001; Galiani, Gertlier, and Schargrodsky 2002; Shirley and Menard 2002; World Bank 2004b, pp. 252–57). Balanced against these positive findings is some evidence of higher water charges and public opposition leading to canceled schemes. The evidence is reviewed in Kirkpatrick and Parker (2005) and by Shirley (2002). The few published econometric analyses of the effects of water privatization in lower-income economies present little evidence that privatization has resulted in marked improvement in performance. Estache and Rossi (2002) compared private and public water companies in 29 Asia and Pacific region countries using 1995 survey data on 50 water enterprises (22 with some form of private-sector participation) from the Asian Development Bank. Adopting stochastic cost frontier modeling and applying error components and technical efficiency effects models, they conclude that efficiency was not significantly different in the private and state water sectors.

A study by Estache and Kouassi (2002), using a sample of 21 African water utilities during 1995–97, estimated a production function from an unbalanced panel data set and used Tobit modeling to relate resulting inefficiency scores to governance and ownership variables. The study found that private ownership was associated with a lower inefficiency score. However, only three firms in the sample had any private capital, and levels of corruption and governance were far more important in explaining efficiency differences between firms than was the ownership variable.

A study of water supply in Africa in the mid- to late-1990s by Clarke and Wallsten (2002) reported greater service coverage under private ownership. On average, they found smaller supplies for lower-income households (proxied by educational attainment) where there was a state-sector operator. While Clarke and Wallsten conclude that private participation leads to more supplies to poorer households, there may be offsetting service difficulties and higher charges when supplies are privatized. Drawing strong conclusions on the desirability of water privatization based on a single measure, such as service coverage, may be misleading. The analysis below uses a range of performance measures in an attempt to address this problem.

II. Assessing Performance in Privatized African Water Utilities

To advance understanding of the results of privatization in water services, data were taken from the Water Utility Partnership for Capacity Building in Africa's Service Providers' Performance Indicators and Benchmarking Network Project (SPBNET) database, which includes 110 water utilities in Africa. The data
collected, usually by questionnaire survey, relate mainly to 2000. The data set used for this study covers 13 countries and 14 utilities that reported private-sector involvement. However, not all of these firms could be included in each stage of the analysis because of incomplete data entries. The descriptive statistics for the sample are given in appendix table A.1.

Suppliers are categorized as either state owned or privately owned, a designation that captures the various institutional options for private-sector involvement in the water sector, including management and leasing contracts. Ideally, the form that private-sector involvement takes would be used to judge the degree of privatization, but the data source permits ownership to be modeled only as a binary variable. This limitation is shared by the earlier econometric studies mentioned above. More generally, the data set is characterized by heterogeneity, small sample size, and a small number of privatized firms. The data limitations mean that the results must be treated as tentative.

Conclusions on the impact of water privatization may be sensitive to the performance measure used. Therefore, to assess the impact of private capital on performance in water services, a range of performance measures were calculated. First, several statistical measures were computed from the data set:

- Labor productivity — ratio of labor costs to total costs, ratio of number of staff to number of water connections, and staff per million cubic meters of water distributed — to reflect the efficiency of labor use.
- Proportion of operating costs spent on fuel and chemicals — to reflect economies in nonlabor operating costs.
- Rate of capital utilization — to reflect capital stock efficiency.
- Average tariffs — to reflect the costs of services to consumers.
- Share of the population served, unaccounted-for water (water losses), and hours of availability of piped water per day — to reflect the quality of service to consumers.

Average figures were computed for both state-owned and privately owned water suppliers for between 61 and 84 utilities depending on the performance measure (table 2). On average, the private sector seems superior in production efficiency. Private-sector water utilities have higher labor productivity (lower ratio of staff to number of connections and amount of water distributed) and a lower share of labor costs in operating costs than do state-owned firms. The

2. The database (http://www.wupafrica.org) was developed with financial and technical support from the UK Department for International Development. Data for a few utilities relate to 1999 or 2001. Given the closeness of the years, all data are treated as applying to 2000 to adopt a cross-sectional analysis of performance.

