THE WORLD BANK ECONOMIC REVIEW

Volume 13 May 1999 Number 2

Household Income and Child Schooling in Vietnam
Jere Behrman and James C. Knowles

Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture
Peter Lanjouw and Martin Ravallion

Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico
William F. Maloney

Tax Incidence in Madagascar: An Analysis Using Household Data
Stephen D. Younger, David E. Sahn, Steven Haggblade, and Paul A. Dorosh

Genuine Savings Rates in Developing Countries
Kirk Hamilton and Michael Clemens

Winners and Losers from the Privatization and Regulation of Utilities: Lessons from a General Equilibrium Model of Argentina
Omar Chisari, Antonio Estache, and Carlos Romero

Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence
Aslı Demirgüç-Kunt and Harry Huizinga

ISSN 0258-6770

Public Disclosure Authorized

Coming in the next issue of

THE WORLD BANK ECONOMIC REVIEW

September 1999
Volume 13, Number 3

- **Calm After the Storms: Income Distribution in Chile, 1987–1994**
  Francisco H.G. Ferreira and Julie A. Litchfield

- **Change in the Perception of the Poverty Line During the Times of Depression: Russia 1993–96**
  Branko Milanovic and Branko Jovanovic

- **Labor Market Analysis and Public Policy: The Case of Morocco**
  Julia Lane, Guillermo Hakim, and Javier Miranda

A symposium on education, including . . .

- **Can Private School Subsidies Increase Schooling for the Poor? The Quetta Urban Fellowship Program**
  Jooseop Kim, Harold Alderman, and Peter Orazem

- **Central Mandates and Local Incentives: The Colombia Education Voucher Program**
  Elizabeth M. King, Peter F. Orazem, and Darin Wohlgemuth

- **Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program**
  Emmanuel Jimenez and Yasuyuki Sawada

- **Schooling Outcomes in Philippine Elementary Schools: The Impact of Four Experiments**
  Jee-Peng Tan, Julia Lane, and Gérard Lassibille

THE WORLD BANK ECONOMIC REVIEW

Volume 13 May 1999 Number 2

Household Income and Child Schooling in Vietnam 211
Jere R. Behrman and James C. Knowles

Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture 257
Peter Lanjouw and Martin Ravallion

Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico 275
William F. Maloney

Tax Incidence in Madagascar: An Analysis Using Household Data 303
Stephen D. Younger, David E. Sahn, Steven Haggblade, and Paul A. Dorosh

Genuine Savings Rates in Developing Countries 333
Kirk Hamilton and Michael Clemens

Winners and Losers from the Privatization and Regulation of Utilities: Lessons from a General Equilibrium Model of Argentina 357
Omar Chisari, Antonio Estache, and Carlos Romero

Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence 379
Aslı Demirgüç-Kunt and Harry Huizinga

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2: 211-56

Household Income and Child Schooling in Vietnam

Jere R. Behrman and James C. Knowles

The stronger are the associations between household income and child schooling, the lower is intergenerational social mobility and the less equal is opportunity. This study estimates the associations between household income and children's school success in Vietnam. The estimates indicate that these associations are considerable. For example, the income elasticity of completed grades is five times the median estimate of earlier studies. Moreover, this association is strongest for grades completed per year of school, not *for completed grades, on which most of the previous literature has focused. There are some gender differences, the most important being a smaller association between income and grades completed per year of school for boys than for girls. This difference implies that schooling of girls is treated as more of a luxury (less of a necessity) than is schooling of boys.*

This article also investigates some ways in which policies relate to household incomes. School fees are progressive, but school fees are only about one-third of what households pay directly to schools and are a much smaller proportion of a household's total school-related expenditures. Total household expenditures paid directly to schools increase with household income less proportionately than do school fees alone, so the overall structure of such payments is less progressive than is the structure of school fees. Because school enrollment is positively related to household income, moreover, the structure of school fees is less progressive for the entire population than for the selected subset that has children enrolled in school. Further, the two school quality measures that have the strongest positive association with children's school success are much more available to higher-income households, meaning that higher-income households have greater school expenditures in part because they are obtaining higher-quality schooling and not because charges for the same quality schooling are progressive across income classes.

Schooling is widely seen as critical to the development process and poverty alleviation. Recent studies confirm that schooling is particularly important when complex new technologies and market options become available (for example, Rosenzweig 1995). Recently, many countries, including Vietnam, have under-

Jere R. Behrman is with the Department of Economics at the University of Pennsylvania, and James C. Knowles is an economic consultant based in Chapel Hill, North Carolina. The analysis that led to this paper was initiated with funding from the Asian Development Bank, through Contract COCS/95-000 under the Financing of Social Services Project: Vietnam (T.A. 2135-VIE) directed by Nguyen Van Tien of the General Statistical Office, Government of Vietnam. Behrman's subsequent input was funded by NIH 5-RO1-HD30907, NIH 1 RO1-AG-11725-01, and NSF SBR95-11955. The authors benefited from discussions with Benjamin E. Diokno from the University of the Philippines, who also worked on the project, and from the comments of Harold Alderman, three anonymous referees, and the editor.

gone considerable macroeconomic stabilization and market liberalization programs. The returns to schooling will probably increase following such programs. Therefore decisions about who is schooled now are likely to be critical in determining Vietnam's future economic growth and distribution of income.

A rising concern for many in Vietnam and other developing countries has been the possibility of greater inequality and reduced intergenerational social mobility under these economic reforms. Part of this concern is that family "dynasties" will be reinforced if children from higher-income households are more likely to receive more and better schooling, and thus reap greater gains from schooling in the future than children from lower-income households. Two different societies with the same income distribution at a point in time may be viewed as having different levels of social welfare if they have different degrees of social mobility. For example, Friedman (1962) argues that a given extent of income inequality that arises in a rigid system in which each family stays in the same position each period may be a cause for more concern than the same degree of income inequality that arises in a fluid system because of the great mobility and dynamic change associated with equality of opportunity.

Because of the concern that schooling could perpetuate social immobility and inequality, the recent policy-related literature has considered targeting public school resources toward children from poorer families (van de Walle and Nead 1995 provide examples and references). The concerns in Vietnam have been about whether family dynasties are becoming more powerful and whether schooling is targeted toward children from poorer households or if it is instead reinforcing the advantages of children from better-off households. Educational reforms have exacerbated these concerns (see World Bank 1996). The reforms are intended to make schools more efficient, but some of their components (such as the introduction of user charges) may affect children differently depending on their household income.

I. Indicators of School Progress

We examine four indicators of individual school progress for children ages 6–17 by income quintile in Vietnam: age when started school, grades passed per year of school, last completed grade, and exam score in last completed grade (table 1). Each of these indicators captures important and different dimensions of schooling from the point of view of individual children and their families.

Schooling in Vietnam

For a given extent of schooling the younger children are when they start school, the sooner they reap post-schooling returns and the longer they have to reap such returns. A few recent studies have emphasized the possible importance of the age when students start school (for example, Alderman, Behrman, Lavy, and Menon 1997; Glewwe and Jacoby 1995a; and Glewwe, Jacoby, and King 1998). In Vietnam this age is inversely associated with income: children from households in the

Table 1. *Means of Individual School Progress Indicators by Predicted Income per Household Member for Children Ages 6–17 in Vietnam, 1996*

	Income quintile				
Indicator	*1 (poorest)*	*2*	*3*	*4*	*5*
Age when started school	6.7	6.6	6.4	6.3	6.2
	(1.3)	(1.1)	(1.1)	(0.7)	(0.6)
	[2.8]*	[2.3]*		[1.3]	[6.1]*
Grades passed per year of school	0.81	0.85	0.87	0.90	0.95
	(0.20)	(0.20)	(0.19)	(0.19)	(0.12)
	[5.0]*	[1.7]**		[2.0]*	[8.8]*
Last completed grade	4.1	4.2	4.3	4.7	6.0
	(2.8)	(2.7)	(2.7)	(2.9)	(3.2)
	[4.1]*	[1.5]		[3.6]*	[11.5]*
Exam score in last completed grade	5.9	6.0	6.2	6.4	6.7
	(1.2)	(1.3)	(1.2)	(1.3)	(1.4)
	[2.6]*	[2.6]*		[2.2]*	[6.8]*

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Standard deviations are in parentheses, and absolute value of *t*-statistics for differences from the third quintile are in square brackets.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

first and second income quintiles (the poorest income quintiles) are significantly older, and children from households in the fifth quintile are significantly younger, than children in the third quintile (table 1). On average, children in the first quintile start school when they are half a year older than children in the fifth quintile.

For a given age at which a child starts school and a given extent of schooling, the more rapidly that child completes his or her schooling, the lower is the private cost of schooling in both direct monetary costs and opportunity costs, the sooner are post-schooling returns reaped, and the longer is the period in which to earn these returns. Grades passed per year of school are positively associated with income. The means for the first two quintiles are significantly below the mean for the third quintile, and the means for the fourth and fifth quintiles are significantly above that for the third quintile. On average, children from the first quintile pass about eight grades in a decade of attending school, while children from the fourth and fifth quintiles pass nine or more grades in a decade.

The last completed grade of schooling is the most emphasized indicator of individual school success in the literature—in fact, in most studies of the determinants and impact of schooling, it is the only indicator used. The last completed grade at the time of the survey is positively associated with income, with the mean for the first quintile significantly below, and the means for the fourth and fifth quintiles significantly above, the mean for the third quintile. On average, children from the fifth quintile had completed almost two more grades than children from the first quintile at the time of the survey.

Children who have completed the same number of grades may perform differently on examinations designed to measure how much they have learned. A few recent studies that have examined this question for developing countries report that cognitive test scores have an explanatory power beyond that of completed grades for estimated labor income relations (Alderman and others 1996b; Glewwe 1996; Knight and Sabot 1990; Lavy, Spratt, and Leboucher 1997; and Moll 1996). Examination scores conditional on the last completed grade of schooling are significantly and positively associated with household income in Vietnam; on average, children from the first two quintiles score significantly lower, and children from the top two quintiles score significantly higher, than children from the third quintile. Children from the fifth quintile score about 17 percent higher than children from the first quintile.

Thus the data suggest that in Vietnam there are systematic associations between important aspects of children's progress in school and household income: children from higher-income households do better in school according to each of the four indicators. These associations raise some important questions. Are such associations large or small? How do they differ among our four indicators? Are they largest for grades completed, as the emphasis in the literature might suggest? How sensitive are the estimated associations to details of their estimation, such as the definition of income and the treatment of students who are still in school? Do the associations differ for girls and boys? To what extent are the associations mitigated or reinforced by education policies? Is the structure of school fees progressive? If so, is this progressivity reinforced or offset by a household's other school-related expenditures? Is it reinforced or offset by the quality of schools attended by children from different segments of the income distribution?

Results of Past Research

Researchers have conducted numerous studies of associations between indicators of household income and schooling for other countries, although we are not aware of such studies for Vietnam. In Behrman and Knowles (1997) we review 42 studies, covering 21 countries (these are summarized in appendix A). Estimates for about three-fifths of the schooling indicators used in these studies show significant associations between household income and schooling. Of the cases for which we can estimate income elasticities, the median elasticity is 0.07. This number suggests that children from higher-income households do better in school than children from poorer households, although the magnitude of the effect is small. The estimates tend to be higher for samples with poorer households, and a number of the studies find small inverse associations between schooling and income. The largest elasticity estimates—those higher than 0.20—are for low-income regions (low-income during the period of the survey): Côte d'Ivoire, Ghana, Nepal, Taiwan (China) for the 1940–49 birth cohort, Northeastern Brazil, and rural Pakistan. But these are the only cases in which the estimates exceed 0.20.

It is not clear how informative these results are for Vietnam because they are from policy and market environments that are much different than those in Viet-

nam, where there has long been concern about equality of schooling. Moreover, most of this literature suffers from at least one of several limitations.

First, the income measures used in most of these studies are probably contaminated by measurement error as well as by the possibility that schooling is endogenous with other decision variables. If there is random measurement error in the income indicator, as would be the case if annual income is used when the relevant household resource constraint is really longer-run income, the estimated association between income and schooling would be biased toward zero. And if households make income and expenditure decisions simultaneously with schooling decisions, there may be a bias toward or away from zero, depending in part on which income measure is used. It may be away from zero, for example, if total income is used and households lower their total income (and possibly their expenditures) when they have school-age children by reducing child labor in order to increase schooling. But even if total income falls, expenditures may rise through dissaving to cover schooling costs, possibly generating the opposite bias. Likewise, the bias may be toward zero if adult (household head, father) income is the most important component of household income and if parents increase their work efforts and income to finance their children's schooling.

Second, studies may mask the true association between schooling and income because their estimations generally include a number of other household, community, and school controls that are correlated with income and may be proxying in part for the income association. To estimate multivariate causal relations and avoid omitted variable bias in the income coefficient estimate, it is important to control for all of the factors correlated with income that the underlying model suggests determine schooling and are predetermined with respect to income. Studies may claim to be estimating causal relations but in fact probably do not succeed in isolating the causal impact of income because of omitted variable biases and other problems. For example, innate ability may affect schooling investments in children, as is found in the studies summarized in Behrman and Rosenzweig (1999), may be correlated across generations through genetic endowments, and may be correlated with parental income, resulting in omitted variable bias. But to examine the extent of the associations between child schooling and parental household income, we do not want to control other characteristics that are correlated with income because the true association with income is likely to be misunderstood if such characteristics are included.

Third, most of the previous studies rely on only one schooling indicator, usually students' amount of time in school or their grade attainment. This narrow focus misses the possibility that there may be separate and important associations between income, on the one hand, and the age of starting school, grades passed per year in school, and performance on examinations in the last completed grade, on the other.

Fourth, in their estimates of associations between schooling and income, many studies that use data on school-age children do not control for the fact that, at the time the survey was conducted, some school-age children may not have started

school yet and others may not have completed school yet, or for other aspects of sample censoring. Because of sample censoring, data summaries, such as that in table 1, simple graphs, or other nonparametric summaries of the data may not reveal the extent of the associations with income.

Fifth, most of the existing studies do not explore whether there are important gender differences in associations between income and schooling, although other aspects of gender differences in schooling have received considerable attention. Finally, most studies do not address how education policies might be related to household income—for example, through targeted policies for school fees and a household's other school-related expenditures or through differences in school quality that may offset or reinforce differences in household income.

A few studies have addressed one or more of these limitations. For example, to address the limitation regarding the use of current annual income or expenditures, Glewwe and Jacoby (1995a, 1995b), Montgomery and Kouame (1993), and Tansel (1997) instrument household expenditures, and Alderman and others (1996a) instrument household income. With regard to the limitation of focusing only on completed schooling, Alderman Behrman, Lavy, and Menon (1997), Glewwe and Jacoby (1995a), and Glewwe, Jacoby, and King (1998) explore the relationship between income and the age of starting school, and Alderman and others (1996a, 1996b) and Alderman, Behrman, Khan, Rose, and Sabot (1997) investigate the relationship between income and cognitive achievement. With regard to the limitation of censored data, King and Lillard (1987) and Alderman and others (1996a, 1996b) control for right-censoring in completed years of schooling. Still, these examples and other studies that address one or a few of these limitations are a small proportion of the literature.

Our contribution in this paper is to explore the associations between household income and the four indicators of school progress for Vietnam. We use 1996 data from a cross-sectional household survey linked with community and schooling surveys that we conducted in collaboration with the Vietnamese General Statistical Office. We move beyond the previous literature by addressing the six limitations summarized above. Through this investigation we illuminate the extent and nature of associations between household income and schooling and related policies in the particular case of Vietnam, and we raise questions concerning what is known about such associations in other countries.

II. The Analytical Framework

We begin with a brief discussion of why schooling might be associated with household income. This discussion points to a number of possible reasons, as well as to the difficulty of disentangling association and causality from cross-sectional data and of determining whether such associations may reflect underlying inefficiencies.

If there were no unobserved differences between low- and high-income households, if schooling were purely an investment (with no current consumption as-

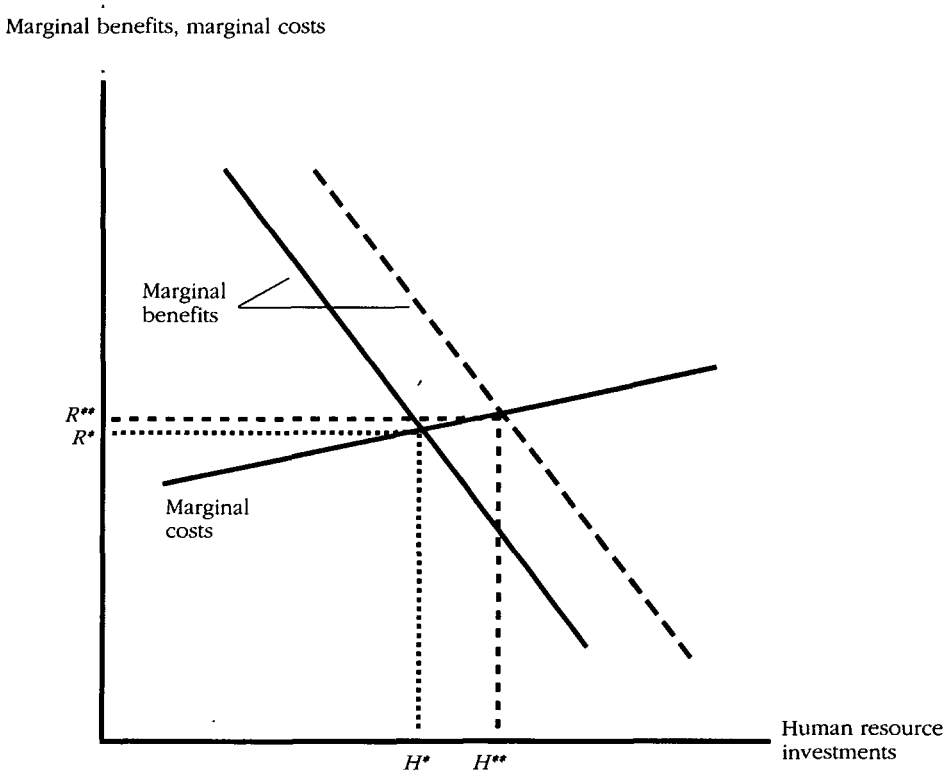
pects), if markets worked perfectly, and if the same prices prevailed in all markets, there would be no differences in schooling investments associated with income once controlling for any observed differences in household characteristics. Therefore it is useful to determine why there might be associations between household income and investments in schooling. The general reasons are that household income is proxying for correlated unobserved determinants of child schooling, such as innate ability, preferences, and family connections; household income is proxying for price variations in school inputs; and household income is playing a causal role in the presence of imperfect markets.

In addition to the investment aspect of schooling, spending time in school may be a current consumption activity that is associated with household income. We emphasize *current* consumption, such as spending time in school rather than spending time at home playing or watching television. Schooling may also affect future consumption (for example, by enriching reading as an adult), but because these effects are obtained in the future, current schooling for such purposes is an investment. If the current consumption of schooling has aspects that are normal goods, ceteris paribus, more household income leads to more schooling for that reason alone.

The relationship between schooling as an investment and household income is multifaceted and more complicated than the relationship between schooling as current consumption and household income. Becker's (1967) Woytinsky lecture on the determinants of human capital investments is a useful starting point for thinking in more detail about possible associations between parental household income and schooling investments. Within this framework schooling investments are made until the private marginal benefits of the investment equal its private marginal costs. Marginal private benefits (the solid benefits curve in figure 1) depend on the expected private gains (such as in wages or salaries or in enriched future consumption) from the human capital investment. The marginal private benefits curve is downward-sloping because of diminishing returns to schooling investments. We might expect diminishing marginal returns (at least at sufficiently high levels of investment) because a given individual has fixed genetic endowments (such as innate ability) and because human capital investments, such as those that extend the time in school, imply greater lags before obtaining postinvestment returns and a shorter postinvestment period in which to reap the returns. Marginal private costs may increase with human resource investments because of increasing opportunity costs of devoting more time to such investments and because of increasing marginal private costs of borrowing on financial markets. The private returns net of costs are maximized where private marginal benefits equal private marginal costs (H^* in figure 1).

If all markets function perfectly, there are no government interventions, and schooling is only an investment, then everyone invests in schooling until the expected rate of return from schooling equals the expected rate of return on alternative investments (at H^*), regardless of household income. In this case there are no or very few channels through which income may be associated with schooling

Figure 1. *Private Marginal Benefits and Private Marginal Costs of Human Resource Investments, with Marginal Benefits Dependent on Income*



(although there still may be some possibilities). But given the range of real-world market imperfections and government interventions, there are many reasons why household income may be associated with schooling, even if schooling is purely an investment.

To illustrate, consider what would happen in the presence of market imperfections. There are several explanations, originating in both policy and market failures (as well as reasons that would persist with perfect markets), why household income may be related to the marginal private benefits and costs of schooling investments and thus to schooling investments themselves. Current consumption effects could also generate associations between income and schooling (with the sign depending on the nature of the consumption effects). Some of these reasons reflect inefficiencies, such as those due to imperfect credit and information markets. Others reflect differing abilities that complement human capital investments or differing prices that are related to household income in different areas given positive transportation costs. Some reflect causal effects of income, such as current consumption demands. And some reflect associations with other variables, such as abilities that are correlated with income and transferred in part intergen-

erationally. With cross-sectional data of the types that are usually available, the relevance of many of these possibilities and the effect of causality compared with association cannot be sorted out conclusively.

A priori it would seem that market failures would be relatively common in Vietnam because of the country's low level of development, which was, until recently, exacerbated by the pervasiveness of relatively centralized command policies for many allocations. As a result, then, schooling investments may be more positively associated with household income in Vietnam than in most societies. However, at least the rhetoric of policy discussions in Vietnam has held that school policies should and do promote relative economic equality and reduce poverty. If the reality reflects the rhetoric, the policy effects would tend to work in the opposite direction.

Marginal Private Benefits Associated with Income

Figure 1 illustrates what happens when marginal private benefits for human capital are associated with household income. Each of the two curves depicted depends on a different income level ceteris paribus. If the (otherwise identical) individual is in the household whose income yields the dashed curve, the private incentives are to invest at H^{**}, which is higher than H^{*}. Why might marginal private benefits of schooling be associated with household income in the presence of government policies or market imperfections? There are several reasons.

First, public policies may affect households with different incomes differently. Policies may favor higher-income households by offering them higher-quality (or more accessible) schooling in response to their greater economic and political power or because prices of some important school inputs may be lower in areas where incomes are higher (for example, teachers may prefer to live and teach in high-income areas and be willing to do so at lower salaries than they would require in low-income areas). If school quality is positively associated with household schooling investments, the dashed marginal private benefits line would represent higher-income households. However, policies may favor poorer households if programs are designed to reduce inequality or to alleviate poverty by allocating better schooling to poorer households or if prices of some school inputs are lower in low-income areas (for example, land for schools). In such cases, if school quality is positively associated with household investments in education, the dashed marginal private benefits line would represent lower-income households.

Second, households may invest in children's education at home directly through tutoring or indirectly through improvements in their health and nutrition. If markets for these investments (or for financing these investments) are imperfect and the costs are lower for wealthier households, the marginal private benefits of schooling will be higher for wealthier households. For instance, the cost of helping with homework may be less for more-schooled parents than for less-schooled parents, and parental schooling is likely to be positively correlated with household income.

Third, children's genetic endowments, for which there are no perfect markets (marriage markets probably serve indirectly as imperfect markets for such endowments), may interact with schooling investments and be correlated with parental endowments that, in turn, are correlated with household income. These relationships arise because such endowments affect income directly and indirectly through parents' human capital stock, including their education. Behrman, Rosenzweig, and Taubman (1994, 1996) present evidence, using data on twins, that schooling investments respond positively to children's genetic endowments in the United States. Behrman and Taubman (1989) present estimates that variations in such endowments are consistent with most of the variance in child schooling for young adults in the United States. The enormous literature on the associations between adults' schooling and their household earnings is surveyed in Psacharopoulos (1994) and Rosenzweig (1995).

Fourth, households may make complementary investments in searching for a job and have contacts that affect their children's job search after completing schooling. If markets for financing such investments are imperfect and the costs are lower for higher-income households, in part because of more attractive possibilities for working in family enterprises and better connections for other employment opportunities, the marginal private benefits would again be higher for such households.

Fifth, higher-income households may have better information (in part because of better family enterprise options and better connections), given imperfect markets for information. As a result, they face less uncertainty about schooling investment decisions and, assuming constant risk aversion, therefore have higher expected marginal private benefits than poorer households.

Sixth, higher-income households may have lower risk aversion. Therefore in the presence of imperfect insurance markets or insurance with positive private costs, their private incentives would be to invest more in schooling than otherwise identical lower-income households.

And lastly, higher-income households may be better able to deal with stochastic events. For example, through their connections (perhaps facilitated by income transfers, including bribes), they may be better able to offset their children's bad performance on admissions examinations than poorer households can. They therefore have private incentives to invest more in schooling than otherwise identical lower-income households.

The first possibility (involving public policies) relates to endogenous policy choices, which, depending on the mechanism, could favor either higher- or lower-income households (see, for example, Rosenzweig and Wolpin 1986). In the other six cases higher-income households have private incentives to invest more in the schooling of otherwise equal children because they cope better with market imperfections, or higher-income households have unobserved characteristics that increase schooling investments and are associated with household income.

Marginal Private Costs Associated with Income

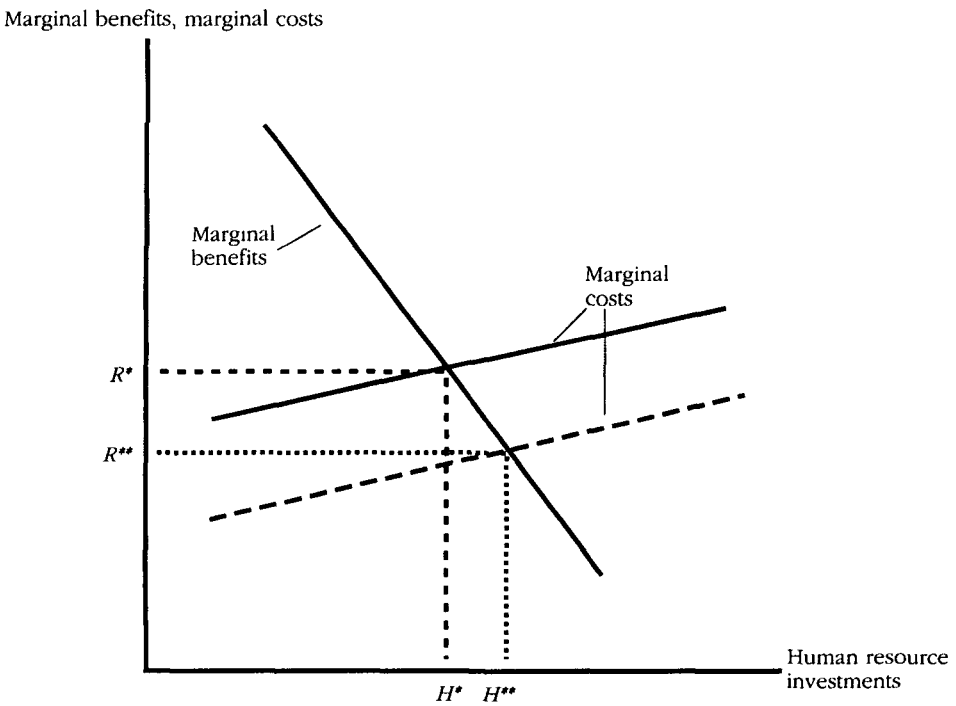
Figure 2 represents two different marginal private cost schedules for schooling investments that depend on household income. With the solid marginal cost curve, the private incentives are to invest at H^*, which is less than the privately optimal level of human capital investment for the dashed marginal cost curve, H^{**}. Why might marginal private costs for human capital investments be associated with household income in the presence of market imperfections? There are two reasons.

Because of capital market imperfections, particularly for human capital investments (in part because human capital is not recognized as collateral), the marginal private costs for such investments are particularly high for individuals from poorer families who cannot as easily finance these investments themselves. In this case the dashed line represents a household with higher income. In another case children from poorer households may be exempt from paying school fees, so that, ceteris paribus, the dashed line represents a poorer household.

III. Data

Our principal data source is the Vietnam Social Sector Financing Survey (VNSSFS), which we and Vietnam's General Statistical Office conducted in 1996 with fund-

Figure 2. *Private Marginal Benefits and Private Marginal Costs of Human Resource Investments, with Marginal Costs Dependent on Income*



ing from the Asian Development Bank. We conducted the VNSSFS in seven provinces, one from each of the country's administrative regions. The data set includes a survey of 1,905 households and a series of commune and school surveys. The VNSSFS is one module within a larger multiround survey that the General Statistical Office is carrying out called the Multi-Objective Household Survey (MOHS). The MOHS is an ongoing survey covering 45,000 households in 1,500 communes located in all 53 provinces of Vietnam. It collects, among other data, information on income and expenditures. The income and expenditure data that we use were collected retrospectively for the previous year (subsequent data were not available to us).

We used a subsample of communes surveyed in the MOHS (we chose three communes from each of three districts in each of seven sample provinces—a total of 63 communes, 19 urban and 44 rural). We administered the VNSSFS to the same 30 households in each sample commune that participated in the MOHS. In addition, we administered a community questionnaire to the Commune People's Committee Chairman in each commune and facility questionnaires to schools (and to other social sector facilities). Households were surveyed by the commune-level MOHS interviewers, who were residents of each commune, and the facility interviews were conducted by commune-level supervisors who were part of the supervisory staff of the MOHS.

The VNSSFS household questionnaire collected information on the use and financing of schooling. For all children in the household questions on schooling included the age they started school, whether they were currently enrolled, the age they last attended school if they were not currently enrolled, the last grade they attended (current grade if currently enrolled), and their last comprehensive examination score conditional on the grade in which the examination was taken (see table 1 for individual school progress indicators). For children currently enrolled in school, we obtained data on the household's school-related expenditures (and exemptions). The units of observation used in this study are primarily the 2,789 children in the 6–17 age range from the 1,844 sample households on whom we had complete data. Usable income data were not available for 3.2 percent of the households. This is a small percentage compared with many samples (for example, in recent years item nonresponses on earnings in the U.S. Current Population Survey have exceeded 20 percent). Of the 2,789 children, 2,203 (79 percent) were enrolled in school at the time of the survey.

Within the 6–12 age range there is some age-related variation in school enrollment and thus in data related to enrollment: among the 1,373 children in the 6–11 age range, 94 percent were enrolled at the time of the survey, while among the 1,416 children in the 12–17 age range, 65 percent were enrolled. Therefore, in addition to estimates based on the full sample, we present estimates that allow the parameters to differ for these two age groups. Because some households in the sample have more than one child, the regression estimates of the associations between schooling and income use the Huber correction for clustering at the household level.

Information on annual household income, expenditures, and assets is available from the MOHS. We merged these data with the VNSSFS household data. To characterize income for this study, we considered four alternative measures: annual household income per household member, annual household expenditures per household member, predicted household income per household member, and predicted household expenditures per household member (table B-1 gives the relations used to construct predicted income and predicted expenditures on the basis of longer-run characteristics). The mean annual income per household member by the income measures is 2.076 million Vietnam dong (D), which translates into 188 U.S. dollars ($) at 1996 exchange rates (table 2). The mean annual expenditures are D1.728 million or $156. The standard deviations are about a third larger for the two income measures than for the parallel expenditure measures. This difference is consistent with the theory that transitory income fluctuations are smoothed somewhat over time so that there is greater measurement error in letting income, rather than expenditures, represent a longer-run household resource constraint.

These standard deviations imply that the distributions of income and expenditures are more equal in Vietnam than in many societies, as is also reported in other studies. World Bank (1995b, table 30), for example, gives the percentage share of consumption by quintile for 22 low-income countries, including Vietnam. The share for countries in the lowest quintile ranges from 2.1 to 9.7 percent. The share for Vietnam is reported to be 7.8 percent, which is ninth highest among the 22 countries. The share for countries in the highest quintile ranges from 38.6 to 63.5 percent. The share for Vietnam is reported to be 44.0 percent, which is ninth lowest. World Bank (1995a) also discusses in some detail income distribution and poverty in Vietnam.

The standard deviations are about twice as large for the actual than for the parallel predicted measures, which is consistent with the possibility that the actual values have considerable measurement error. All but one of the correlations are between 0.50 and 0.80, meaning that most of these measures share substantial variance, although each has some independent variation. The one exception is the correlation of 0.99 between predicted income and predicted expenditures—apparently the two predicted variables are similar linear combinations of the

Table 2. *Summary of Alternative Income Measures*
(thousands of dong per capita)

Alternative income indicators	*Mean*	*Standard deviation*	*Correlations*			
			Income	*Expenditures*	*Predicted Income*	*Predicted Expenditures*
Income	2,076	1,740	1.00			
Expenditures	1,728	1,320	0.80	1.00		
Predicted income	2,076	924	0.53	0.51	1.00	
Predicted expenditures	1,728	684	0.52	0.51	0.99	1.00

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

same underlying variables. Because the correlation is almost perfect, we have nothing to gain in presenting estimates for both predicted values because they imply virtually identical elasticities. This high correlation is also consistent with the possibility that both income and expenditures represent the same underlying longer-run household resource constraint, with random measurement errors drawn from differing distributions (with greater variance in the distribution for income).

The predicted measures are preferable to the others if there is random measurement error or if current income or expenditure decisions are made simultaneously with current schooling decisions and it is the longer-run household resource constraint that is relevant for schooling decisions. If the relations being estimated are causal, then, in addition, the instruments used in the first-stage relations cannot be correlated with the disturbance term in the relation of interest, which is a condition that often is difficult to satisfy (for example, demands for children's schooling probably respond to children's unobserved ability, which is likely to be correlated with first-stage variables, such as parents' schooling). For the present purpose of characterizing associations between schooling and household income, there is no such condition to satisfy, because we want to find the full association between longer-run income and schooling whether that association arises because of effects of income or because of correlated unobserved characteristics in the disturbance. If some consumption smoothing is possible, expenditures are likely to be a better measure than income. Therefore we use predicted expenditures per household member for all of our estimates.

In addition to the VNSSFS household questionnaire, the community questionnaire collected information on the location of schools used by the population in the commune. Most of these schools were in the commune, although some upper-secondary schools were located outside. We administered school questionnaires to heads of 209 schools. These questionnaires collected information on characteristics of the personnel, current inputs, physical structures, and finances. We merged these data with the data on children in the 6–17 age range to give the nature of the school options available to each child, depending on the commune of residence.

IV. Associations between Parental Household Income and Children's Progress in School

Here, we measure how strongly the four indicators of children's school progress are associated with parental household income in order to ascertain the schooling advantages that children from higher-income households have over children from lower-income households. We estimate elasticities of each of these four indicators with respect to income per household member.

Preferred Elasticity Estimates

Our preferred estimates use predicted income to represent the longer-run resource constraint. They control for right-censoring for children who have not yet

started school or have not yet completed school and control for the mass point at 1 for grades completed per year of school (table 3). Right-censoring is relatively uncommon for the age of starting school indicator (only 8 of 2,625 observations) but is quite common for the last completed grade indicator (2,173 of 2,615 observations). The mass point at 1 for grades completed per year of school is considerable (1,442 of 2,615 observations).

Two major points about these estimates merit emphasis. First, these income associations are large. The estimate for last completed grade alone—0.353—is seven times the median found for the countries surveyed in Behrman and Knowles (1997). Thus in Vietnam children from higher-income households have a considerable advantage in schooling over children from lower-income households. These elasticities imply, for example, that compared with a child from a household whose income is one standard deviation below the mean, a child from a household whose income is one standard deviation above the mean starts school a quarter of a year earlier, successfully passes 94 percent rather than 80 percent of her or his classes, completes 2.2 years more of schooling, and scores 7 percent higher on examinations, controlling for grades completed (or about a third of the standard deviation in such scores). The mean completed grades of schooling for this calculation is based on 17-year-olds, whose mean is 7.2 years and whose enrollment rate is 0.34, under the assumption that further schooling will be equivalent to one more completed year of schooling for everyone enrolled at the time of the survey. Under plausible assumptions and in combination with other estimates from the data (presented in appendix C), the combined implications of these differences is equal to 13.1 percent of the present discounted value of lifetime income.

Second, the association between income and last grade completed is an important part of the story, but only part of the story. Most of the literature focuses on grades completed, often exclusively. In this case, under the assumptions elaborated in appendix C, changes in grades completed for the two households have a significant association (3.4 percent) with a change in the present discounted value of lifetime income. But this amount is only about a quarter of the total association of 13.1 percent. The association between income and grades passed per year of school accounts for about three-fifths of the total, more than twice as much as the last completed grade. Although such a comparison is based on particular assumptions for a particular sample, it suggests that focusing only on the last completed grade or on years of schooling may substantially underestimate the association between children's success in school and parental household income and that other studies should give more attention to the grades passed per year of school.