3. Concession and management and lease contracts, together with privately owned assets, are categorized as private utilities. The utilities classified as private were cross-checked with the World Bank's PPI Database. The countries in the database with private water utilities are Cameroon, Cape Verde, Côte d’Ivoire, Gabon, Ghana, Kenya, Morocco, Nigeria, Republic of Guinea, Senegal, South Africa, Tunisia, and Zambia.
<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Average for State-Sector Operations (SD)</th>
<th>Average for Private-Sector Operations (SD)</th>
<th>F-test for Between-Group Difference in means (Probability Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labor productivity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor costs in total costs (percent)</td>
<td>29 (17)</td>
<td>21 (27)</td>
<td>1.45 (0.23)</td>
</tr>
<tr>
<td>Number of staff per 1,000 water connections</td>
<td>20.1 (19.9)</td>
<td>13.1 (14.4)</td>
<td>0.22 (0.65)</td>
</tr>
<tr>
<td>Number of staff per million cubic meters of water distributed</td>
<td>123 (519.7)</td>
<td>78 (151.8)</td>
<td>0.18 (0.68)</td>
</tr>
<tr>
<td><strong>Operating costs (percent)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share spent on fuel</td>
<td>20 (16)</td>
<td>11 (12)</td>
<td>0.44 (0.51)</td>
</tr>
<tr>
<td>Share spent on chemicals</td>
<td>17 (16)</td>
<td>4 (5)</td>
<td>2.37 (0.13)</td>
</tr>
<tr>
<td><strong>Capital</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital utilization (percent)</td>
<td>60 (21.6)</td>
<td>67 (21.8)</td>
<td>0.076 (0.79)</td>
</tr>
<tr>
<td><strong>Consumer charges</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average tariff (US$ per cubic meter)</td>
<td>168 (473)</td>
<td>305 (440)</td>
<td>1.9 (0.17)</td>
</tr>
<tr>
<td>Share of customers metered (percent)</td>
<td>60 (41.5)</td>
<td>79 (38.4)</td>
<td>1.45 (0.23)</td>
</tr>
<tr>
<td><strong>Quality of service</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of population served (percent)</td>
<td>63 (29.8)</td>
<td>64 (30.2)</td>
<td>0.22 (0.64)</td>
</tr>
<tr>
<td>Unaccounted-for water (percent of total)</td>
<td>34.8 (13.5)</td>
<td>29.0 (13.1)</td>
<td>0.63 (0.43)</td>
</tr>
<tr>
<td>Availability of piped water (hours per day)</td>
<td>17 (6.7)</td>
<td>16 (9.3)</td>
<td>0.25 (0.62)</td>
</tr>
</tbody>
</table>

The private sector is also more economic in its use of other inputs (fuel and chemicals) and achieves a slightly higher capital utilization rate of 67 percent as against 60 percent in the public sector firms.

Charges are on average 82 percent higher in the private sector, and more customers have their water consumption metered under privatized service.\footnote{Tariff figures have to be viewed with care since tariff levels are affected by public policy toward subsidies.} Metering water can increase revenues derived from consumers by linking payments to the volume of water used. The private sector also has lower water losses (probably assisted by greater use of metering), averaging 29 percent as against 35 percent for state-owned firms. Other measures of customer service suggest smaller differences between the private and state sectors, however. On average, state-owned firms supply piped water for 17 hours a day, while the private-sector records a slightly lower figure of 16 hours. The state and private sectors serve about the same share of the population in their areas, at 63 and 64 percent, respectively.

The standard deviations show a high degree of variance in performance within both the state and the private-sector categories for each of the measures, implying the need for care in interpreting conclusions based on average performance. Similarly, the F-test results for the difference in means for the public and private utilities’ performance ratios show that none are statistically significant (table 2). Also, data from the SPINET database suggest that privately owned water utilities in Africa are on average more than twice as large as state-owned utilities in terms of the total volume of water distributed (92 million cubic meters a day as against 36.4 million cubic meters a day) and have more connections in their systems (averaging 159,600 for private utilities and 94,500 for state-owned firms). This may partially account for the private utilities’ somewhat higher labor productivity.

To provide a fuller appraisal of relative performance, two further sets of performance measures were calculated, drawing on the same database, one using stochastic frontier analysis and one using data envelopment analysis.

Stochastic Cost Function Analysis

Because most water utility firms are required to meet demand and so are not free to choose the level of output, the analysis is based on a cost frontier instead of a production frontier. With output set exogenously, the firm is expected to minimize the costs of producing a given level of output. The coefficients of the cost function can be estimated by ordinary least squares (OLS) regression analysis, or a stochastic cost frontier model can be estimated by the maximum likelihood method. The stochastic cost frontier model decomposes the error term into stochastic noise ($\nu_i$) and cost inefficiency ($pi$).

Various distributions have been suggested for the inefficiency term in the stochastic cost function. Two of the most commonly used are the half-normal
distribution (Aigner, Lovell, and Schmidt 1977) and the truncated normal distribution (Stevenson 1980). The truncated normal distribution is a generalization of the half-normal distribution, obtained by truncating the normal distribution at 0, with mean \( \mu \) and variance \( \sigma^2 \). Preassigning \( \mu \) to be 0 reduces the truncated distribution to half normal. The appropriate model for estimation can be determined by testing the null hypothesis, \( H_0: \mu = 0 \). If the hypothesis \( \mu = 0 \) is rejected, the assumption of the truncated distribution is correct. If \( \mu \) is not significantly different from 0, a model assuming a half-normal distribution should be estimated instead.

As in the parameterization proposed in Battese and Correa (1977), \( \sigma^2 \) and \( \sigma^2_v \) are replaced by \( \sigma^2 = \sigma^2_\mu + \sigma^2_v; \gamma = \sigma^2_\mu / (\sigma^2_\mu + \sigma^2_v) \), to allow application of maximum likelihood estimates. The parameter \( \gamma \) lies between 0 and 1, with 0 indicating that the deviation from the frontier is due entirely to noise and 1 indicating that the deviation is due entirely to inefficiency. The superiority of a stochastic frontier can be tested by the null hypothesis, \( H_0: \gamma = 0 \). If the null hypothesis cannot be rejected, this indicates that the inefficiency term should be removed from the model, leaving a specification with parameters that can be consistently estimated using OLS.

The stochastic cost function has been widely specified as a Cobb–Douglas function or as a translog cost function. A generalized likelihood ratio test is used to determine whether a Cobb–Douglas function is appropriate. The result shows that the null hypothesis of the Cobb–Douglas specification cannot be rejected. In addition, Leamer’s extreme bound analysis shows that the range of the coefficients of the key variables for the Cobb–Douglas function is much smaller than that of the translog mode, confirming that use of the Cobb–Douglas specification is appropriate. To account for variable returns to scale, the quadratic term of the output variable is included. The coefficient is statistically insignificant, however.

A likelihood ratio test also points to the standard Cobb–Douglas specification.

As in the literature, the cost function is estimated using data on the cost level, the output level, and input prices. Operating and maintenance costs (COST) are used as the dependent variable in the cost frontier because adequate capital cost data are not available to compute total costs. An arbitrary cost function is therefore formulated that excludes the price of the capital input. Accordingly, the output and input variables were treated as focus variables and the control variables as doubtful variables. The bounds from the Cobb–Douglas model were much narrower than those from the translog model. In addition, in the translog model, the bound for the material input variable spanned zero. The results suggested that the coefficients for the Cobb–Douglas model were more robust than those for the translog specification.

Accordingly, the output and input variables were treated as focus variables and the control variables as doubtful variables. The bounds from the Cobb–Douglas model were much narrower than those from the translog model. In addition, in the translog model, the bound for the material input variable spanned zero. The results suggested that the coefficients for the Cobb–Douglas model were more robust than those for the translog specification.

Accordingly, the output and input variables were treated as focus variables and the control variables as doubtful variables. The bounds from the Cobb–Douglas model were much narrower than those from the translog model. In addition, in the translog model, the bound for the material input variable spanned zero. The results suggested that the coefficients for the Cobb–Douglas model were more robust than those for the translog specification.

5. Leamer’s extreme bound analysis was applied to the Cobb–Douglas and the translog specifications. Accordingly, the output and input variables were treated as focus variables and the control variables as doubtful variables. The bounds from the Cobb–Douglas model were much narrower than those from the translog model. In addition, in the translog model, the bound for the material input variable spanned zero. The results suggested that the coefficients for the Cobb–Douglas model were more robust than those for the translog specification.

6. The full results of these tests are available from the authors.

7. Estache and Rossi (2002) follow a similar procedure. In response to a referee’s comment that the inclusion of a fixed-capital measure might result in a misspecification of the cost function, an alternative specification of the cost function was tested that used the number of water treatment plants as a proxy variable for capital costs. The results for the ownership variable were unaffected.
personnel cost per employee (MP) is used to reflect the cost of labor, and material cost per unit of water distributed (MAT) is included as an additional determinant of noncapital costs. The amount of water distributed per year (WD) is included in the cost function as the output variable. Also included is a quality variable, measured by the hours of piped water available per day (QUAL).8

A number of control or environmental variables are also included to capture cross-country heterogeneity in the political, legal, and economic environment.9 Good governance, in the form of sound finance and regulatory systems and protection of property rights, has been found to be an important explanation for differences in economic performance (North 1990; Jalilian, Kirkpatrick, and Parker 2002; Kaufmann, Kray, and Zoido-Lobatón 2002), including in water services (Estache and Kouassi 2002). The freedom variable (FRD) developed by the Fraser Institute (http://www.freetheworld.com) is therefore included to capture wider governance or regulatory effects on performance in water utilities that might otherwise be attributed to ownership. An index of property rights (PROPERTY) is used as a measure of the quality of the investment environment (http://www.freetheworld.com). The fiscal balance variable (BALANCE) proxies the quality of macroeconomic management (http://www.freetheworld.com). A density variable, measured by population served per connection (DEN), drawn from the SPBNET database, is included because it plays an important role in defining the network infrastructure.10 Annual water resources per capita (WRS) is used as another control (WRI 2003). GDP per capita (GDP) is included to capture the extent of economic development (World Bank 2002). Finally, a dummy variable (ONS) is included to account for the effects of ownership on performance, taking a value of 1 if the utility had private capital.