Sensitivity to Selected Aspects of the Estimation

For some of the indicators of individual school progress, the use of annual income with no control for censoring results in much smaller estimates of the association with household income. In particular, our preferred estimates are

Table 3. *Elasticities of Individual School Progress Indicators with Respect to Predicted Income per Household Member*

	Elasticity with respect to income				*Basic estimate as a percentage of estimate with*		
Indicator	*Basic: predicted income, control for censoring*	*Annual income, no control for censoring*	*Predicted income, no control for censoring*	*School/ commune fixed effects*	*Annual income, no control for censoring*	*Predicted income, no control for censoring*	*School/ commune fixed effects*
Age when started school	–0.051	–0.051	–0.050	–0.040	100	102	128
	(9.6)	(6.1)	(5.5)	(7.0)			
Grades passed per year of school	0.193	0.070	0.118	0.134	276	164	144
	(12.1)	(5.7)	(9.0)	(6.3)			
Last completed grade	0.356	0.240	0.178	0.353	148	200	101
	(7.4)	(9.0)	(4.9)	(6.1)			
Exam score in last completed grade	0.092	0.085	0.092	0.087	108	100	106
	(6.1)	(6.2)	(6.1)	(3.7)			

Note: Absolute values for *t*-statistics are given in parentheses beneath point estimates. The estimates are all significant at the 5 percent level. The underlying estimates are given in table B-2. For the basic estimates, censored normal regressions are used for age when started school and for last completed grade because of right-censoring, and upper-limit tobits are used for grades passed per year of school because of the mass point at 1.0.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

176 percent greater for grades passed per year of school and 48 percent greater for last completed grade than those obtained with no control for censoring (table 3). The differences are very small for age when started school and for exam score in last completed grade, however. This result suggests that, at least for grades passed per year of school and for last completed grade, if the true resource constraint is the long-run constraint represented by predicted income, the use of uncensored annual income considerably underestimates the income association because of transitory fluctuations in income and endogenous aspects of income. Thus we obtain higher income associations than most of the previous literature partly because we use predicted income with control for censoring. This explanation might hold if the use of predicted income is preferable because of endogeneity, but it cannot be the full explanation if the use of predicted income is preferable only because of measurement error due to transitory fluctuations. In that case we would expect similar changes in the estimates for all four of our indicators.

Control for censoring (or for the mass point at 1 for the second indicator) by itself increases some of the estimated income associations considerably. For the first three indicators for which we control for such censoring, our preferred estimates are, respectively, 2, 64, and 100 percent higher than what we obtain with the same income variable but without such controls (table 3). Therefore, if we do not control for censoring, we underestimate considerably the extent of the association with income for the last two of these variables because censoring is selective with regard to income (for example, observations on children from relatively high-income households are more likely to be censored for last completed grade or for successfully passing one grade every year they have been in school). Controlling for censoring does not much affect the estimate for age when started school, probably because censoring is relatively limited for this variable.

Controlling for school and commune fixed effects lessens somewhat the estimated associations between school progress and income. As noted earlier, one channel through which income may be associated with schooling success is through the relationship between schools and communities and household incomes. Estimates that incorporate the full association between income and schooling (including factors related to local communities and school characteristics) are from 1 to 44 percent higher than estimates that control for school and commune fixed effects (table 3). Therefore part of the association between income and our indicators, particularly for age when started school and for grades passed per year of school, reflects differences among schools and communes that are associated with household income. Further, the direction of the changes in the income associations when controlling for school and commune fixed effects is consistent with school resources being allocated to reinforce income differentials rather than to compensate for them.

Sex and Age Differences

In alternative estimates we allowed the income elasticities and the constants to vary with sex and age group (6–11 and 12–17 age brackets). The age that boys

Table 4. *Estimated Differences by Sex and Age in Elasticities of Individual School Progress Indicators with Respect to Predicted Income per Household Member*

Indicator	Elasticity with respect to income: Female, ages 12–17	Difference if: Male	Difference if: Ages 6–11	F-test for significance of parameters being dependent on: Male	F-test for significance of parameters being dependent on: Ages 6–11
Age when started school					
Male and age parameter differences	–0.048	–0.021	0.020	7.7	50.8
	(5.8)*	(2.0)*	(1.9)**	[0.0005]	[0.0000]
Male and age parameter differences plus school/commune fixed effects	–0.046	–0.008	0.023	5.5	98.3
	(5.8)*	(1.0)	(2.7)*	[0.0043]	[0.0000]
Grades passed per year of school					
Male and age parameter differences	0.198	–0.011	0.007	8.0	59.4
	(7.9)*	(0.3)	(0.2)	[0.0003]	[0.0000]
Male and age parameter differences plus school/commune fixed effects	0.131	–0.023	0.024	7.9	69.8
	(4.7)*	(0.8)	(0.8)	[0.0004]	[0.0000]
Last completed grade					
Male and age parameter differences	0.426	–0.100	–0.122	0.4	20.9
	(5.8)*	(0.9)	(0.7)	[0.6502]	[0.0000]
Male and age parameter differences plus school/commune fixed effects	0.400	–0.056	–0.120	0.2	20.5
	(5.3)*	(0.6)	(0.7)	[0.8266]	[0.0000]
Exam score in last completed grade					
Male and age parameter differences	0.094	–0.020	0.024	16.2	22.6
	(5.0)*	(1.0)	(1.2)	[0.0000]	[0.0000]
Male and age parameter differences plus school/commune fixed effects	0.083	–0.018	0.037	14.3	26.3
	(3.6)*	(1.0)	(1.8)**	[0.0000]	[0.0000]

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Absolute values for *t*-statistics are given in parentheses beneath point estimates. Probabilities are given in square brackets. The underlying estimates are given in table B-2. Censored normal regressions are used for age when started school and for last completed grade because of right-censoring. Upper-limit tobits are used for grades passed per year of school because of the mass point at 1.0.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

start school is more income-responsive than is the age that girls start school (table 4). But otherwise, if anything, boys' school progress indicators have smaller associations with parental household income than do girls'. Although the individual point estimates are imprecisely estimated for both the income-interactive impact of sex and for the additive component, *F*-tests indicate that the sex differences are significantly nonzero at the 1 percent level for age when started school, grades passed per year of school, and exam score in last completed grade, although not for last completed grade. Therefore, once again, if we were to focus only on last completed grade, our inference would be misleading: we would infer that there are no significant gender differences, when in fact there are, according to the other three indicators of school progress.

The point estimates suggest, moreover, that these effects may be fairly large—the income elasticity for boys is more than 40 percent larger in absolute magnitude for age when started school and is 6 and 22 percent smaller, respectively, for grades passed per year of school and exam score in last completed grade (including the effects generated by school and commune characteristics). The income elasticity of total schooling expenditures that households pay directly to schools is also significantly greater for boys than for girls, even though the elasticities for school fees alone and households' school-related expenses not paid directly to schools do not differ significantly by sex (see table B-3).

Thus there seem to be some gender differences in income associations with three of the four indicators, although age when started school pushes in a different direction than grades completed per year of school and exam score in last completed grade. The net effect of these gender differences is to imply a smaller association between income and school progress for boys than for girls because the income effects are much larger for grades completed per year of school than for age when started school. The association between enrollment and parental household income is also significantly smaller for boys than for girls (compare elasticities of 0.092 and 0.196 in table B-4). Thus if income elasticities are higher for luxuries than for necessities, girls' schooling is treated as more of a luxury than is boys' schooling.

The absolute magnitude of the elasticities of income associated with age when started school and last completed grade are smaller for children in the age range 6–11 compared with children in the age range 12–17, but the elasticities of income associated with grades passed per year of school and exam score in last completed grade are larger. The individual income-multiplicative and additive age effects (similar to those for sex) tend to be imprecisely estimated, but *F*-tests indicate that the age effects are significantly nonzero at high levels of significance for all four indicators (table 4). Also, the point estimates suggest that there may be fairly large sex effects on the gross income elasticities for the 6–11 age group compared with the 12–17 age group. These effects reduce the absolute magnitude of the income elasticity by more than 40 percent for age when started school and by more than 25 percent for last completed grade, and they increase the income elasticity for exam score in last completed grade by more than 25 percent

(although they increase the income elasticity for grades passed per year of school only 4 percent).

The smaller absolute magnitude of the income elasticity of age when started school for younger children suggests a reduction in the importance of income in determining age of entry between these two cohorts. The larger elasticities of the older group for last completed grade and exam score are consistent with income being a less important constraint for children of primary school age, given that primary school is almost universal, but being a more important constraint for older children (12–17 years), for whom most decisions at the margin about continuing school are made. The pattern is the opposite for grades completed per year of school, although, as usual with cross-sectional data, it is difficult to disentangle life-cycle and cohort effects for such variables. Also, it should be noted that the interaction between age group and income is not significantly nonzero in the enrollment probits (table B-4).

V. Associations between Parental Household Income and School Policies

School policies may compensate for differences in household income or may reinforce them. Two major aspects of such policies are income-related school fees and school quality in terms of the quality of teachers, current inputs, and school structures, and the amount of congestion. Both school fees and school quality may affect which children enroll in school and how well they succeed once enrolled. Whatever distributional effect these policies have on children who enroll in school, they have a less equalizing effect on all children if enrollment itself (inclusive of the effects of these policies) is inversely associated with income. We find that, indeed, enrollment is positively associated with household income (table 5). Children in the first two quintiles have significantly lower enrollment rates, and those in the fifth quintile have significantly higher enrollment rates, than those in the third quintile. We return to this point at the end of this section.

School Fees and Other Household School-Related Expenditures

Actual school fees paid tend to be inversely associated with income: the means for the first and second quintiles are significantly below, and the mean for the fifth quintile is significantly above, the mean for the third quintile (table 5). The structure of school fees is progressive mainly because there are a primary school fee exemption and a higher concentration of children from lower-income households in primary school. Of the total number of children in the sample who receive exemptions from school fees, 80.3 percent receive them because they are in primary school, 8.0 percent because they live in mountainous regions, 4.3 percent because they are members of ethnic minorities, and only 1.0 percent because they are from poor households.

For this reason the income elasticities that depend on age are of primary interest. These estimates yield an income elasticity of 2.312 for the age range 12–17

Table 5. *Means of Enrollment, School Fees, and Other Household School-Related Expenditures Conditional on Enrollment by Predicted Income per Household Member*

Enrollment, fees, and school-related expenditures	*Income quintile*				
	1 (poorest)	*2*	*3*	*4*	*5*
Current enrollment	0.72	0.78	0.78	0.81	0.87
	(0.45)	(0.41)	(0.41)	(0.39)	(0.34)
	[2.8]*	[2.3]*		[1.3]	[6.1]*
Conditional on enrollment at time of survey					
School fees	15	12	21	22	36
	(35)	(32)	(52)	(48)	(68)
	[2.5]*	[2.3]*		[0.5]	[3.6]*
Total expenditures paid to schools[a]	38	41	55	67	102
	(45)	(42)	(65)	(59)	(88)
	[4.7]*	[3.1]*		[3.1]*	[8.6]*
School-related expenditures not paid to schools[b]	123	146	175	287	523
	(140)	(198)	(181)	(362)	(411)
	[4.8]*	[1.7]**		[5.4]*	[15.3]*
Total school-related expenditures	162	187	230	354	625
	(163)	(220)	(216)	(385)	(437)
	[5.5]*	[2.3]*		[5.7]*	[16.4]*

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Standard deviations are in parentheses, and absolute values of *t*-statistics for differences from the third quintile are given in square brackets. All fees and expenditures are measured in thousands of dong per year in 1996 prices.

a. Includes school fees, school improvement fees, parent association fees, and insurance.

b. Includes textbooks, supplies and materials, uniforms, tutoring fees, travel, meals, and miscellaneous.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

years, but close to zero for the primary school age group (6–11 years; table 6). The latter elasticity is close to zero because exemptions for primary school are widespread regardless of household income. That the income elasticity for the probability of fee exemptions is much larger in absolute magnitude for older children suggests that the income-related exemptions have a much greater effect. The estimated elasticity of school fees with respect to income in the simplest specification (with only income and not age) is smaller and more imprecise, whether or not we control for school and commune fixed effects (table B-3). If we do control for age, the estimated income elasticity does not change much for the age group 12–17 years if school and commune fixed effects are added, although the estimated change for the age group 6–11 years is much smaller in absolute magnitude and much more imprecisely estimated (table 6). Gender effects are not significantly nonzero in any of these specifications.

Table 6. *Estimated Elasticities of Household School-Related Expenditures with Respect to Predicted Income per Household Member*

Private school-related expenditures	*Base*	*Change for*	
		Ages 6–11	*Male*
School fees			
With all parameters dependent on age and sex[a]	2.312	–2.228	0.190
	(2.4)*	(1.6)	(0.2)
With school/commune fixed effects and age and sex[a]	2.123	–0.448	–0.729
	(2.1)*	(0.4)	(0.7)
Total expenditures paid to schools			
With all parameters dependent on age and sex[b]	1.063	–0.338	0.319
	(6.2)*	(1.7)*	(1.6)
With school/commune fixed effects and age and sex[b]	0.243	0.015	0.221
	(1.7)*	(0.1)	(1.5)

* Significant at the 5 percent level.

Note: The full lower-limit tobit estimates are given in table B-3. Absolute values for *t*-statistics are given in parentheses beneath point estimates.

a. *F*-tests indicate that parameter dependence on age is significant even though individual additive and multiplicative parameter estimates are imprecise (F = 223.5, probability 0.0000 for all parameters; F = 267.0, probability 0.0000 for fixed effects). *F*-tests indicate that parameter dependence on sex is insignificant (F = 1.1, probability 0.3458 for all parameters; F = 0.9, probability 0.4163 for fixed effects).

b. *F*-tests indicate that parameter dependence on age and sex is significant even though individual additive and multiplicative parameter estimates are imprecise (F = 23.8, probability 0.0000 for age, and F = 4.5, probability 0.0111 for sex in all parameters; F = 55.9, probability 0.0000 for age, and F = 4.3, probability 0.0130 for sex in fixed effects).

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

Thus school fees tend to favor poorer households among those households that enroll children in school; however, they are not targeted only to such households because they work substantially through exemptions of primary school fees, which are given broadly regardless of household income. But this progressive fee structure may have a limited effect because school fees are a relatively small proportion of total household expenditures paid to schools, to say nothing of total school-related household expenditures. School fees average only 34 percent of total school-related household expenditures paid directly to schools (school improvement fees also average 34 percent, insurance averages 12 percent, and parent association fees average 10 percent).

Total expenditures paid to schools, like school fees alone, increase systematically with income (see table 5), but with an income elasticity that is much smaller than the 2.312 estimated for school fees—it is 1.063 with comparable controls for age and sex and 0.243 with additional controls for school and commune fixed effects (table 6). Therefore, the relatively limited response to income of household expenditures paid directly to schools (other than school fees) means that, despite fairly high income elasticities for school fees, total household expenditures paid directly to schools are much less income-responsive than are school fees alone. School fees, moreover, are an even smaller share of total school-related household expenditures—the mean for the full sample is only 6.9 percent because school-related household expenditures not paid directly to schools

are almost four times as large as those paid directly to schools (16.4 and 4.2 percent, respectively, of total household expenditures per household member). Thus even school fee exemptions that are much better targeted would have only a limited impact on the relationship between total household school-related expenditures and income and therefore, presumably, on poor households' decisions about schooling. To have more of an impact, policies would have to extend exemptions to household expenditures paid directly to schools beyond school fees or amend payment structures to make them much more strongly related to household income (possibly including negative fees or subsidies for children from poorer households).

School Quality

Governments also might alter schooling options for different households by affecting the quality of public schools (which dominate in Vietnam, accounting for 91.5 percent of students in the sample). Much of the recent economic literature on schooling focuses on the importance of school quality for school outcomes (see, for example, Card and Krueger 1996 and Moffitt 1996 for recent surveys on studies in the United States; Alderman and others 1996b, Behrman and Birdsall 1983, Behrman, Birdsall, and Kaplan 1996, and Hanushek 1995 for studies and surveys of developing countries).

We have constructed four indicators of school quality for school staff, current inputs, congestion, and facilities (see the note to table 7 for details). A priori and on the basis of other studies, congestion seems to be negatively associated with school success, and the other three quality indicators seem to be positively associated. We calculate simple log-linear estimates of the associations between our four indicators of school success and our four quality indicators. These estimates look like production functions, and similar relations have often been interpreted as production functions (see, for example, most of the studies surveyed in Hanushek 1995). But they can be interpreted as production functions only under the assumption that the right-side variables are predetermined in a statistical sense and therefore are independent of the disturbance term in the relation. This seems highly unlikely a priori and in light of the estimates in table 9, suggesting that the right-side variables are significantly correlated with household income (and therefore are most likely correlated with other determinants of school success, such as home learning environments, the effects of which are in the disturbance term of this relation). In our cross-sectional data set (as in most cross-sectional data sets), moreover, plausible instruments that would enable us to use some technique, such as instrumental variables, to control for determinants of school quality, are not available.

Our estimates suggest that the strongest associations are between the quality of school staff and children's success in school (table 8). All four elasticities are significantly nonzero with the a priori expected signs and with absolute magnitudes ranging from 0.176 for exam score in last completed grade to 1.094 for last completed grade. The quality of current inputs is significant and positively asso-

Table 7. *Mean School Quality Indicators by Predicted Income per Household Member*

	Income quintile				
Indicator	*1 (poorest)*	*2*	*3*	*4*	*5*
Staff	0.66	0.70	0.71	0.73	0.76
	(0.11)	(0.09)	(0.10)	(0.08)	(0.07)
	[8.0]*	[1.6]		[3.4]*	[8.3]*
Current inputs	0.30	0.33	0.39	0.45	0.56
	(0.12)	(0.15)	(0.16)	(0.15)	(0.12)
	[9.4]*	[6.4]*		[6.2]*	[18.7]*
Congestion	0.19	0.19	0.20	0.20	0.21
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
	[3.4]*	[1.6]		[0.1]	[4.3]*
Facilities	0.10	0.18	0.24	0.33	0.36
	(0.14)	(0.18)	(0.22)	(0.25)	(0.27)
	[11.7]*	[5.1]*		[5.4]*	[7.5]*

* Significant at the 5 percent level.