All variables except the ownership variable and those in index or percentage terms are logged. In total, the estimations include 76 observations, including 10 private-sector operations. The program FRONTIER 4.1 is used to obtain the maximum likelihood estimates of the parameters and efficiency measures. The procedure for estimation is as follows. An error-component model is first estimated with the assumption of a half-normal distribution for the inefficiency term.11 To

8. Alternative quality indicators (unaccounted-for water and share of samples that fail to meet quality standards) were also tested, with similar results.


10. As pointed out by a referee, this density measure does not fully capture the dispersion of connections since it does not allow for the number of connections per building. Data on more common measures of dispersion, such as connections per kilometer of main lines or connections per square kilometer, were not available.

11. The error component model is the standard form of stochastic frontier model used in the literature. It decomposes the error term into stochastic noise and cost inefficiency. The truncated-distribution assumption yields $\mu = 0.47$, with a standard error of 2.56. A likelihood ratio test shows that the hypothesis $\mu = 0$ could not be rejected at the 10 percent level. Consequently, the results from the model with the half-normal assumption were adopted.
test the robustness of the results on ownership, a technical efficiency effects frontier is then estimated in which the inefficiency effects are expressed as a function of the ownership dummy variable.12

The value of $\gamma$ in the error-component model suggests a high ratio of the variance of inefficiency to the total residual variance (0.98; table 3).13 Analogously, the high value of $\gamma$ means that the stochastic frontier is superior to OLS modeling in explaining the cost structure of water utilities (both results are presented in table 3, for comparison). This is also confirmed by the generalized likelihood ratio statistic, which exceeds the critical value at the 1 percent level.14

The results of the half-normal error-component model show that the output variable, water distributed annually ($\ln WD$), has a positive and significant effect

---

**Table 3. The Stochastic Cost Frontier Results**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Error-Component Model (Half-Normal Distribution)</th>
<th>Technical Efficiency Effects Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ordinary Least Squares</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>Constant</td>
<td>4.17 (2.60)**</td>
<td>1.18 (1.6)</td>
</tr>
<tr>
<td>$\ln WD$</td>
<td>0.76 (13.22)**</td>
<td>0.88 (29.49)**</td>
</tr>
<tr>
<td>$\ln QUALI$</td>
<td>0.12 (0.81)</td>
<td>0.14 (1.88)**</td>
</tr>
<tr>
<td>$\ln MP$</td>
<td>0.26 (3.76)**</td>
<td>0.15 (4.33)**</td>
</tr>
<tr>
<td>$\ln MAT$</td>
<td>0.56 (8.20)**</td>
<td>0.65 (15.84)**</td>
</tr>
<tr>
<td>$\ln WRS$</td>
<td>-0.001 (0.01)</td>
<td>-0.09 (1.48)**</td>
</tr>
<tr>
<td>$\ln DEN$</td>
<td>-0.02 (0.44)</td>
<td>0.00003 (0.001)</td>
</tr>
<tr>
<td>$\ln GDP$</td>
<td>0.09 (0.85)</td>
<td>-0.01 (0.26)</td>
</tr>
<tr>
<td>$\ln FRD$</td>
<td>-0.13 (1.28)</td>
<td>-0.08 (0.22)</td>
</tr>
<tr>
<td>$\ln PROPERTY$</td>
<td>-0.11 (1.38)</td>
<td>-0.05 (4.03)**</td>
</tr>
<tr>
<td>$\ln BALANCE$</td>
<td>0.02 (0.64)</td>
<td>-0.004 (0.32)</td>
</tr>
<tr>
<td>$\ln ONS$</td>
<td>0.42 (2.00)**</td>
<td>0.15 (1.05)</td>
</tr>
<tr>
<td>$\delta ONS$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.98 (0.63E+07)</td>
<td>0.98 (0.21E+06)</td>
</tr>
</tbody>
</table>

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

Source: Authors' analysis based on data described in text.

12. The technical efficiency effects model can be used to investigate the determinants of technical inefficiencies among firms. The technical efficiency effects frontier is a stochastic frontier model that explicitly formulates technical inefficiency effects in terms of firm-specific factors. All parameters are estimated in a single-stage maximum likelihood procedure.

13. A referee has pointed out that the error term may be capturing more than just inefficiency where there is misspecification because of heterogeneity or measurement problems.

14. The critical value was obtained from Kodde and Palm (1986).
on operating costs. This is in line with expectations. Similarly, the variables of service quality (\(\ln\text{QUALI}\)), labor price (\(\ln\text{MP}\)), and material cost (\(\ln\text{MAT}\)) are all significant and correctly signed. The negative and statistically significant (at the 10 percent level) coefficient for the water resource variable (\(\ln\text{WRS}\)) is also consistent with expectations. The costs of water production and distribution would be expected to be lower in countries where water resources are abundant. The negative coefficients of income per capita (\(\ln\text{GDP}\)) and the freedom index (\(\text{FRD}\)) suggest that the operational costs of the utilities may be lower in countries that are wealthier, with sounder institutional governance. However, the effects are not statistically significant. More robust evidence of the influence of institutional development is provided by the property rights variable (\(\text{PROPERTY}\)), which shows negative and significant effects on the cost level, indicating that costs are lower in countries where property rights and therefore private investment are better protected. The impact of the government fiscal management measure (\(\text{BALANCE}\)) appears to be trivial. Contrary to expectations, however, the results for the density variable (\(\ln\text{DEN}\)) are statistically insignificant. The coefficient of the ownership dummy variable (\(\text{ONS}\)) is positive, suggesting that private ownership is associated with higher costs. However, the result is not statistically significant.

To assess the robustness of these results, a technical efficiency effects model is estimated in which the inefficiency term is expressed as a function of the ownership dummy variable. In this model, the inefficiency error, \(\mu_i\), has a mean of \(m_i\), and \(m_i = \delta \tilde{x}_i\), where \(\tilde{x}_i\) is a vector of variables that may influence the efficiency of a firm. This is taken as the ownership dummy variable in the estimation. The maximum likelihood estimation shows that the coefficient \(\delta_{\text{ONS}}\) is positive but not statistically significant (table 3). This finding is consistent with the ownership outcome from the error-component model.

**The Data Envelopment Analysis**

A data envelopment analysis was also undertaken.\(^{15}\) Water distributed is represented by the volume of output produced, and the number of hours of piped water available per day is used as the proxy for the quality of water services. (Unaccounted-for water and the share of samples that failed to meet the quality standards were also used as a proxy for quality of service, and the results were very similar.) An input-oriented variable returns to scale model was adopted to allow for variations in the size of the utilities.\(^{16}\) The analysis includes 66 utilities, nine of them private. The inputs are personnel cost per employee (because number of staff would not reflect the average skill level of staff\(^{17}\)), material cost per unit of water distributed, and number of water treatment works. The efficiency scores from the initial data envelopment analysis are regressed on the control variables (\(\text{DEN}, \text{WRS}, \text{GDP}, \text{FRD}\),

---

15. The authors are grateful to Catarina Figueira for assistance with the data envelopment analysis.
16. A constant returns to scale model produced a similar set of results but with lower overall scores.
17. The authors thank a referee for drawing this to their attention.
TABLE 4. Summary of the Data Envelopment Analysis Results

<table>
<thead>
<tr>
<th>Ownership</th>
<th>Utilities With 100 Percent Efficiency</th>
<th>Utilities With 90–99 Percent Efficiency</th>
<th>Utilities With 80–89 Percent Efficiency</th>
<th>Utilities With 70–79 Percent Efficiency</th>
<th>Utilities With Less Than 70 Percent Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>32 (53)</td>
<td>7 (12)</td>
<td>9 (16)</td>
<td>5 (9)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Private</td>
<td>6 (67)</td>
<td>1 (11)</td>
<td>1 (11)</td>
<td>1 (110)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Values are expressed as n (%).
Note: The lowest score, 52.5, was recorded by a state-owned water utility in South Africa.
Source: Authors' analysis based on data described in text.

Property, and Balance as defined earlier) using a Tobit model. Only population served by connection, DEN, and the property rights variable, Property, are statistically significant, and these two variables are included as control variables in a second-stage data envelopment analysis.18

The final data envelopment analysis results were tabulated by efficiency scores: the number of private and state utilities that achieved a score of 100 percent efficiency, 90–99 percent, and 80–89 percent (table 4).19 Significantly, state-owned firms help to form the efficiency frontier, suggesting that state ownership does not necessarily lead to low relative efficiency. More than half of the state-owned firms in the data set (32 of 57) were on the frontier. Six of the nine private operations included in the analysis populated the frontier. Therefore, the data envelopment analysis results appear to be consistent with the stochastic frontier analysis in suggesting that the efficiency performance of state-owned water firms in Africa is comparable to that of private enterprises. However, the results provide stronger evidence for possible higher relative efficiency in the private sector as a whole. For example, no utilities with private-sector involvement have less than 70 percent relative efficiency, and 67 percent of private as against 53 percent of state operations populate the frontier. It should be noted, however, that there are only nine private firms in the sample.

III. TRANSACTION COSTS AND WATER CONCESSIONS

With the results of the analysis presented here, it is interesting to consider why privatization of water services may be problematic in lower-income economies. The answer seems to lie in a combination of the technology of water provision

18. The inclusion of control variables in data envelopment analysis is widely practiced in empirical studies; see, for example, Ruggiero (1996, 2004) and Paradi and Schaffnit (2004). Wang and Schmidt (2002), however, are critical of this two-step procedure in data envelopment analysis.