Note: Standard deviations are in parentheses, and absolute values of *t*-statistics for differences from the third quintile are in square brackets. The results are based on local schools for all children ages 6–17 in the sample. The "facilities" variable is the proportion of buildings occupied by the school that are "permanent construction" (as opposed to "semi-permanent" or "temporary"). The "staff," "current inputs," and "congestion" indicators are each based on the sum of a number of components. For each component the range of responses for all schools has been normalized to between 0 and 1 for each of three schooling levels (primary, lower secondary, upper secondary) and then averaged across the three levels. The components for "staff" are whether the head of school is qualified to teach at the level of the school, the years of experience as head of the school, the proportion of the school's teachers who are qualified by the Ministry of Education and Training to teach at that level, the proportion of teachers ranked "excellent" or "good" by the head of the school, the proportion of teachers who are full-time, and the average teaching experience of teachers. The components for "current inputs" are the proportion of classrooms wired for electricity, the proportion of classrooms with legible blackboards, the proportion of classrooms with ceiling fans, whether the school has safe water, whether the school has hygienic latrines, whether the school has a library, whether the school has a science laboratory, the number of computers per student, and the number of textbooks per student for rental or loan to poor students. The components for "congestion" are the proportion of classes taught in the third shift, the number of students per class, and the number of students per teacher.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

ciated with grades passed per year of school and exam score in last completed grade and is associated less precisely with last completed grade (significant at the 15 percent level), but with much smaller elasticities (between 0.067 and 0.096). Moreover, the association with age when started school is significant and opposite in sign to that expected a priori. Congestion is significant and negatively associated with exam score in last completed grade, with an elasticity of –0.050, but is not significantly nonzero for the other indicators. Facilities are significantly nonzero only for last completed grade, with the opposite sign to that expected a priori.

What are the associations between these indicators of school quality and parental income? When we compared the estimates of the elasticities of the four school success indicators with respect to household income with and without controls for school and commune fixed effects, we found that, on net, school quality indicators are positively associated with income (that is, the basic estimates in table 3 are larger without than with controls for school and commune

Table 8. *Regressions of the Logarithm of Individual School Progress Indicators on School Quality Indicators*

	Elasticity with respect to							
Indicator	*Staff*	*Current inputs*	*Congestion*	*Facilities*	*Constant*	*R-squared*[a]	*Chi² test*[b]	*Number of observations*
Age when started school[c]	–0.206	0.026	–0.011	0.0004	1.770	0.048	157.9	2,308
	(11.7)*	(4.3)*	(1.1)	(0.1)	(109.9)*	[0.115]	[0.0000]	
Grades passed per year of school	0.420	0.090	0.022	0.012	0.288	0.048	128.3	2,403
	(7.4)*	(4.5)*	(0.7)	(1.1)	(5.2)*	[0.355]	[0.0000]	
Last completed grade[c]	1.094	0.096	0.109	–0.114	2.958	0.029	51.4	2,306
	(6.0)*	(1.6)	(1.1)	(3.4)*	(17.3)*	[0.732]	[0.0000]	
Exam score in last completed grade[d]	0.176	0.067	–0.050	–0.009	1.836	0.025	6.4	2,094
	(3.3)*	(3.6)*	(2.0)*	(1.1)	(41.4)*	[0.251]	[0.0000]	

* Significant at the 5 percent level.

Note: Absolute values for *t*-statistics are given in parentheses beneath point estimates. Because the variable log facilities is in the data for only 1,957 observations, a dummy variable is included to control for observations for which this variable is missing.

a. Standard error given in square brackets.

b. Probability given in square brackets.

c. Censored normal regression (with 1 right-censored and 2,307 uncensored observations for age when started school and with 2,173 right-censored and 442 uncensored observations for last completed grade). The R^2 is a pseudo R^2.

d. Regressions are with robust standard errors, and the standard deviations are corrected for clustering at the household level. The root mean standard error is given beneath R^2, and the *F*-test is given in the penultimate column.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

fixed effects). All of the estimated elasticities of the four quality indicators with respect to household income are positive, although that for congestion is significantly nonzero only at the 10 percent level (table 9). This result means that three of the four quality measures do not compensate poorer households; rather, they reinforce the advantages of higher-income households. The one exception is congestion, which weakly favors poorer households but is significantly associated with only one of our indicators. Because policies result in children from higher-income households on net having higher-quality facilities, part of the positive association between parental household income and household expenditures paid to schools may result from households' paying for higher-quality schooling and not from a progressive school fee structure for a given school quality.

Combined Implications

We now illustrate how the associations between household income and school-related expenditures paid directly to schools, school quality, and school enrollment can be combined to obtain an estimate of the association between the price of school and household income holding quality constant. Let household school-related expenditures paid directly to schools be the product of three components: the price for a given school quality, school quality, and quantity. That is, we posit the effective quantity of school to be the quantity adjusted for the quality by a multiplicative factor. Thus the expenditure elasticity with respect to income ($E_{exp,\ inc}$) is the sum of the elasticity of the constant-quality price with respect to income ($E_{price,\ inc}$), the elasticity of the quantity with respect to income ($E_{quan,\ inc}$), and the elasticity of the quality with respect to income ($E_{qual,\ inc}$):

$$E_{exp,\ inc} = E_{price,\ inc} + E_{quan,\ inc} + E_{qual,\ inc} \tag{1}$$

Table 9. *Regressions of the Logarithm of School Quality Indicators on the Logarithm of Predicted Income per Household Member*

Indicator	*Elasticity with respect to income*	*Constant*	*R-squared*[a]	*F-test*[b]	*Number of observations*
Staff	0.086	–0.977	0.091	12.8	2,388
	(3.6)*	(5.3)*	[0.140]	[0.0005]	
Current inputs	0.435	–4.180	0.201	36.1	2,388
	(6.0)*	(7.6)*	[0.445]	[0.0000]	
Congestion	0.072	–2.184	0.021	2.9	2,388
	(1.7)**	(7.2)*	[0.253]	[0.0937]	
Facilities	0.686	–6.652	0.125	15.2	1,942
	(3.9)*	(5.2)*	[0.935]	[0.0002]	

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Regressions are with robust standard errors. Absolute values for *t*-statistics are given in parentheses beneath point estimates. The standard deviations are corrected for clusters at the commune level.

a. Root mean squared error given in square brackets.

b. Probability given in square brackets.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

If the first component—the elasticity of price with respect to income—is positive, the marginal cost curve is lower for poorer households, so pricing favors the poor. A positive value of the second component, the elasticity of quantity with respect to income, implies that higher-income households purchase more inputs than poorer households, which, ceteris paribus, translates into greater benefits for higher-income households. A positive value of the third component, the elasticity of quality with respect to income, implies that higher-income households purchase higher-quality inputs than poorer households, which, ceteris paribus, translates into greater benefits for higher-income households.

Let $E_{exp,\ inc}$ be the elasticity of household school-related expenditures paid directly to schools (1.063 in table 6), $E_{quan,\ inc}$ be the elasticity of enrollment with respect to household income (0.152 in table B-4), and $E_{qual,\ inc}$ be the average of the elasticities of the two components of school quality that are most related to child school success in table 8—the quality of staff and the quality of current inputs (0.261, the average of the first two rows of table 9). Equation 1 can then be solved for the constant-quality price elasticity with respect to income, $E_{price,\ inc}$ = 0.650. This value is considerably less than—in fact only about three-fifths as large as—the income elasticity of household school-related expenditures paid directly to schools. The structure of household fees paid directly to schools thus appears to be much less progressive once we correct for the fact that children from poor households are less likely to enroll in school and children from higher-income households have higher-quality schools available to them.

VI. Conclusions

Many societies are concerned about the association between parental income and child schooling because it is perceived that the stronger this association is, the less is intergenerational social mobility and the less equal is opportunity. Such concerns have been especially strong in Vietnam, although there is a perception that some important policies, such as school fee exemptions, have substantially weakened the association between household income and school success.

We investigated the magnitudes of these associations and found them to be considerable. For example, our estimate for the income elasticity of completed grades is five times the median estimate of 42 earlier studies. We found, moreover, that this association is strongest *not* for completed grades, on which most of the previous literature has focused, but for grades completed per year of school. That is, the most important relationship takes into account failed and repeated grades, not just the total number of grades finished. We also found significant, but smaller, associations between parental household income and age when started school and exam score in last completed grade. There is some evidence of gender differences in the income associations, the strongest being the smaller income association with grades completed per year of school for boys than for girls. This result implies that schooling of girls is treated as more of a luxury (less of a necessity) than is schooling of boys.

We also explored the sensitivity of our estimates to a number of different estimation choices made in previous studies. We found that the estimated associations with income are significantly larger when we use predicted income (to represent longer-run income), control for censoring, and represent additional channels through which there may be effects beyond just the number of completed grades. If our results generalize to other societies, much of the previous literature may understate the true associations between household income and school success and therefore overstate true intergenerational social mobility and equality of opportunity.

Finally, we explored some dimensions of how policies relate to parental household income. School fees are progressive in the sense that they favor children from lower-income households among those children enrolled in school, particularly because of the primary school fee exemption. But school fees are only about one-third of what households pay directly to schools and are a much smaller proportion of households' total school-related expenditures. Total expenditures paid directly to schools increase with household income much less proportionately than do school fees alone, so the overall structure of such payments is much less progressive than is the structure of school fees. Moreover, because school enrollment is positively correlated with household income, the structure of school fees is less progressive for the entire population of households with school-age children than it is for the selected subset of that population with children enrolled in school. Further, the two school quality measures that are most strongly and positively associated with our four indicators—the quality of the staff and the quality of current inputs—are much more available to higher-income households, meaning that higher-income households have greater school expenditures in part because they are obtaining higher-quality schooling, not because of progressive charges for the same quality of schooling across income classes. Therefore, although the school fee structure attempts to equalize schooling options, it has only a limited impact.

Appendix A. A Summary of Studies Exploring the Association between Income and Schooling

Table A-1. *Estimated Income Elasticities for Schooling*

Country and year	*Schooling indicator*	*Income elasticity*	*Source*	*Notes*
Bangladesh, 1980–81	School attendance	0.20	Hossain (1989)	Father's income; control for household and community characteristics
Bolivia, 1989	Grade repetition		Patrinos and Psacharopoulos (1993)	Income; control for household and community characteristics
Bolivia, 1990	Grade attained Days missed Grades failed	0.04* –0.06** –0.02	Behrman, Li, and Murillo (1995)	Expenditure per household member; control for household characteristics and community fixed effects
Brazil, 1970	Completed years	0.09 to 0.16*	Birdsall (1985)	Father's income; control for household and teacher characteristics
Brazil, 1972–74	Enrollment	–0.17*	Singh (1992)	Rural only; income (excluding children's income); control for household characteristics
Brazil, 1982	Completed years	0.06 to 0.14*	Barros and Lam (1996)	For São Paulo and the Northeast; household head's income; control for parental schooling
Brazil, 1982	Completed years	0.12 to 0.22*	Thomas, Schoeni, and Strauss (1996)	Urban only; household income; control for household characteristics
Côte d'Ivoire, 1985–87	Completed years Current enrollment*	0.19*	Montgomery and Kouame (1993)	Consumption per adult; control for mother's characteristics and region
Côte d'Ivoire, 1985–87	School attainment	0.14 to 0.42*	Tansel (1997)	Expenditure per capita (instrumented); control for age, sex (significant for females, lower and insignificant or significant at 10 percent level for males), school distance, wages, rural area
Egypt, 1980	Ever attended; currently attending; and years attended		Cochrane, Mehra, and Osheba (1986)	Income per capita; control for household and school characteristics
Ghana, 1988–89	Grade attainment;* reading;* attending other than nearest school;* delayed enrollment;* mathematics; dropping out age		Glewwe and Jacoby (1994, 1995a)	Expenditure per capita (instrumented); control for household and (numerous) school characteristics and selectivity

(Table continues on the following page.)

Table A-1. *(continued)*

Country and year	*Schooling indicator*	*Income elasticity*	*Source*	*Notes*
Ghana, 1987	Ever-attended school; school attainment		Lavy (1996)	Income per capita; control for household and (numerous) school and community characteristics
Ghana, 1987–89	School attainment	0.18 to 0.56*	Tansel (1997)	Expenditure per capita (instrumented); control for age, sex (significant for males, lower and insignificant or significant at 10 percent level for females), school distance, wages, rural area
Guatemala, 1989	Repetition**		Patrinos and Psacharopoulos (1993)	Income; control for household and community characteristics
India, 1980–81	Enrollment		Duraisamy (1988)	Rural Tamil Nadu; non-labor market income; control for household and community characteristics
India, 1991	Achievement test	0.12*	Kingdon (1996)	Lucknow, Uttar Pradesh; index of consumer durables; control for household and school characteristics
Indonesia, 1987	Probability of attainment		Deolalikar (1993)	Nonlabor income; control for household and community characteristics
Indonesia, 1989	Transition probabilities for initial enrollment and to next school level*		King (1995)	Expenditure per capita; control for household and community characteristics
Jamaica, 1989	Enrollment*		Handa (1994)	Expenditure per capita; control for household and community characteristics
Jamaica, 1990	Mathematics	0.07*	Glewwe and others (1995)	Expenditure per capita; control for household and numerous school characteristics and selectivity
Kenya, 1994	Reading Enrollment;* household school expenditure;* student-teacher ratio*	0.04	Deolalikar (1997)	Expenditure per capita; interacts with household and school characteristics

Malaysia, 1975–76	Current enrollment* (Malays**)		De Tray (1984)	Income; control for household and school characteristics
Malaysia, 1975–76	Completed schooling (Malays,* Chinese)		King and Lillard (1987)	Income; control for household and school characteristics and right-censoring
Malaysia, 1988	Schooling progression probabilities		Lillard and Willis (1994)	Father's earnings; control for household characteristics, distance to primary schools, and region
Mali, 1981–82	Enrollment		Birdsall and Orivel (1996)	Income; control for school characteristics
Nepal, 1980–81	Grade attainment Enrollment*	0.38*	Moock and Leslie (1986)	Value of crop output; control for household characteristics, presence of local primary school, and region
Nicaragua, 1977–78	Grades completed	0.02 to 0.07	Behrman and Wolfe (1987)	Income or predicted mother's earnings plus other household income; control for household and community characteristics
Pakistan, 1991	Ever-enrolled;* expenditure on primary school*		Sather and Lloyd (1994)	Expenditure; control for household characteristics, distance to primary schools and region
Pakistan, 1989	Numeracy Literacy Attendance (males,* females)	0.05 0.23*	Alderman and others (1996a)	Household income (instrumented); control for household and community variables
Pakistan, 1986–92	Enrollment*		Alderman, Behrman, Khan, Ross, and Sabot (1997)	Rural only; expenditure (average over three years); control for household, school, and community characteristics
Paraguay, 1990	Current enrollment;* years; grade repetition; dropping out		Patrinos and Psacharopoulos (1995)	Income; control for household characteristics
Peru, 1985–86	Ever-enrolled;* early enrollment		Ilon and Moock (1991)	Rural only; household expenditure; control for household and community variables

(Table continues on the following page.)

Table A-1. *(continued)*

Country and year	*Schooling indicator*	*Income elasticity*	*Source*	*Notes*
Peru, 1985–86	Progression on time through school*		Jacoby (1994)	Income other than self-employed income; control for household and school variables
South Africa, 1993	Years	-0.01 to 0.10*	Case and Deaton (1996)	Expenditure per household member; ages 10–18 for years of school, 8–24 for enrollment; control for household characteristics including race and pupil–teacher ratio; years of schooling lower and insignificant for whites; reason not attending: expense, illness, completed, insignificant for pregnancy and cannot cope; expenditure share insignificant for secondary and, for whites, for primary
	Enrollment (blacks)	0.01*		
	Reason not attending (blacks)	–0.02 to 0.02*		
	Expenditure share—school	–0.02 to 0.12*		
	Test scores* (blacks; for whites literacy not significant and numeracy significant at 10 percent level)			
Taiwan (China), 1989	Years of schooling	0.12 to 0.33*	Parish and Willis (1993)	Father's income; birth cohorts 1940–49 through 1970–75, with declining elasticity for more recent cohorts; control for household characteristics
United States, 1957	Years of schooling	0.04*	Hauser and Daymont (1977)	Wisconsin high school graduates; parental income; control for household characteristics

United States, 1978	Years of schooling	–0.03 to 0.00	Datcher (1982)	Income; control for household and community characteristics
United States, 1981	Years of schooling	0.02*	Behrman and Taubman (1986)	Income; control for household characteristics
United States, 1982	Years of schooling	0.09*	Hill and Duncan (1987)	Income averaged over three; control for household and community characteristics
United States, 1981	Years of schooling	0.04*	Behrman and Taubman (1989)	Income; control for household characteristics
United States, 1979–86	Negative or insignificant for college grades and graduate probabilities		Datcher Loury and Garman (1995)	Income; control for individual test scores, race, and college mean test scores
Venezuela, 1987	Years Negative for currently in school,* repeating grade,* and illiterate*	0.01**	Psacharopoulos and Yang (1991)	Family income; control for age, sex, and father's schooling

* The underlying point estimates are significantly nonzero at the 5 percent level.

** The underlying point estimates are significantly nonzero at the 10 percent level.

Note: Income elasticities calculated at point of sample means for what appear to be preferred estimates. If information is not provided with which to calculate elasticities, the dependent variables and significant levels are indicated.

Source: Behrman and Knowles (1997).