19. Data envelopment analysis provides scores relative to peers with similar operating characteristics based on an estimated efficiency frontier. The resulting scores are relative, not absolute, scores. Therefore, a score of 100 percent does not imply absolute efficiency but merely efficiency compared with the other units in the analysis. Similarly, a stochastic cost frontier approach creates a frontier based on actual performances in the data set.
and the nature of the product, the costs of organizing long-term concession agreements, and regulatory weaknesses.

As explained, the technology of water supply and the nature of the product severely restrict prospects for competition in the market and therefore the efficiency gains that can result from encouraging competition following privatization. This leaves rivalry under privatization mainly in the form of competition for the market—competition to win the contract or concession agreement. However, serious problems can arise related to pervasive transaction costs in contracting for water services provision. These include the costs of organizing the bidding process, monitoring contract performance, and enforcing contract terms where failures are suspected (Williamson 1985). The economics literature suggests that such costs are likely to be high where there are serious information asymmetries at the time of contract negotiation.

Information imperfections are especially likely when contracts have to be negotiated to cover service provision over long periods of time. Many future events that could affect the economic viability of the contract and the acceptability of the service offering are unforeseen, and some may be unforeseeable. Concession agreements in water are typically negotiated for 10–20 years or more. Inevitably, therefore, the contracts will need to permit periodic adjustment of such variables as price, volume, and quality during the contract life. The contract will be incomplete in terms of specifying all of the contingencies that may trigger such adjustments and the form the renegotiation might take. This requires considerable skills on the part of both government and companies when operating water concessions, to ensure that the outcome is as mutually beneficial as possible.

The usual approach in water concessions is a two-part bidding process: selection of approved bidders based on technical capacity and then selection of a winner based on such criteria as the price offered and the service targets. However, the smaller the number of bidders, the greater the scope for actual or tacit collusion in bidding and the less competitive will be the bidding process. The evidence suggests that water concessions in developing countries are subject to small-numbers bidding (McIntosh 2003, p. 2). For example, in 2001 in Nepal, 18 companies expressed interest in operating a water contract in the first stage of the process, but only two serious bidders remained in the final stage (cited in Mitlin 2002, p. 17). In Argentina, there have usually been only a small handful of applicants for water concessions, typically between two and four (Estache 2002). To stimulate greater interest, concessions can include sovereign (government or donor agency) guarantees of profitability, but this introduces obvious moral hazard risks—with profits guaranteed, what incentive does the concession winner have to operate efficiently?

The literature on transaction costs also suggests that small-numbers bidding is a source of opportunistic behavior (Williamson 1985), leading to both adverse selection and moral hazard. Adverse selection takes the form of suboptimal contracts at the outset, as one of the contracting parties acts opportunistically to arrange especially favorable terms. Moral hazard occurs when one of the
contracting parties renegotiates the terms of the contract in its favor during the lifetime of the contract. During contract renegotiation, either the company or the government could be the loser, depending on the results.20

Guasch (2004) concludes that 75 percent of water and sanitation concession contracts in Latin America and the Caribbean were renegotiated significantly within a few years of being signed. In Buenos Aires, prices were raised within months of the start of the water concession (Alcazar, Abdala, and Shirley 2000). But even the ability to renegotiate terms may not be sufficient to overcome investor reluctance to participate in water privatizations, thus reinforcing the small-numbers bargaining problem. Difficulties arise especially when private investors fear that there is no long-term political commitment to water privatization (Rivera 1996). Moreover, cronyism and corrupt payments to win concessions may compromise the legitimacy of the privatization process. For example, in Lesotho, the Highlands Water Project was associated with allegations of bribes to government officials (Bayliss 2000, p. 14). Esguerra (2002, p. 2) shows how the water concessions in Manila were backed by the Philippines' two wealthiest families with support from multinationals: "It appears that the two companies' approach was to win the bid at all costs, and then deal with the problems of profitability later."

Studying concession contracts in developing countries, Harris and others (2003) find that water and sewerage concessions have the second highest incidence of contract cancellation after toll roads. This is not surprising given the substantial potential sunk costs in the water industry. Tamayo and others (1999, p. 91) note that the specificity of assets in the water industry is three to four times that in telecommunications and electricity. Handley (1997) stresses the problems caused by inadequate risk management techniques in developing countries. The preference by the private sector for the state to remain responsible for the infrastructure in water contracting reflects the desire of companies to minimize their sunk costs. Transaction costs in water concessions reinforce serious weaknesses in government-regulatory capacity in developing countries (Spiller and Savedoff 1999, pp. 1–2). For example, in India, there have been some local moves to attract private capital into water supply, notably in Gujarat, Maharashtra, and Tiruppur. But regulatory systems are underdeveloped, and in Tiruppur, they appear to be largely under the indirect control of the water operator (TERI 2003, pp. 171–221). As Mitlin (2002, pp. 54–55) concludes on the experience in Manila:

The gains [from privatization] may be less than anticipated because the assumption that the involvement of the private sector would remove political interference from the water sector was wrong. It may be that processes and outcomes have simply become more complex because the water supply industry now has the interests of private capital in addition to a remaining level of politicisation and an acute level of need amongst the poorest citizens.