Appendix B. Estimation Results for Predicted Income and Associations between Income and Schooling

Table B-1. *Estimates of Current Income per Household Member and Total Household Expenditures per Household Member as a Function of Longer-Run Household Characteristics*

Right-side variables	*Current income*	*Current expenditures*
Parent's schooling		
Mother's schooling	2.49	5.38
	(2.3)	(2.2)
Father's schooling	2.20	1.90
	(1.9)	(2.1)
Mother's schooling missing	8.16	9.23
	(0.7)	(1.0)
Father's schooling missing	10.60	15.50
	(0.9)	(1.8)
Number of household members		
Male	−8.43	−6.31
	(3.5)	(3.4)
Female	−15.69	−11.23
	(7.0)	(6.5)
Ethnic group	4.72	4.88
	(0.5)	(0.7)
Formal religion	66.89	35.31
	(4.9)	(3.4)
Catholic	−56.95	−28.81
	(3.8)	(2.5)
Buddhist	−27.90	−12.41
	(2.1)	(1.2)
Assets per dong10,000	0.359	0.186
	(6.1)	(4.2)
Income primarily from primary sector	16.81	7.59
	(2.0)	(1.2)
Employment		
State	−30.67	10.17
	(1.1)	(0.5)
Cooperative	−104.84	−58.45
	(4.0)	(2.9)
Private sector	−31.20	−4.70
	(1.2)	(0.2)
Infrastructure availability or use		
Electricity	45.27	28.49
	(6.3)	(5.2)
Good water	5.51	6.89
	(0.8)	(1.2)
Latrine	62.49	41.90
	(8.1)	(7.1)
Good transportation	−2.98	−3.49
	(0.5)	(0.7)
Regular news	37.88	31.28
	(5.5)	(6.0)
Constant	166.90	119.06
	(5.7)	(5.3)
Root mean squared error	124.00	94.78
Adjusted R^2	0.27	0.26

Note: Sample includes all 1,844 households with relevant data. Absolute *t*-values are in parentheses.
Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

Table B-2. *Regressions of the Logarithm of School Progress Indicators on the Logarithm of Predicted Income per Household Member*

Indicator	Elasticity with respect to income: Base: female ages 12–17	Elasticity: Difference from base if Male	Elasticity: Difference from base if Ages 16–11	Constant: Base: female ages 12–17	Constant: Difference from base if Male	Constant: Difference from base if Ages 6–11	R^2[b]	Statistical tests[a]: Chi² test for overall significance	F-test: School/commune fixed effects	F-test: Male	F-test: Ages 6–11
Age when started school (N = 2,625)[c]											
Basic	–0.051			2.22			0.030	90.7			
	(9.6)*			(57.4)*			[0.133]	[0.0000]			
Annual income, no control for censoring[d]	–0.051			2.22			0.056	37.4			
	(6.1)*			(34.9)*			[0.131]	[0.0000]			
Predicted income, no control for censoring[d]	–0.050			2.22			0.034	30.5			
	(5.5)*			(32.1)*			[0.132]	[0.0000]			
Basic plus school/commune fixed effects	–0.040			2.16			0.433	1312.8	25.1		
	(7.0)*			(48.2)*			[0.105]	[0.0000]	[0.0000]		
Basic plus male and age parameter differences	–0.048	–0.021	0.020	2.22	0.17	–0.20	0.068	205.8		7.7	50.8
	(5.8)*	(2.0)*	(1.9)**	(36.3)*	(2.3)*	(2.6)*	[0.130]	[0.0000]		[0.0005]	[0.0000]
Basic plus school/commune fixed effects and male and age parameter differences	–0.046	–0.008	0.023	2.21	0.07	–0.22	0.499	1512.8	27.3	5.5	98.3
	(5.8)*	(1.0)	(2.7)*	(38.8)*	(1.2)	(3.6)*	[0.101]	[0.0000]	[0.0000]	[0.0043]	[0.0000]
Grades passed per year of school (N = 2,774)[e]											
Basic	0.193			–1.398			0.046	148.0			
	(12.1)*			(12.1)*			[0.372]	[0.0000]			
Annual income, no control for censoring[d]	0.070			–0.666			0.036	30.2			
	(5.7)*			(7.3)*			[0.227]	[0.0000]			

(Table continues on the following page.)

Table B-2. *(continued)*

Indicator	Elasticity with respect to income: Base: female ages 12–17	Elasticity: Difference from base if Male	Elasticity: Difference from base if Ages 16–11	Constant: Base: female ages 12–17	Constant: Difference from base if Male	Constant: Difference from base if Ages 6–11	R^2[b]	Statistical tests[a]: Chi² test for overall significance	F-test: School/commune fixed effects	F-test: Male	F-test: Ages 6–11
Predicted income, no control for censoring[d]	0.118			–1.015			0.063	81.4			
	(9.0)*			(10.3)*			[0.224]	[0.0000]			
Basic plus school/commune fixed effects	0.134			–1.100			0.175	556.3	6.2		
	(6.3)*			(7.1)*			[0.342]	[0.0000]	[0.0000]		
Basic plus male and age parameter differences	0.198	–0.011	0.007	–1.477	0.016	0.129	0.091	290.1		8.0	59.4
	(7.9)*	(0.3)	(0.2)	(8.1)*	(0.1)	(0.5)	[0.367]	[0.0000]		[0.0003]	[0.0000]
Basic plus school/commune fixed effects and male and age parameter differences	0.131	–0.023	0.024	–1.150	0.110	0.011	0.226	720.4	6.4	7.9	69.8
	(4.7)*	(0.8)	(0.8)	(5.7)*	(0.5)	(0.1)	[0.336]	[0.0000]	[0.0000]	[0.0004]	[0.0000]
Last completed grade (N = 2,615)[c]											
Basic	0.356			–0.068			0.025	54.4			
	(7.4)*			(0.2)			[0.819]	[0.0000]			
Annual income, no control for censoring[d]	0.240			–0.367			0.046	80.7			
	(9.0)*			(1.8)**			[0.680]	[0.0000]			
Predicted income, no control for censoring[d]	0.178			0.090			0.016	23.7			
	(6.3)*			(0.3)			[0.691]	[0.0000]			
Basic plus school/commune fixed effects	0.353			0.192			0.191	425.7	5.9		
	(6.1)*			(0.4)			[0.651]	[0.0000]	[0.0000]		
Basic plus male and age parameter differences	0.426	–0.100	–0.122	0.620	0.729	1.52	0.052	116.5		0.4	20.9
	(5.8)*	(0.9)	(0.7)	(1.2)	(0.9)	(1.1)	[0.881]	[0.0000]		[0.6502]	[0.0000]

Basic plus school/commune fixed effects and male and age parameter differences	0.400	–0.056	–0.120	–0.135	0.407	1.43	0.219	486.2	5.5	0.2	20.5
	(5.3)*	(0.6)	(0.7)	(0.2)	(0.6)	(1.1)	[0.693]	[0.0000]	[0.0000]	[0.8266]	[0.0000]
Exam score in last completed graded (N = 2,374)											
Basic	0.092			1.292			0.029	37.1			
	(6.1)*			(10.0)*			[0.263]	[0.0000]			
Annual income, no control for censoring[d]	0.085			1.177			0.038	27.1			
	(6.2)*			(16.6)*			[0.263]	[0.0000]			
Predicted income, no control for censoring[d]	0.092			1.292			0.029	37.1			
	(6.1)*			(10.0)*			[0.263]	[0.0000]			
Basic plus school/commune fixed effects	0.087			1.225			0.116	7.0	5.1		
	(3.7)*			(7.5)*			[0.255]	[0.0000]	[0.0000]		
Basic plus male and age parameter differences	0.094	–0.020	0.024	1.111	0.085	–0.107	0.060	21.6		16.2	22.6
	(5.4)*	(1.0)	(1.2)	(8.8)*	(0.6)	(0.7)	[0.259]	[0.0000]		[0.0000]	[0.0000]
Basic plus school/commune fixed effects and male and age parameter differences	0.083	–0.018	0.037	1.246	0.077	–0.200	0.143	7.4	5.5	14.3	26.3
	(3.6)*	(1.0)	(1.8)**	(7.7)*	(0.6)	(1.3)	[0.251]	[0.0000]	[0.0000]	[0.0000]	[0.0000]

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Absolute values for *t*-statistics are given in parentheses beneath point estimates.

a. Probabilities in square brackets.

b. Standard errors in square brackets.

c. Censored normal regression (with 8 right-censored and 2,617 uncensored observations for age when started school and with 2,173 right-censored and 442 uncensored observations for last completed grade). The R^2 is a pseudo R^2.

d. Regressions are with robust standard errors and with standard deviations corrected for household clusters. The overall test is an *F*-test, not a Chi^2 test. Root mean standard error is given beneath R^2.

e. Upper-limit tobit with upper limit at 1.0 and with 1,332 uncensored observations and 1,442 censored observations. The R^2 is a pseudo R^2.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

Table B-3. *Tobit Estimates of the Logarithm of Household School-Related Expenditures on the Logarithm of Predicted Income per Household Member*

	Elasticity with respect to income					
		Change for			*Pseudo*	
Private school-related expenditures	*Base*	*Ages 6–11*	*Male*	*Constant*	R^2[a]	*Chi² test*[b]
School fees (N = 1,463)	1.528			−14.96	0.0001	3.7
	(1.9)**			(2.5)*	[12.4]	[0.0543]
School/commune fixed effects[c]	1.185			−12.35	0.068	346.2
	(1.1)			(1.5)	[10.6]	[0.0000]
Age and sex[d]	2.312	−2.228	0.190	−11.67	0.123	627.7
	(2.4)*	(1.6)	(0.2)	(1.6)*	[9.0]	[0.0000]
School/commune fixed effects and age and sex[c, d]	2.123	−0.448	−0.729	−12.28	0.202	1032.3
	(2.1)*	(0.4)	(0.7)	(1.7)**	[7.2]	[0.0000]
Total expenditures paid to schools (N = 1,553)	1.055			2.593	0.016	103.0
	(10.3)*			(3.4)*	[1.95]	[0.0000]
School/commune fixed effects[e]	0.300			4.61	0.182	1202.7
	(2.7)*			(5.7)*	[1.38]	[0.0000]
Age and sex[f]	1.063	−0.338	0.319	2.99	0.024	157.4
	(6.2)*	(1.7)*	(1.6)	(2.4)*	[1.92]	[0.0000]
School/commune fixed effects and age and sex[e, f]	0.243	0.015	0.221	5.45	0.200	1317.1
	(1.7)*	(0.1)	(1.5)	(5.2)*	[1.33]	[0.0000]
School-related expenditures not paid to schools (N = 1,553)	1.285			2.308	0.028	179.1
	(13.8)*			(3.4)*	[1.78]	[0.0000]
School/commune fixed effects[g]	0.527			6.84	0.209	1338.0
	(5.3)*			(9.5)*	[1.24]	[0.0000]
Age and sex[h]	1.223	0.133	0.039	3.12	0.034	216.0
	(7.8)*	(0.7)	(0.2)	(2.7)*	[1.76]	[0.0000]
School/commune fixed effects and age and sex[g, h]	0.336	0.508	0.038	8.58	0.227	1455.3
	(2.6)*	(3.9)*	(0.3)	(9.2)*	[1.19]	[0.0000]

* Significant at the 5 percent level.

** Significant at the 10 percent level.

Note: Conditional on current enrollment at time of survey. Absolute values for *t*-statistics are in parentheses beneath point estimates. Lower-limit tobit estimates are used because of left-censoring at zero (for school fees, 954 left-censored observations and 509 uncensored observations; for total school expenditures, 42 left-censored observations and 1,511 uncensored observations; for school-related expenditures not paid to schools, 26 left-censored observations and 1,527 uncensored observations).

a. Standard errors in square brackets.

b. Probabilities in square brackets.

c. *F*-tests indicate that school/commune fixed effects are significant ($F = 3.1$ for school/commune fixed effects and $F = 3.6$ for school/commune fixed effects and age and sex, probability in both cases 0.0000).

d. *F*-tests indicate that parameter dependence on age is significant even though individual additive and multiplicative parameter estimates are imprecise ($F = 223.5$, probability 0.0000 for age and sex; $F = 267.0$, probability 0.0000 for school/commune fixed effects and age and sex). *F*-tests indicate that parameter dependence on sex is insignificant ($F = 1.1$, probability 0.3458 for age and sex; $F = 0.9$, probability 0.4163 for school/commune fixed effects and age and sex).

e. *F*-tests indicate that school/commune fixed effects are significant ($F = 10.9$ for school/commune fixed effects and $F = 1.3$ for school/commune fixed effects and age and sex, probability in both cases 0.0000).

f. *F*-tests indicate that parameter dependence on age and sex is significant even though individual additive and multiplicative parameter estimates are imprecise ($F = 23.8$, probability 0.0000 for age and $F = 4.5$, probability 0.0111 for sex in age and sex; $F = 55.9$, probability 0.0000 for age and $F = 4.3$, probability 0.0130 for sex in school/commune fixed effects and age and sex).

g. *F*-tests indicate that school/commune fixed effects are significant ($F = 19.8$ for school/commune fixed effects and $F = 21.6$ for school/commune fixed effects and age and sex, probability in both cases 0.0000).

h. *F*-tests indicate that parameter dependence on age (though not on sex) is significant even though individual additive and multiplicative parameter estimates are imprecise ($F = 17.1$, probability 0.0000 for age and $F = 2.0$, probability 0.1368 for sex in age and sex; $F = 59.7$, probability 0.0000 for age and $F = 2.1$, probability 0.1255 for sex in school/commune fixed effects and age and sex).

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

Table B-4. *Probit for School Enrollment on the Logarithm of Predicted Income per Household Member*

Alternative specifications	Coefficient of income: Base: female ages 12–17	Coefficient of income: Difference from base if Male	Coefficient of income: Difference from base if Ages 16–11	Constant: Base: female ages 12–17	Constant: Difference from base if Male	Constant: Difference from base if Ages 6–11	Pseudo R^2	Statistical tests[a]: Chi^2 test for overall significance	Statistical tests[a]: F-test: School/commune fixed effects	Statistical tests[a]: F-test: Male	Statistical tests[a]: F-test: Ages 6–11
Basic	0.000256 (6.3)* [0.152]			0.402 (5.6)*			0.017	39.3 [0.0000]			
Plus school/commune fixed effects	0.000332 (5.3)* [0.184]			1.167 (4.73)*			0.115	372.4 [0.0000]	314.2 [0.0000]		
Plus male and age parameter differences	0.000377 (5.6)* [0.194]	–0.000239 (2.9)* [0.071]	0.000141 (1.4) [0.267]	–0.277 (2.4)*	0.502 (3.7)*	0.959 (5.5)*	0.155	265.3 [0.0000]		14.9 [0.0006]	220.6 [0.0000]
Plus school/commune fixed effects and male and age parameter differences	0.000441 (4.5)* [0.196]	–0.000223 (2.5)* [0.092]	0.000112 (0.9) [0.246]	0.511 (1.9)	0.535 (3.4)*	1.214 (6.1)*	0.268	715.9 [0.0000]	358.0 [0.0000]	15.1 [0.0005]	288.8 [0.0000]

* Significant at the 5 percent level.

Note: N = 2,789. Standard errors are corrected for clustering at household level. Absolute values for *z* statistics are in parentheses beneath point estimates. Elasticity at means (enrollment = 0.791) are in brackets beneath standard errors (for males and ages 6–11 elasticities are for these categories, not for difference from base).

a. Probabilities are in brackets.

Source: Calculated from data from the 1996 Vietnam Social Sector Financing Survey.

Appendix C. Estimation of the Impact of Changes in Household Income on Future Income

We illustrate here the impact of changes in household income on future income for a child from a household that is one standard deviation above the mean and a child from a household that is one standard deviation below the mean of the distribution of household income per member (the example given in the fourth section of the paper). For these calculations we assume that the mean completed grades of schooling, S, is 7.5 grades, based on the completed schooling and enrollment rates of 17-year-olds in the sample; the real interest rate, r, is 5 percent; the retirement age, R, is 60 years; and the private cost of attending school for children is the children's time and the school-related expenditures of the household. Further, we assume that the students generate no income while in school and that, after school, annual income (or other returns from schooling), Y_S, is dependent on completed grades of schooling and is constant until retirement.

Delay in Starting School

The basic cost of starting school when a child is older is the delay in obtaining post-school returns. Consider the cost of a delay in finishing S grades of school in terms of the present discounted value of future income at an interest rate r evaluated at age six due to starting school when older. This cost is the difference in the present discounted value of future income with the delay and without the delay:

$$\text{(C-1)} \qquad \int_{S+D}^{R-6} e^{-rt} Y_S dt - \int_{S}^{R-6} e^{-rt} Y_S dt = \frac{Y_S}{r}\left[e^{-r(S+D)} - e^{-rS}\right]$$

where D is the delay in completing school beyond age $S + 6$. For the illustrative example considered in the fourth section of the article, the delay in starting school is 0.25 years, so that this expression equals $0.171 Y_S$.

Reduced Rate of Completing Grades

Reducing the rate of completing grades has two major effects on income. The first is a delay in obtaining the post-schooling returns from school, which has the same effect as a delay in starting school, so expression C-1 can be used to calculate this effect. For the illustrative example considered in the fourth section, D is 1.40 years for passing the average of 7.5 grades at a rate of 0.80 instead of 0.94, so this expression equals $0.929 Y_S$. The second effect is the monetary cost incurred by the household in terms of school-related expenditures. At the sample means, households' total school-related expenditures are 0.076 of annual parental household income per household adult, Y_H, which must be multiplied by 1.40 and discounted back to age six, yielding $0.071 Y_H$. Under the added assumption that average income does not change over time, so that $Y_S = Y_H$, the total present discounted value of the cost of this reduced rate of passing courses is $0.992 Y_S$.

Reduced Number of Completed Grades

A reduction in completed grades from S to S', also has two effects. First, the present discounted value of post-schooling income falls:

$$\text{(C-2)} \quad \int_{S+D}^{R-6} e^{-rt} Y_S dt - \int_{S'}^{R-6} e^{-rt} Y_{S'} dt = \frac{Y_S}{r}\left[e^{-rS} - e^{-r(R-6)}\right] - \frac{Y_{S'}}{r}\left[e^{-rS'} - e^{-r(R-6)}\right].$$

For the illustrative example considered in the fourth section, $S - S' = 2.2$. An estimated semilog income relation for the parents of the children in the sample yields $Y_S = 1{,}584$ and $Y_S' = 1{,}353$ for $S - S' = 2.2$, centered around the mean parental schooling level of 6.7 grades. [ln income = 0.072 schooling (t = 14.7) + 6.807 (t = 156.4), $R^2 = 0.227$, root mean squared error = 0.452, F = 215.3.] If these values are used to evaluate expression C-2 (again under the assumption that $Y_S = Y_H$), they imply a loss of income equal to $0.536Y_S$ at the sample means. Second, the household school-related expenditures also fall, under assumptions parallel to those above, by $0.120Y_S$ at the sample means. The net effect is a loss of $0.416Y_S$ at the sample means.

Reduced Exam Score in Last Completed Grade

For the illustrative example considered in the fourth section, exam performance falls about 7 percent, or about one-quarter of the sample variance. Because the data set does not have exam scores and income for the same individuals, we are not able to estimate the relationship between income and exam performance from the sample. We assume here that the effect on the present discounted value of post-schooling income is about one-quarter of the effect of completing an additional year of schooling, which implies a loss of $0.045Y_S$ at the sample means.