20. For example, in the concession involving Maynilad in Manila, the company terminated the concession when it was refused a rate adjustment to which it felt entitled. By contrast, in Dolphin Bay, South Africa, the municipality believed that it had little alternative but to agree to an unplanned price rise when the private-sector supplier threatened to withdraw services (Bayliss 2002, p. 16).
To assess the effects of regulation on water privatization in Africa, the stochastic cost function analysis was repeated, this time incorporating a regulatory variable as a dummy variable alongside the freedom variable (representing wider good governance in a country). The SPINET database provides information on regulation of prices, water quality, and customer services. The different regulatory indicators are included separately in the regressions and are also combined into a composite regulation dummy variable to reflect the presence or absence of regulation in the water sector.

Regulation is expected to influence costs depending on its form. For example, a good regulatory regime should create more investor certainty and may reduce the costs of capital. Alternatively, regulation could raise costs by imposing higher and more expensive quality standards or by raising uncertainty for investors. The regression results show a negative sign for the composite regulation dummy variable and for the water quality and services dummy variables, suggesting that regulation lowered operating costs. However, these results are not statistically significant. The regulation dummy variable for tariff regulation is positively signed and statistically significant, suggesting that regulation of prices increased costs.

The findings from this stage of the analysis are therefore inconclusive.\textsuperscript{21} Regulation, both sector specific and as reflected in the general standards of governance in a country, are statistically insignificant. The single exception is related to tariff regulation, and the result is consistent with recent concerns that state regulation can raise costs (World Bank 2004a). However, the regulation variables used are far from ideal, and future research would benefit from developing a set of superior regulatory variables that more closely reflect the form of regulation rather than simply its existence.

IV. Conclusions

In principle, privatization has the potential to improve water services in developing countries, reversing decades of underinvestment and low productivity under state supply. However, the results, taken together, do not provide strong evidence of differences in the performance of state-owned water utilities and water utilities involving some private capital in Africa. While the data envelopment analysis results point tentatively to private-sector superiority, the stochastic cost frontier analysis provides some evidence that state-owned utilities have better cost performance, but the results are statistically insignificant. The descriptive statistics suggest no statistically significant differences.

The results therefore complement those of Estache and Rossi (2002), who also failed to find evidence that the performance of privately owned water

\textsuperscript{21} The detailed results can be obtained from the authors. A Tobit model was used to assess the impact of the regulation variables on the data envelopment analysis scores discussed earlier. The results were also statistically insignificant.
utilities in developing countries is superior to that of state-owned firms. Estache and Kouassi (2002) report a statistically significant result for the effect of privatization. However, this is based on data for only three privatized utilities in a total sample of 21 water utilities in Africa, and governance and institutional factors were found to be much more significant in explaining performance.

Admittedly, the results here contrast with the findings of Crampes and Estache (1996) and Galiani, Gertlier, and Schargrodsky (2002), who concentrate on service coverage. They conclude that privatization increased the number of people provided with safe water and sanitation. This study found no real difference in the share of the population served between private and state-owned utilities, but the limited availability of data precluded detailed exploration of this dimension of service. As with any study, the findings are dependent on the data used, and these were far from ideal. There is also the possibility that governments in Africa turned to private capital for the worst performing water utilities, thus making it less likely that the private sector would exhibit superior performance.

Other reasons why water privatization might prove problematic in lower-income economies were also identified and may help to explain why this and earlier studies have not found an unequivocally positive effect of private ownership on performance. Regulation dummy variables were included in the stochastic cost frontier model to shed further light on the importance of regulation, but most results were statistically insignificant. This outcome may reflect the crudity of the regulatory variables used, which simply measure the existence of water regulation not its impact on the management of utilities. Under conditions of perfect competition, perfect information, and complete contracts, ownership does not matter (Shapiro and Willig 1990) and the regulatory environment becomes inconsequential. However, none of these conditions applies to water services, and governance and regulatory variables are expected to be important in determining performance before and after privatization.

Finally, it needs to be stressed that providing affordable, safe, and accessible water to the poor is a fundamental priority for low-income countries. Policy-makers and regulators are likely to face difficult tradeoffs between ensuring that poor households are provided with affordable water supplies and allowing firms to charge prices high enough to recover costs and attract the foreign capital and technical capabilities needed to upgrade and expand the water supply network.