Impact as a Proportion of Present Discounted Value of Total Lifetime Wealth

Under the above assumptions the sum of these four effects is $1.624Y_S$ at the sample means. This compares with the present discounted value of lifetime income at the sample means:

$$\text{(C-3)} \quad \int_{S}^{R-6} e^{-rt} Y_S dt = \frac{Y_S}{r}\left[e^{-rS} - e^{-r(R-6)}\right]$$

which equals $12.402Y_S$ at the sample means. Therefore the sum of these four effects is 13.1 percent of lifetime wealth at the sample means.

References

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

Alderman, Harold, Jere R. Behrman, Shahrukh Khan, David Ross, and Richard Sabot. 1997. "The Income Gap in Cognitive Skills in Rural Pakistan." *Economic Development and Cultural Change* 46(1):97–122.

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon. 1997. "Child Nutrition, Child Health, and School Enrollment: A Longitudinal Analysis." Policy Research Working Paper 1700. World Bank, Policy Research Department, Washington, D.C. Processed.

Alderman, Harold, Jere R. Behrman, David Ross, and Richard Sabot. 1996a. "Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy." *Journal of Human Resources* 31(1):229–54.

———. 1996b. "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labor Market." *Oxford Bulletin of Economics and Statistics* 58(1):29–55.

Barros, Ricardo, and David Lam. 1996. "Income and Educational Inequality in Children's Schooling Attainment." In Nancy Birdsall and Richard H. Sabot, eds., *Opportunity Foregone: Education in Brazil.* Baltimore, Md.: Johns Hopkins University Press for the Inter-American Development Bank.

Becker, Gary S. 1967. *Human Capital and the Personal Distribution of Income: An Analytical Approach.* Woytinsky Lecture. Ann Arbor: University of Michigan Press. Republished in Gary S. Becker. 1975. *Human Capital.* Second Edition. New York: National Bureau of Economic Research.

Behrman, Jere R., and Nancy Birdsall. 1983. "The Quality of Schooling: Quantity Alone Is Misleading." *American Economic Review* 73(5):928–46.

Behrman, Jere R., Nancy Birdsall, and Robert Kaplan. 1996. "The Quality of Schooling and Labor Market Outcomes in Brazil: Some Further Explorations." In Nancy Birdsall and Richard Sabot, eds., *Opportunity Forgone: Education in Brazil.* Baltimore, Md.: Johns Hopkins University Press.

Behrman, Jere R., Masako Ii, and David Murillo. 1995. "How Family and Individual Characteristics Affect Schooling Demands in Urban Bolivia: Multiple Schooling Indicators, Unobserved Community Effects, Nonlinearities, and Interactions." UPADE/Grupo Social, La Paz, Bolivia. Processed.

Behrman, Jere R., and James C. Knowles. 1997. "How Strongly Is Child Schooling Associated with Household Income?" Department of Economics, University of Pennsylvania, Philadelphia. Processed.

Behrman, Jere R., and Mark R. Rosenzweig. 1999. "'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates." *Economics of Education Review* 18(2):159–67.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102(6):1131–74.

———. 1996. "College Choice and Wages: Estimates Using Data on Female Twins." *Review of Economics and Statistics* 73(4):672–85.

Behrman, Jere R., and Paul Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics* 4(4):S121–45.

———. 1989. "Is Schooling 'Mostly in the Genes'? Nature-Nurture Decomposition with Data on Relatives." *Journal of Political Economy* 97(6):1425–46.

Behrman, Jere R., and Barbara L. Wolfe. 1987. "Investments in Schooling in Two Generations in Pre-Revolutionary Nicaragua: The Roles of Family Background and School Supply." *Journal of Development Economics* 27(1–2):395–420.

Birdsall, Nancy. 1985. "Public Inputs and Child Schooling in Brazil." *Journal of Development Economics* 18(1):67–86.

Birdsall, Nancy, and François Orivel. 1996. "Demand for Primary Schooling in Rural Mali: Should User Fees Be Increased?" *Education Economics* 4(3):279–96.

Card, David, and Alan B. Krueger. 1996. "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina." *Journal of Economic Perspectives* 10(4):31–50.

Case, Anne, and Angus Deaton. 1996. "School Quality and Educational Outcomes in South Africa." Department of Economics, Princeton University, Princeton, N.J. Processed.

Cochrane, Susan Hill, Kaplana Mehra, and Ibrahim Taha Osheba. 1986. *The Educational Participation of Egyptian Children.* Education and Training Series Report EDT45. Washington, D.C.: World Bank.

Datcher, Linda. 1982. "Effect of Community and Family Background on Achievement." *Review of Economics and Statistics* 64(1):32–41.

Datcher Loury, Linda, and David Garman. 1995. "College Selectivity and Earnings." *Journal of Labor Economics* 13(2):289–308.

Deolalikar, Anil B. 1993. "Gender Differences in the Returns to Schooling and School Enrollment Rates in Indonesia." *Journal of Human Resources* 28(4):899–932.

———. 1997. "The Determinants of Primary School Enrollment and Household Expenditures in Kenya: Do They Vary by Income?" Department of Economics, University of Washington, Seattle. Processed.

DeTray, Dennis. 1984. "Schooling in Malaysia: Historical Trends and Recent Enrollments." Report N-2011-AID. Rand Corporation, Santa Monica, Calif. Processed.

Duraisamy, Palanigounder. 1988. "An Econometric Analysis of Fertility, Child Schooling, and Labour Force Participation of Women in Rural Indian Households." *Journal of Quantitative Economics* 4(2):293–316.

Friedman, Milton. 1962. *Capitalism and Freedom.* Princeton, N.J.: Princeton University Press.

Glewwe, Paul. 1996. "The Relevance of Standard Estimates of Rates of Return to Schooling for Education Policy: A Critical Assessment." *Journal of Development Economics* 51(2):267–90.

Glewwe, Paul, Margaret Grosh, Hanan Jacoby, and Marlaine Lockheed. 1995. "An Eclectic Approach to Estimating the Determinants of Achievement in Jamaican Primary Education." *The World Bank Economic Review* 9(2):231–58.

Glewwe, Paul, and Hanan Jacoby. 1994. "Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana." *Journal of Human Resources* 29(3):842–64.

———. 1995a. "An Economic Analysis of Delayed Primary School Enrollment and Childhood Malnutrition in a Low-Income Country." *Review of Economics and Statistics* 77(1):156–69.

———. 1995b. "School Enrollment and Completion in Vietnam: An Investigation of Recent Downward Trends." World Bank, Development Research Group, Washington, D.C. Processed.

Glewwe, Paul, Hanan Jacoby, and Elizabeth King. 1998. "Estimating an Economic Model of Nutrition and Learning Using Longitudinal Data." World Bank, Development Research Group, Washington, D.C. Processed.

Handa, Sudhanshu. 1994. "The Determinants of Teenage Schooling in Jamaica: Rich vs. Poor, Females vs. Males." University of the West Indies, Kingston. Processed.

Hanushek, Eric A. 1995. "Interpreting Recent Research On Schooling in Developing Countries." *The World Bank Research Observer* 10(2):227–46.

Hauser, Robert M., and Thomas N. Daymont. 1977. "Schooling, Ability, and Earnings: Cross-Sectional Findings 8 to 14 Years after High School Graduation." *Sociology of Education* 50(July):182–206.

Hill, Martha S., and Greg J. Duncan. 1987. "Parental Family Income and the Socioeconomic Attainment of Children." *Social Science Research* 16(1):39–73.

Hossain, Shaikh. 1989. "Effect of Public Programs on Family Size, Child Education, and Health." *Journal of Development Economics* 30(1):145–58.

Ilon, Lynn, and Peter Moock. 1991. "School Attributes, Household Characteristics, and Demand for Schooling: A Case Study of Rural Peru." *International Review of Education* 37(4):429–51.

Jacoby, Hanan. 1994. "Borrowing Constraints and Progress through School: Evidence from Peru." *Review of Economics and Statistics* 76(1):151–60.

King, Elizabeth M. 1995. "Does the Price of Schooling Matter? Fees, Opportunity Costs, and Enrollment in Indonesia." World Bank, Development Research Group, Washington, D.C. Processed.

King, Elizabeth M., and L. A. Lillard. 1987. "Education Policy and Schooling Attainment in Malaysia and the Philippines." *Economics of Education Review* 6(2):167–81.

Kingdon, Geeta Gandhi. 1996. "Student Achievement and Teacher Pay: A Case Study of India." Development Economics Research Programme 74. London School of Economics and Political Science. Processed.

Knight, John B., and Richard H. Sabot. 1990. *Educational Productivity and Inequality: The East African Natural Experiment*. New York: Oxford University Press.

Lavy, Victor. 1996. "School Supply Constraints and Children's Educational Outcomes in Rural Ghana." *Journal of Development Economics* 51(2):291–314.

Lavy, Victor, Jennifer Spratt, and Nathalie Leboucher. 1997. "Patterns of Incidence and Change in Moroccan Literacy." *Comparative Education Review* 41(2):120-41.

Lillard, Lee, and Robert Willis. 1994. "Intergenerational Education Mobility: Effects of Family and State in Malaysia." *Journal of Human Resources* 29(4):1126–66.

Moffitt, Robert A. 1996. "Introduction." Symposium on School Quality and Educational Outcomes. *Review of Economics and Statistics* 78(4):559–61.

Moll, Peter G. 1996. "Primary Schooling, Cognitive Skills, and Wages in South Africa." Chicago. Processed.

Montgomery, Mark R., and Aka Kouame. 1993. "Fertility and Child Schooling in Côte d'Ivoire: Is There a Tradeoff?" Department of Economics, State University of New York at Stony Brook. Processed.

Moock, Peter R., and Joanne Leslie. 1986. "Childhood Malnutrition and Schooling in the Terai Region of Nepal." *Journal of Development Economics* 20(1):33–52.

Parish, William L., and Robert J. Willis. 1993. "Daughters, Education, and Family Budgets: Taiwan Experiences." *Journal of Human Resources* 28(4):863–98.

Patrinos, Harry Anthony, and George Psacharopoulos. 1993. "Socioeconomic and Ethnic Determinants of Grade Repetition in Bolivia and Guatemala." World Bank, Human Development Network Education Team, Washington, D.C. Processed.

———. 1995. "Educational Performance and Child Labor in Paraguay." *International Journal of Educational Development* 15(1):47–60.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22(9):1325–44.

Psacharopoulos, George, and Hongyo Yang. 1991. "Educational Attainment among Venezuelan Youth: An Analysis of Its Determinants." *International Journal of Educational Development* 11(4):289–94.

Rosenzweig, Mark R. 1995. "Why Are There Returns in Schooling?" *American Economic Review* 85(2):153–58.

Rosenzweig, Mark R., and Kenneth J. Wolpin. 1986. "Evaluating the Effects of Optimally Distributed Public Programs." *American Economic Review* 76(3):470–87.

Sathar, Zeba A., and Cynthia B. Lloyd. 1994. "Who Gets Primary Schooling in Pakistan: Inequalities among and within Families." *Pakistan Development Review* 33(2):103–34.

Singh, Ram D. 1992. "Underinvestment, Low Economic Returns to Education, and the Schooling of Rural Children: Some Evidence from Brazil." *Economic Development and Cultural Change* 40(3):645–64.

Tansel, Aysit. 1997. "School Attainment, Parental Education, and Gender in Côte d'Ivoire and Ghana." *Economic Development and Cultural Change* 45(4):825–56.

Thomas, Duncan, Robert F. Schoeni, and John Strauss. 1996. "Parental Investments in Schooling: The Roles of Gender and Resources in Urban Brazil." Rand Corporation, Santa Monica, Calif. Processed.

van de Walle, Dominique, and Kimberly Nead, eds. 1995. *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press for the World Bank.

World Bank. 1995a. "Vietnam Poverty Assessment and Strategy." World Bank Report 13442-VN. Washington, D.C. Processed.

———. 1995b. *World Development Report 1995: Workers in an Integrating World*. New York: Oxford University Press.

———. 1996. "Vietnam Education Financing Sector Study: A Sector Report." Report 15925-VN, Washington, D.C. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL 13, NO. 2· 257-73

Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture

Peter Lanjouw and Martin Ravallion

Assessments of the distributional effects of public spending reforms have generally been based on average rates of program participation by income or expenditure group. This practice can be deceptive because the socioeconomic composition of participants can change as a social program expands or contracts. The geographic variation found in 1993–94 household survey data for rural India is used to estimate the marginal odds of participating in schooling and antipoverty programs. The results suggest early capture of these programs by the nonpoor. It is shown that conventional methods for assessing benefit incidence underestimate the gains to the poor from higher public outlays and underestimate the losses from cuts.

Benefit incidence analysis is widely used to assess the distributional impact of public spending. Typically, the average participation rate for each public program is tabulated against household income or expenditure per capita, using data from a household survey. The public subsidy rate for each program is then applied to the participation rates to determine the incidence of program spending—to assess, for example, whether the poor gain more than the rich. Discussions of the method and examples for developing countries can be found in Meerman (1979), Selowsky (1979), Meesook (1984), Hammer, Nabi, and Cercone (1995), Selden and Wasylenko (1995), van de Walle (1995, 1998), and Demery (1997).

This article examines whether this now-standard methodology provides a reliable guide to the distributional impact of public spending reforms. These reforms typically entail marginal changes in spending across one or more programs. But benefit incidence calculations are based on averages. To see why this difference matters, consider a publicly supplied private good, such as subsidized schooling or food. Unlike a private good obtained in a competitive market, one cannot buy or sell as much of a publicly supplied good as one wants—consumers are quantity constrained. And unlike a pure public good (in which everyone faces the same quantity constraint), the way in which a program's outlays are allocated

Peter Lanjouw and Martin Ravallion are with the Development Research Group at the World Bank.
This article was prepared as an input to the World Bank's 1998 Poverty Assessment for India. The authors gratefully acknowledge the support of the World Bank's Research Committee (under RPO 681-39) and the comments of Zoubida Allaohua, Francisco Ferriera, Jenny Lanjouw, Valerie Kozel, Ricardo Paes de Barros, Lant Pritchett, Kalanidhi Subbarao, Dominique van de Walle, and three anonymous referees.

across consumers comes into play. That allocation is typically the outcome of a political process. The distributional impact of a change in supply will then depend on the abilities of different socioeconomic groups to influence that political process. Those abilities, in turn, will depend in part on the history of allocations made under the program at the time reforms commence.

In these circumstances the averages used in traditional benefit incidence analysis can be deceptive in predicting the marginal changes that would arise from public spending reforms. For example, suppose that the nonpoor were able to capture most of the benefits when the program was first introduced but are now virtually satiated at the margin. Then the poor will gain a large share of the extra benefits from an expansion of the program and may lose heavily from its contraction, even though they receive a small share of the average benefits.

We use household survey data on participation in primary school and the main antipoverty programs in rural India. First, we estimate average participation rates, using standard methods. We then estimate marginal participation rates, exploiting the fact that there are large differences across Indian states in the scale of each type of public program. Thus we are able to compare the average and marginal odds of participation and so test for bias in estimates of the distributional impact of spending changes based on standard benefit incidence analysis.

I. Does the Composition of Program Participation Vary with Program Size?

The population is divided into two or more groups according to consumption expenditure (or some other welfare indicator). For each expenditure group we know the participation rate in a public program. We examine the effect of changing the overall size of the program as part of a public spending reform. If the group-specific participation rates stay the same, then the composition of program participation is said to be *homogeneous* (strictly, homogeneous of degree one). This means that if, say, 40 percent of participants are poor when 1,000 participants are covered, then 40 percent are also poor when 2,000 are covered. Policy conclusions drawn from standard benefit incidence analyses implicitly assume homogeneity. But there is no obvious reason why this should hold.

Nonhomogeneous participation can arise when the poor are able to capture program benefits at certain times in the program's history, but not others. This can occur even when the program is ostensibly targeted solely to the poor. Either the government is unable to target perfectly—because of information or incentive problems—or targeting only the poor is not a political equilibrium in that the government relies on the support of the nonpoor.[1] The timing of program capture will depend on how the costs and benefits of participation vary with the scale of the program. Social programs invariably impose costs on participants. These costs could take the form of taxes or fees for financing the programs. Or they could be hidden,

1. For an overview of these issues, see Besley and Kanbur (1993); for a model of the political economy of targeting in which perfect targeting is not an equilibrium, see Gelbach and Pritchett (1997).

deadweight losses arising from financing methods or from participation itself. Such losses include the opportunity cost to parents of children's time in school or the cost to a nonpoor person of illegally securing participation in a means-tested program. These costs are likely to vary with the scale of the program.

For example, the geography of program placement can generate nonhomogeneous participation. Consider a country in which poor areas tend to be more remote and hence less convenient for program staff to reach. Initial placement is in less remote areas. When the program is first set up, the poor will find it more costly to participate than when the program eventually expands into remote areas. The nonpoor will be able to capture the program early because it is more accessible to them. But after some point marginal gains start to favor the poor.

General equilibrium effects could also produce rising costs of participation that differ between the poor and the nonpoor. For example, although a small public works program may not affect wages in alternative work, a large program may bid up wages and hence increase the expected forgone income of program participants. To the extent that the nonpoor face better chances of getting work, they will find the public employment program less and less attractive as it expands. Again, early capture by the nonpoor can be expected.

Late capture is also possible. For example, it may be far easier for the (theoretically ineligible) nonpoor to bribe officials to gain access to the program once it is widely available, when a nonpoor participant would be less conspicuous.

A Model

We illustrate these arguments more formally in a simple political economy model. The model assumes that the government wants to reduce poverty, but that it faces a political constraint in that it cannot impose a welfare loss on the nonpoor. The program does impose costs on the nonpoor (such as taxes for financing the program) that depend on average participation of the poor and nonpoor. Let the cost to the nonpoor be $C(X)$, where X is average participation and C is a positive, smoothly increasing function with $C(0) = 0$. (Other factors may enter this cost function, such as the proportion of the population who is not poor, but these can be ignored for the purpose of our model.) Marginal cost is $C'(X)$, which could either increase or decrease with X (depending on whether C is convex or concave). The program's benefits are allocated between nonpoor households, who participate at the rate X_n, and poor households, who participate at the rate X_p. The corresponding per capita benefits are $B(X_n)$ and $B(X_p)$, where the function B is increasing from $B(0) = 0$. (To simplify the notation, we assume that B is the same for the poor and the nonpoor. Allowing the functions to differ does not affect our analysis, although it could affect the assessment of benefit incidence in practice.) The utility of a nonpoor household is $U[Y_n + B(X_n) - C(X)]$, where Y_n is the household's exogenous income, and the function U is strictly increasing.

Political feasibility requires that the nonpoor do not lose from the program. In other words, a necessary condition for the program to continue is that:

$$U[Y_n + B(X_n) - C(X)] \geq U(Y_n), \tag{1}$$

where $U(Y_n)$ is the utility of the nonpoor without the program ($X_n = X = 0$). Since the government values gains to the poor, the political economy constraint in equation 1 will be binding in equilibrium. (If it were not binding, there would be a politically feasible change that benefited the poor.) Because equation 1 must then hold with equality, we have:

$$B(X_n) = C(X). \tag{2}$$

Solving equation 2 for

$$X_n = \Psi(X) \tag{3}$$

tells us how program participation by the nonpoor varies with average participation, given the political economy constraint. The participation rate of the poor is also a function of X, namely:

$$X_p = [X - N_n\Psi(X)] / N_p, \tag{4}$$

where N_n and N_p are the proportions of the population who are nonpoor and poor, respectively. The marginal change in participation by the nonpoor as the program expands—the marginal participation rate of the nonpoor—is given by

$$\Psi'(X) = C'(X) / B'(X_n). \tag{5}$$

To see what this model implies for the timing of program capture, consider first the special case in which B is linear (constant marginal benefits). Then, it is plain from equation 5 that X_n will be concave (convex) in X—implying that the marginal gains to the poor tend to rise (fall) as the program expands—whenever the marginal cost to the nonpoor is decreasing (increasing). Early (late) capture will occur when the cost function is concave (convex).

With declining marginal benefits to the nonpoor (B is concave), a convex cost function still implies late capture. For early capture the cost function must be sufficiently concave. Differentiating equation 5 with respect to X, we can see that

$$\Psi''(X) = [C''(X)\, B'(X_n) - X_n'(X)\, C'(X)\, B''(X_n)] / B'(X_n)^2. \tag{6}$$

For early capture, $\Psi''(X) < 0$, to be the only politically feasible option in this model, it is necessary and sufficient that the (absolute) elasticity of marginal cost, $-XC''(X) / C'(X)$, exceed the elasticity of participation by the nonpoor, $X\Psi'(X) / X_n$, times the elasticity of the marginal benefit from the program's allocation, $-X_nB''(X_n) / B'(X_n)$.

Early versus Late Capture

The above model illustrates how, for public programs with relatively large start-up costs, early capture by the nonpoor may be the only politically feasible option, particularly when start-up costs must be financed domestically. For example, in exchange for paying taxes to cover these costs, the nonpoor may de-

mand a sizable share of the initial benefits, such as by requiring that the program not be located in inaccessible, poor areas. Only later, when the marginal costs of program expansion are lower, will it be politically feasible to reach the poor.

Figure 1 illustrates the case of early capture. The figure shows the group-specific participation rate as it varies with the average rate, that is, the function $\Psi(X)$ and $2X - \Psi(X)$ for the nonpoor and poor, respectively. (For convenience, the figure is drawn assuming that there are equal numbers of poor and nonpoor.) The nonpoor capture the bulk of the gains initially but become progressively satiated. Imagine we are at point A, where the poor and nonpoor are participating equally. Given the average participation rates, the standard benefit incidence analysis would conclude that expanding the program would not benefit the poor relative to the nonpoor. This conclusion is plainly wrong: most of the gains from a small aggregate expansion at point A (to, say, point B) would go to the poor. Similarly, a small cut in the program at point A would be borne mostly by the poor.

Now consider the case of late capture as illustrated in figure 2. Suppose we are initially at the average participation rate A. Although the poor are participating more than the nonpoor, it is the nonpoor who would capture most of the gains from increasing the level of average participation (to, say, point B) and incur most of the loss from retrenchment.

We would expect some public programs to be more like the early capture model and others to be more like the late capture model. For example, it is likely that children of better-off parents will be the first to gain from public spending

Figure 1. *Early Capture*

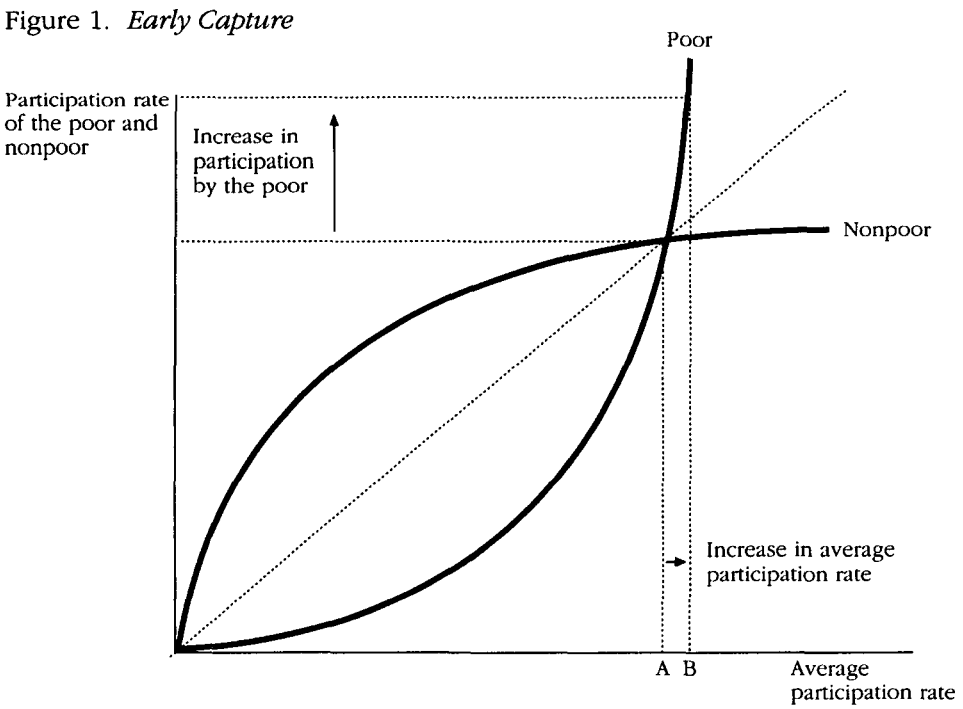
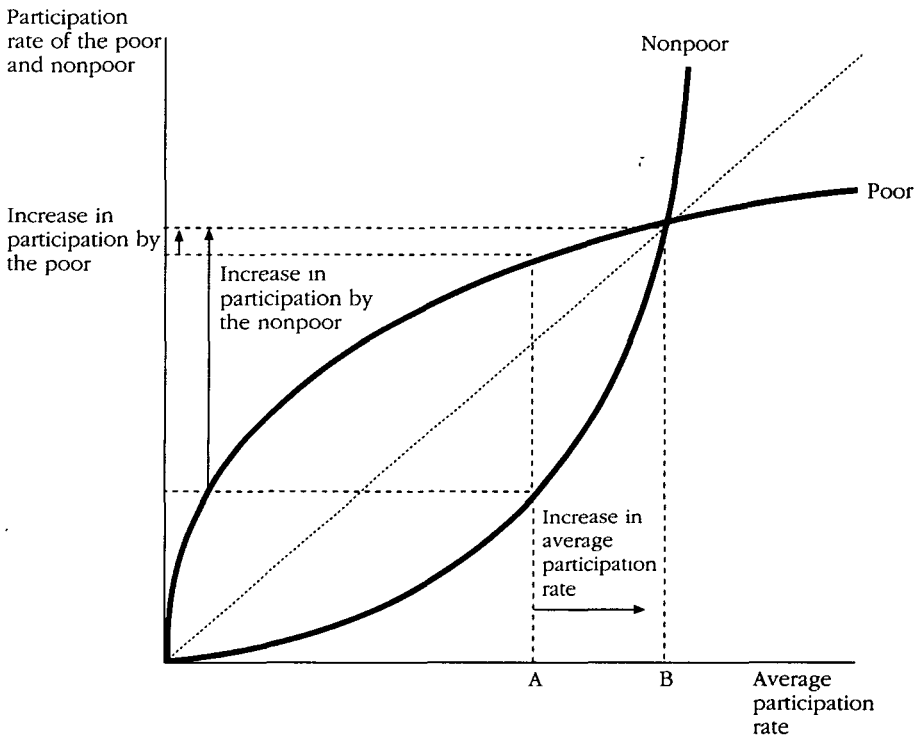


Figure 2. *Late Capture*



on education (early capture). But they will become satiated in due course, at which point marginal gains will go to the poor. By contrast, consider a food rationing scheme that is initially targeted to the poor. In time, political pressures to favor middle-income groups may lead to higher marginal gains for the nonpoor (late capture).

This discussion suggests that the political economy of program capture may contain important clues to some poorly understood issues concerning the welfare impacts of changes in public spending. One such issue is whether or not there are politically feasible ways of protecting the poor from cuts in social spending. Ravallion (forthcoming) explores this issue further, providing evidence that spending cuts were borne more heavily by the poor in an antipoverty program in Argentina.

The political economy of program capture can also help us to understand the empirical relationship between intercountry differences in public spending on social programs and aggregate outcome indicators. Aggregate data often show this relationship to be weak. But Bidani and Ravallion (1997) have developed an econometric specification that allows them to compare the effects of differences in countries' public health spending on health indicators for the poor and the nonpoor. They find that differences in health spending matter far more to the

health outcomes of the poor than of the nonpoor. This is what we would expect if the nonpoor capture the inframarginal gains.

II. Measuring Participation Rates

The average participation rate is the proportion of households in a given expenditure or income quintile that participates in the program. The average odds ratio of participation (herein, the average odds of participation) is defined as the ratio of the participation rate of one quintile to the overall average. The marginal odds ratio of participation (herein, the marginal odds of participation) is the increment in the program participation rate of a given quintile when there is a change in aggregate participation. Differences between the marginal and average odds of participation reflect differences in the incidence of inframarginal spending. Only if participation is homogeneous will the two be everywhere the same.

The average odds of participation can be calculated from the survey data in a straightforward way. How can we estimate the marginal odds of participation? We have only a single cross-sectional survey (as is typically used in benefit incidence analysis), including data on program participation across geographic areas ("regions") within states. We can readily calculate the average participation rate for a given program for each quintile and each region. The participation rate for a given quintile varies across regions according to the level of public spending on the program in the state to which each region belongs. To estimate the marginal odds of participation by program and expenditure quintile, we can regress the quintile-specific participation rates across regions on the state's average participation rate (for all quintiles and all regions) for each program.[2]

Ordinary least squares will give a biased estimate of the marginal odds of participation, because the specific region and quintile participation rates (on the left side) are implicitly included when calculating the state's overall mean participation rate (on the right side). To deal with this problem, we use the "leave-out mean" as an instrumental variable for the state's average participation rate. The leave-out mean is the mean for the state excluding the specific region and quintile participation rates that correspond to each observation in the data. For example, if we are using the data for quintile three in region five within state ten, then the leave-out mean is the average for all regions and quintiles within state ten, excluding quintile three in region five.

How can we interpret the marginal odds of participation? As with average participation rates, to infer overall incidence we must also know the program's subsidy rate. In conventional benefit incidence analysis the subsidy rate for each program is typically assumed to be constant across geographic areas and income groups. For example, it is assumed that the cost to the government is the same if

2. Our method appears to be new, although models of outcome indicators stratified by socioeconomic group are familiar from past work. Deolalikar (1995), for example, studies the cross-sectional differences in health outcomes for poor and nonpoor children in Indonesia.

a poor person participates or if a rich person participates. Using that assumption, we can infer from the marginal odds of participation how an increase in public spending on a given program will affect each quintile. We will be able to make partial tests of that assumption.

III. Data

Our analysis is based on the household-level data from India's National Sample Survey (NSS) for 1993–94. This survey includes standard data on consumption expenditures, demographics, and educational attainment, including school enrollment. In addition, this round of the NSS also asked about participation in three antipoverty programs: public works schemes, a means-tested credit scheme called the Integrated Rural Development Programme (IRDP), and a food rationing scheme called the Public Distribution System (PDS). We collate data on participation in these programs with data on total consumption expenditures per person at the household level. A household is said to have participated in a public works program if any household member worked for at least 60 days on public works in the preceding 365 days. A household participated in the IRDP program if it received any assistance from IRDP in the past five years. And a household participated in the PDS program if it purchased any commodity from a ration or fair price shop in the past 30 days.

We ranked sampled households by total consumption expenditure per person (including imputed values of consumption from own production) normalized by state-specific poverty lines. Quintiles are defined over the entire rural population, with an equal number of people in each. Thus the poorest quintile includes the poorest 20 percent of the nation's rural population in terms of consumption expenditure per capita.

These data are not ideal. The relationship between participating in the IRDP over the past five years and consumption expenditure over the past month may be a poor indication of the program's incidence because participants' living standards may have changed considerably over five years. There are also concerns about the adequacy of participation as an indicator of use of the PDS. For example, the rich may buy only a small quantity of the rationed good (although this conjecture is not consistent with other data on the incidence of PDS purchases; see Radhakrishna and Subbarao 1997). Another possibility is that the individual participant may have a different standard of living than the household as a whole. In the case of public works projects the data will likely include people who participate in public works projects but are not part of antipoverty programs.

The sample size (for rural areas) of the 1993–94 NSS is 61,464 households. We conduct the analysis at the level of the NSS region, of which there are 62 in India, spanning 19 states. Each NSS region belongs to only one state. So, in the basic model, for any given combination of quintile and program, we regress the sample participation rates from the 62 regions on the average participation rate (irrespective of quintile) from each of the 19 states.

Recall that in benefit incidence analysis predicting the incidence of changes in public spending from the estimated marginal odds of participation requires the assumption that the average subsidy rate (given participation) is constant. For public works programs and the IRDP there is no obvious way in which the subsidy rate conditional on participation could vary by household expenditure per person within a given state. But we can expect variation between states. For the PDS income effects on demand for the rationed goods could create differences in the subsidy rate across quintiles within a given state.

We are able to test the assumption of a constant subsidy rate for public works programs, the IRDP, and primary school enrollment, but we do not have the data to do so for the PDS. For each program we regress the state's per capita spending on its average participation rate plus four of the quintile-specific participation rates. We are unable to reject the null hypothesis that the parameter estimates on the quintile-specific participation rates are jointly zero. The probability values for the *F*-tests are, respectively, 0.57, 0.11, and 0.23 for primary schooling, public works, and the IRDP. The coefficients on a state's average participation rate are highly significant, as we would expect. Thus we find no evidence that the subsidy rate varies significantly by quintile. This helps us to justify the constant-subsidy assumption when interpreting our results.

IV. Results

We begin with primary school enrollment for children ages 5–9 years (table 1). Our calculated enrollment rates from the NSS are appreciably lower than those obtained from schools themselves, on which official enrollment rates are based. The official primary enrollment rate for India was higher than 100 percent in 1993. Although there are differences in definition—for example, we have confined attention to the age group 5–9 and so excluded late starters—there are reasons to believe that biases in official sources lead to overestimation of enrollments in India (Kingdon 1996).

Enrollment rates rise with household expenditures per capita nationally and in all states, and they tend to be higher for boys than for girls. (Lanjouw and Ravallion 1998 give full results by state.) But there are marked differences among states. In Kerala, for example, there is less difference among the quintiles and between boys and girls (indeed, enrollment rates are slightly higher for girls from the poorest quintile) than in Bihar or Punjab.

The average odds of enrollment suggest that subsidies to primary schools mildly favor the nonpoor. Notice, however, that we cannot split public and private schooling in the data; public school enrollment may be lower for the nonpoor. While the average enrollment rate is higher for the richest quintile, the relationship between region-specific enrollment rates and states' average rates is steeper for the poorest quintile (Lanjouw and Ravallion 1998 give scatter plots). Thus the marginal odds of participation are higher for the poor, even though their average participation rate is lower.

Table 1. *Average Primary School Enrollment in Rural India*

	Boys		*Girls*		*Total*	
Quintile	*Enrollment rate (percent)*	*Average odds of enrollment (mean = 1.0)*	*Enrollment rate (percent)*	*Average odds of enrollment (mean = 1.0)*	*Enrollment rate (percent)*	*Average odds of enrollment (mean = 1.0)*
1 (poorest)	42.6	0.75	31.6	0.66	37.2	0.71
2	53.4	0.93	43.1	0.91	48.6	0.90
3	60.5	1.07	50.3	1.06	55.8	1.08
4	66.1	1.16	58.6	1.26	62.6	1.21
5	69.9	1.23	65.2	1.38	67.7	1.31

Note: The table gives the average primary school enrollment rates as a percentage of children aged 5–9 and the average odds of enrollment, defined as the ratio of the quintile-specific enrollment rate to the mean enrollment rate. Households are ranked by total expenditure per person in forming the quintiles.

Source: Authors' calculations based on India's 1993–94 National Sample Survey.

We estimate the marginal odds of being enrolled by regressing the participation rates of each quintile across regions on the states' average participation rates (table 2). The numbers in table 2 can be interpreted as the gain in subsidy incidence per capita for each quintile from a one-rupee increase in aggregate spending on each program. For example, if an extra 100 rupees per capita is spent on primary schools, public expenditures per capita going to the poorest quintile will rise by 110 rupees. These are instrumental variables estimates in which the leave-out mean is the instrument for the state average participation rate.

The estimates of the marginal odds of participation suggest that expanding primary schooling would be decidedly propoor at the margin. (As in standard benefit incidence analysis, future earnings gains from better education are not factored into this calculation.) The implication for the incidence of subsidies to primary education is clear (given our inability to reject the constant unit-subsidy assumption). The average odds of participation given in table 1 suggest that the share of the total subsidy going to the poorest quintile is only 14 percent (0.71 times one-fifth). By contrast, the marginal odds of participation, given in table 2, imply that the poorest quintile would obtain about 22 percent of an increase in the total subsidy going to primary education.

There is also a gender difference between the average and marginal odds of participation. The average odds of poor children attending school are higher for boys (0.75 compared with 0.66 for girls). However, the marginal odds are almost identical (1.09 compared with 1.08). These results are clearly not consistent with homogeneous participation. Marginal gains from expanding primary schooling in rural India are much better distributed than average gains.