This study found that private operation of water facilities is associated with much higher average water charges and with greater use of water metering. But what are the impacts of this on water consumption and health? Water privatization usually means the involvement of a handful of major international companies. But what effect does this have on the development of indigenous ownership and regulation of socially important assets?22 Also, if privatization of

---

### Table A.1. Descriptive Statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of piped water (hours per day)</td>
<td>2.00</td>
<td>24.00</td>
<td>17.17</td>
<td>6.99</td>
</tr>
<tr>
<td>Labor cost per employee (in PPP units)</td>
<td>134.49</td>
<td>88,478.92</td>
<td>12,806.64</td>
<td>17,851.00</td>
</tr>
<tr>
<td>Material cost per unit water distributed (in PPP units)</td>
<td>0.00024</td>
<td>0.67</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>Number of connections (in thousands)</td>
<td>0.01</td>
<td>526.14</td>
<td>61.78</td>
<td>100.34</td>
</tr>
<tr>
<td>Total operating cost (in PPP units)</td>
<td>62,812.45</td>
<td>1,107,688,842.80</td>
<td>53,038,864.01</td>
<td>157,294,171.22</td>
</tr>
<tr>
<td>Total volume of water distributed per year (cubic meters)</td>
<td>8200</td>
<td>668,000,000</td>
<td>48,258,663.55</td>
<td>95,605,864.54</td>
</tr>
</tbody>
</table>

Note: PPP is purchasing power parity.
Source: Authors' analysis using data from the Water Utility Partnership for Capacity Building in Africa's Service Providers' Performance Indicators and Benchmarking Network Project database (http://www.wupafrica.org).

Water services leads to full cost recovery, is this outcome compatible with poverty reduction and what are the environmental implications of privatization? Clearly, water privatization raises a complex set of political economy questions that deserve fuller exploration than has been possible here because of data limitations.

### References


The Journal of African Economies is a vehicle to carry rigorous economic analysis, focused entirely on Africa, for Africans and anyone interested in the continent - be they consultants, policymakers, academics, traders, financiers, development agents or aid workers.

Each issue of the Journal of African Economies contains applied research of the highest standard, together with a comprehensive book review section, and a listing of the current working papers from around the world.

Now with Advance Access

Advance Access enables us to publish papers online soon after they have been accepted for publication and well ahead of their appearance in the printed journal, thus greatly reducing publication times and ensuring our readers stay in touch with the very latest developments in the field. See our website for more details.

JAE Online

Visit the Journal of African Economies website to read a free online sample issue and find articles of interest by searching across the contents of past issues using keywords and author's names. The website also contains advance access papers, subscription prices, submission information, and our free content alerting service.

www.jae.oxfordjournals.org
Accelerated Online Publication

This journal allows advanced access to accepted papers with Accelerated Online Publication.

The advantage for authors:
See your paper receiving citations faster than ever before.

Visit Oxford Journals's website at www.oupjournals.org, select the journal from the list to reach its homepage, and click on Advance Access - Accelerated Online Publication. Bookmark the page and regularly check for the latest accepted papers.
For information on advertising in this journal please contact:

Helen Pearson
Oxford Journals Advertising
PO Box 347
Abingdon, OX14 1GJ, UK

Tel: +44 (0) 1235 201904
Fax: +44 (0) 8704 296864

e-mail: helen@oxfordads.com

Information available online
Visit www.oxfordads.com
Forthcoming papers in

THE WORLD BANK ECONOMIC REVIEW

Volume 20, Number 2, 2006

PREFERENCE EROSION: A SYMPOSIUM

• Introduction
  Bernard Hoekman

• Doha Merchandise Trade Reform: What’s at Stake for Developing Countries?
  Kym Anderson, Will Martin, and Dominique van der Mensbrugghe

• Preference Erosion and Multilateral Trade Liberalization
  Joseph F. Francois, Bernard Hoekman, and Miriam Manchin

• Trade Preferences to Small Developing Countries and the Welfare Costs of Lost Multilateral Liberalization
  Nuno Limão and Marcelo Olarreaga

• Price Effects of Preferential Market Access: Caribbean Basin Initiative and the Apparel Sector
  Çağlar Özden and Gunjan Sharma

• Aid and the Supply Side: Public Investment, Export Performance and Dutch Disease in Low Income Countries
  Christopher S. Adam and David L. Bevan

• Infrastructure, Externalities, And Economic Development: A Study of Indian Manufacturing Industry
  Charles R. Hulten, Esra Bennathan, and Sylaja Srinivasan
THE WORLD BANK
1818 H Street, NW
Washington, DC 20433, USA
E-mail: wber@worldbank.org