Turning now to the antipoverty programs, we see that for both public works programs and the IRDP participation rates fall as expenditures per person rise (table 3). But the rate of decline is not large; the odds of the poorest quintile participating in public works programs are 1.23 compared with 0.83 for the

Table 2. *Marginal Odds of Primary School Enrollment in Rural India*

Quintile	*Boys*	*Girls*	*Total*
1 (poorest)	1.09	1.08	1.10
	(6.90)	(9.65)	(8.99)
2	0.91	0.91	0.97
	(6.05)	(6.99)	(7.92)
3	0.92	0.84	0.87
	(5.85)	(6.54)	(7.65)
4	0.66	0.66	0.67
	(4.10)	(4.28)	(4.77)
5	0.53	0.70	0.67
	(4.08)	(5.53)	(5.69)

Note: The table gives the instrumental variables estimates of the regression coefficients of the quintile-specific primary school enrollment rates across regions on the average rate by state. The leave-out mean enrollment rate is the instrument for the actual mean. The numbers in parentheses are *t*-ratios.

Source: Authors' calculations based on the 1993–94 National Sample Survey.

Table 3. *Average Participation Rates for India's Main Antipoverty Programs in Rural Areas*

	Public works programs		*Integrated Rural Development Program*		*Public Distribution System*	
Quintile	*Participation rate (percent)*	*Average odds of participation (mean = 1.0)*	*Participation rate (percent)*	*Average odds of participation (mean = 1.0)*	*Participation rate (percent)*	*Average odds of participation (mean = 1.0)*
1 (poorest)	5.0	1.23	6.5	1.03	69.5	0.92
2	4.6	1.13	7.1	1.13	76.7	1.01
3	4.2	1.04	6.4	1.03	77.9	1.03
4	3.5	0.86	6.0	0.96	78.1	1.00
5	3.4	0.83	5.6	0.89	76.1	1.00

Note: The table gives the average participation rates and the average odds of participation, defined as the ratio of the quintile-specific participation rate to the mean participation rate for each program.

Source: Authors' calculations based on the 1993–94 National Sample Survey.

richest quintile. The rate of decline is even lower for the IRDP. Participation rates among the richest 20 percent of the population are high even for public works programs. For the PDS the participation rate is lowest for the poorest quintile and highest for the second-richest quintile.

Keep in mind that these figures are national aggregates. We find large differences among states (full details are available in Lanjouw and Ravallion 1998). In Orissa, for example, the proportion of households in the poorest quintile participating in public works programs is more than four times that of the richest quintile; the odds of the poorest quintile participating are 1.6, well above the national mean, 1.23 (table 3). In Maharashtra the odds of the poorest quintile participating in public works programs are also well above the national average. At the other extreme, in states such as Andhra Pradesh, Gujarat, Kerala, and Tamil Nadu, the poorest quintile has lower than average participation rates.

The marginal odds of participation for the poorest quintile are highest for public works programs, while the IRDP dominates for the three middle quintiles; the marginal odds of participation for the richest quintile are higher for the PDS (table 4). (The regional plots for all programs and the poorest and richest quintiles are available from the authors.) The estimated marginal odds of participation broadly confirm the conclusion drawn from the average odds of participation—public works programs are best at reaching the poorest, while the IRDP is more effective at reaching the middle quintiles, including those living at India's poverty line (at roughly the fortieth percentile).

The difference between the marginal odds of participation for any two programs gives the estimated gain from transferring one rupee between them. For example, transferring 100 rupees per capita from the PDS to public works programs would raise public spending per capita on the poorest quintile by 10 rupees (116 – 106 = 10, using the basic model).

Table 4. *Marginal Odds of Participation for India's Main Antipoverty Programs in Rural Areas*

Quintile	*Public works programs*	*Integrated Rural Development Program*	*Public Distribution System*
1 (poorest)	1.16	1.11	1.06
	(3.27)	(15.49)	(8.14)
2	0.93	1.28	0.99
	(3.64)	(17.73)	(7.26)
3	0.80	1.21	0.91
	(2.98)	(23.52)	(6.88)
4	0.92	0.96	0.86
	(4.32)	(19.09)	(7.16)
5	0.55	0.39	0.81
	(3.29)	(8.06)	(6.27)

Note: The table gives the instrumental variables estimates of the regression coefficients of the quintile-specific program participation rates across regions on the average rate by state for that program. The leave-out mean participation rate is the instrument for the actual mean. The numbers in parentheses are *t*-ratios.

Source: Authors' calculations based on the 1993–94 National Sample Survey.

For both the public works programs and the IRDP it is notable that the marginal odds of participation tend to fall more rapidly moving from the poorest to the richest quintile than do the average odds. Thus the average odds of participation underestimate how propoor an increase in average spending on each of these programs will be. This difference is particularly strong for the IRDP: the average odds of participation are only slightly higher for the poorest quintile than for the richest (1.03 and 0.89), whereas the marginal odds are much higher for the poorest quintile than for the richest (1.11 and 0.39). Compared with the average odds of participation, the marginal odds of participation raise the share of total IRDP spending imputed to the poorest 40 percent of the population by 11 percent, while that imputed to the richest 20 percent falls by 56 percent. For the PDS, however, there is less difference between the average and marginal odds, so the former are a better guide to PDS incidence relative to the other programs.

As with primary school enrollment, these results are inconsistent with the homogeneity assumption. Unlike schooling, for the antipoverty programs studied here, both average and marginal odds of participation tend to be higher for the poor. But, like schooling, marginal gains from these programs tend to be better distributed than average gains.

V. Caveats

It is worth reviewing some of the assumptions that underpin our efforts to estimate the marginal incidence of spending on these programs. We estimate the marginal odds of participation by regressing quintile-specific participation rates across regions on the state's average participation rate (all quintiles, all regions) for each program. We do not include any other explanatory variables (such as state-level poverty rates). To the extent that other variables affect quintile-specific participation rates via their influence on states' average participation rates, they are not of concern, because it is the effect of expansion in the overall size of the programs that we are interested in evaluating. There is, however, one way in which our specification may be unsatisfactory. In the first section we outlined how political economy factors could influence program incidence by determining the timing of program capture. Yet by not including political economy variables as separate explanatory variables in our regression, we implicitly assume that they are identical across states or vary in ways that are uncorrelated with state-level average participation rates.

We are unable to control for regional fixed effects in our estimations because we do not have time series data. But we are able to examine the extent to which interstate differences in average participation rates account for interstate differences in quintile-specific participation rates. First, for each program we reestimate each model by regressing quintile- and region-specific participation rates on a full set of state dummy variables. (The effect of the state participation rate is not identified, because it is predicted perfectly by the state dummy variables.) We then compare the R^2 values from these regressions to those that we obtained from our regressions

on state average participation rates. We find that in most cases the R^2 values from the state participation rate specifications are 70 percent or more of those from the state fixed-effect regressions. This suggests that states' average participation rates capture a large share of the variance in the dependent variable attributable to state effects. In the case of primary schooling of boys the R^2 values from our specification decline from about 75 percent of the R^2 values of the state fixed-effect specification for the lowest two quintiles to an average of about 45 percent for the top two quintiles. In the case of public works programs the ratio of R^2 values averages about 50 percent for the three lowest quintiles, rises to 70 percent for the fourth, and then declines to 31 percent for the top quintile.

Second, we examine the residuals from our regression using states' average participation rates to see whether, for any given state and quintile, the average of the residuals across regions is significantly higher or lower than that observed for other states. For example, we ask whether the participation rate of the bottom quintile in Kerala or West Bengal (both of which have had long periods of left-wing governments) is unusually high given the state's average participation rate, reflecting a difference in political economy.

We find no obvious patterns in the residuals. In very few cases (looking at the average residuals per state for each of the quintile-specific and program-specific regressions) does the state's average residual exceed in absolute value the standard error of the regression as a whole. And in the few cases in which this does occur, there is no discernible pattern showing that one state appears to be consistently more effective in reaching a particular quintile across programs. The only pattern that does emerge is for primary school enrollment in both Haryana and Punjab (for boys, girls, and the full sample): the average residuals for the bottom quintile are uniformly negative and larger than one standard error. Assuming that the political economy in these two states is appreciably different, we drop them. Our estimate of the marginal incidence of additional education spending on the poorest quintile is slightly higher than that given in table 2; the marginal odds of participation are 1.16 for boys, 1.12 for girls, and 1.13 overall. For the second poorest quintile the marginal odds of participation are 0.98 for boys, 0.99 for girls, and 1.01 overall. The direction of change strengthens our main result for the comparison of average and marginal odds of participation.

VI. Conclusions

We have used a simple model of the political economy of the timing of program capture to argue that conventional benefit incidence analysis can be deceptive about the distributional impacts of public spending returns. Motivated by this model, we used regional data for India to study how the composition of program participation varies with the size of a social program. This provided a relatively simple method of estimating the marginal odds of program participation. The method can be implemented with the same basic data used in conventional benefit incidence analysis.

Our results for India suggest that average participation rates are not a reliable guide to the distributional impacts of changes in aggregate public outlays or reallocations among programs. Our estimates of the marginal odds of participation broadly confirm the qualitative conclusion drawn from the average participation numbers for the three poverty programs that we studied. However, the average odds of participation greatly understate how propoor extra spending on either public works programs or the means-tested credit scheme is likely to be. Similarly, conventional methods underestimate the loss to the poor from program cuts. The average odds also underestimate how propoor a switch from, say, the Public Distribution System to public works programs would be in India.

In the case of primary schooling the average odds of participation give the wrong qualitative result. Although the average odds of enrollment rise with expenditures per person, the marginal odds fall sharply, indicating that aggregate expansion is decidedly propoor. Indeed, the marginal odds suggest that higher subsidies to primary education are about as propoor as the best programs directed (explicitly) at fighting poverty.

For both primary schooling and poverty programs (except the food rationing scheme) our results are more consistent with the early capture model than with the late capture model. The geographic pattern of participation suggests that the nonpoor tend to be the first to gain when a program is introduced, but that high marginal gains to the poor emerge later.

These findings are tentative. The fact that we had to rely on a single cross-sectional survey meant that we were not able to eliminate the possibility of omitted state-level effects that influence distributional outcomes and are correlated with average program participation rates. Geographic panel data on program participation would allow more robust tests.

To the extent that further work supports our findings, serious doubts are raised about assumptions routinely made in discussions of the distributional impacts of social programs. The timing of program capture can mean that the poor obtain larger gains from extra spending, and are hurt more by cuts, than data on average participation rates would suggest.

References

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

Besley, Timothy, and Ravi Kanbur. 1993. "Principles of Targeting." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor*. Washington, D.C.: World Bank.

Bidani, Benu, and Martin Ravallion. 1997. "Decomposing Social Indicators using Distributional Data." *Journal of Econometrics* 77(1):125–40.

Demery, Lionel. 1997. "Benefit Incidence Analysis." World Bank, Institutional and Social Policy, Africa Technical Family, Washington, D.C. Processed.

Deolalikar, Anil B. 1995. "Government Health Spending in Indonesia: Impacts on Children in Different Economic Groups." In Dominique van de Walle and Kimberly Nead,

eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

Gelbach, Jonah, and Lant Pritchett. 1997. "Redistribution in a Political Economy: Leakier Can Be Better." World Bank, Development Research Group, Washington, D.C. Processed.

Hammer, Jeffrey S., Ijaz Nabi, and James A. Cercone. 1995. "Distributional Effects of Social Sector Expenditures in Malaysia, 1974–89." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

Kingdon, Geeta G. 1996. "Private Schooling in India: Size, Nature, and Equity Effects." Working Paper 72. Development Economics Research Programme, London School of Economics. Processed.

Lanjouw, Peter, and Martin Ravallion. 1998. "Benefit Incidence and the Timing of Program Capture." Policy Research Working Paper 1956. World Bank, Development Research Group, Washington, D.C. Processed.

Lipton, Michael, and Martin Ravallion. 1995. "Poverty and Policy." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics.* Vol. 3. Amsterdam: North-Holland.

Meerman, Jacob. 1979. *Public Expenditure in Malaysia: Who Benefits and Why.* New York: Oxford University Press.

Meesook, Oey A. 1984. *Financing and Equity in the Social Sectors in Indonesia.* World Bank Staff Working Paper 703. Washington, D.C.: World Bank.

Radhakrishna, Rokkam, and Kalanidhi Subbarao. 1997. "India's Public Distribution System: A National and International Perspective." World Bank Discussion Paper 380. World Bank, Poverty Reduction and Economic Management Network, Washington, D.C. Processed.

Ravallion, Martin. Forthcoming. "Is More Targeting Consistent with Less Spending?" *International Tax and Public Finance.*

Selden, Thomas M., and Michael J. Wasylenko. 1995. "Measuring the Distributional Effects of Public Education in Peru." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

Selowsky, Marcelo. 1979. *Who Benefits from Government Expenditures? A Case Study of Colombia.* New York: Oxford University Press.

van de Walle, Dominique. 1995. "The Distribution of Subsidies through Public Health Services in Indonesia, 1978–87." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

———. 1998. "Assessing the Welfare Impacts of Public Spending." *World Development* 26(3):365–79.

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2: 275-302

Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico

William F. Maloney

This article offers an alternative to the traditional dualistic view of the relationship between formal and informal labor markets. For many workers inefficiencies in formal sector protections and low levels of labor productivity may make informal sector employment a desirable alternative to formal sector employment. The analysis offers the first study of worker transitions between sectors using detailed panel data from Mexico and finds little evidence in favor of the dualistic view. Traditional earnings differentials cannot prove or disprove segmentation in the developing-country context. The patterns of worker mobility do not suggest a rigid labor market or one segmented along the formal/informal division.

In developing countries roughly 40 percent of urban workers are not protected by labor legislation and work in small, informal firms.[1] The origins and dynamics of the informal sector have attracted renewed interest for at least two reasons. First, increasing labor market efficiency and flexibility are considered essential complements to the market-based reforms under way throughout the developing world (see, for example, World Bank 1995). To the degree that the informal sector exists because of segmentation driven by government- or union-imposed distortions, its large size stands as a measure of the magnitude of market inefficiencies and required reforms. Second, a related literature with a very different emphasis sees informality as the result of an ongoing effort by large modern enterprises to evade mandated protections through subcontracting to unprotected workers, a process accelerated by heightened global competition in labor-intensive manufactures (see Portes, Castells, and Benton 1989). The existence and behavior of the sector are thus directly relevant to the debate over establish-

1. See, for example, Harris and Todaro (1970), Sabot (1977), Mazumdar (1983), or Fields (1990). See Fields (1990), Tokman (1992), Portes (1994), Rosenzweig (1988), and Thomas (1992) for excellent overviews of the informality literature. See Stiglitz (1974) and Esfahani and Salehi-Isfahani (1989) for efficiency wage models of dualism in developing countries.

William F. Maloney is with the Poverty Reduction and Economic Management Unit of the Latin America and Caribbean Region at the World Bank. The author is grateful to Enríque Dávila Capalleja, Aslan Cohen, René Cortázar, Bill Dickens, Hadi Esfahani, Ariel Fiszbein, Alec Levenson, Frank Lysy, Doug Marcouiller, Gustavo Márquez, David Nielson, Ana Revenga, Eric Rice, Bill Savadoff, and Guilherme Sedlacek for helpful discussions. The author thanks Roberto Flores Lima and Agustín Ibarra Almada of the Mexican Secretariat of Labor and Social Welfare for their advice and help with the project, and the Mexican National Institute of Statistics, Geography, and Information (INEGI) for use of the data.

ing common labor standards through free trade agreements. Yet whatever safeguards may be enshrined in labor codes, firms in developing countries can still employ an effectively unprotected workforce to compete with firms in industrial countries (see Portes, Castells, and Benton 1989; Tokman 1992; and U.S. Department of Labor 1992, 1993).

In both cases the informal sector represents the disadvantaged sector in a dualistic labor market—a view shared by much of the literature. For example, a recent World Bank document argues that, "Protected workers in the 'modern' or 'formal' sector . . . enjoy high wages, social security, vacation, pension, and employment security as mandated by legislation. By contrast, those unable to find work in such firms resort to the next best alternative, the so-called 'informal sector,' in small firms or self-employment, engaged in labor-intensive activities, without job security or benefits" (Ozorio de Almeida, Alves, and Graham 1995). According to this view, government or union intervention pushes wages in the formal sector above their market-clearing level; therefore, migrants, young workers, and laid-off workers queue up for good jobs and, in the meantime, subsist in the inferior informal sector.

In this article I argue that the traditional conflation of issues of formality and dualism is probably conceptually inappropriate. The labor market for relatively unskilled workers may be well integrated with both formal and informal sectors, offering desirable jobs with distinct characteristics from which workers may choose, with little queuing. Two factors may deter workers from choosing formal salaried work. First, labor protection laws often levy an implicit tax on workers. In a market with downwardly flexible wages the cost of nonwage benefits reduces wages by an equivalent amount. For several reasons, including the inefficiencies and rigidities that often accompany such protections, workers may prefer to evade these labor taxes by operating in the informal sector. Second, the various modalities of informal work may offer other desirable characteristics, such as greater flexibility or possibilities for training, that suit some workers better than those of a formal salaried job. This article examines three such modalities—self-employment, informal salaried employment, and contract work—and explains how they can exist in unsegmented labor markets.

Self-employed workers own informal firms, with or without additional employees. The literature has long recognized that some self-employed workers prosper as dynamic entrepreneurs and coexist with workers rationed out of the formal sector. The large size of the informal sector in developing countries has often seemed pathological and indicative of distortions. It may be, however, that workers in developing countries who choose to start small firms have the same desire for independence or the same entrepreneurial ambition as their counterparts in industrial countries. In fact, they may have greater incentives to try self-employment, and not just to avoid labor taxes. The low level of formal sector productivity that the mass of poorly educated workers in developing countries face may reduce the opportunity costs of being independently employed. Thus the voluntary "upper tier" of self-employment may in fact be the majority of the sector.

Informal salaried workers are employed by informal firms. The literature views informal salaried workers as the least advantaged group in a segmented urban workforce. They receive neither the benefits of self-employment nor the benefits of formal employment. Yet the unique demographics of the sector suggest that informal salaried workers do not represent a reserve army of mature workers who were rationed out of the formal sector. Instead, informal salaried work may serve as the entry point and perhaps training area for young workers.

Contract workers do not receive a regular wage or salary; they are paid by percentage, by piece, on commission, or by fixed contract, and they are often connected to larger firms. Here, contract work corresponds most closely to the overall definition of informality embraced by Portes, Castells, and Benton (1989). In line with the traditional dualistic view they argue that informal workers fall into the category of "degraded labor," receiving fewer benefits, earning lower wages, and enduring worse working conditions than workers in the formal sector. But contract work may serve as a Pareto-improving arrangement between informal workers and firms in the formal sector. The arrangement enables both to evade not the labor legislation per se, but the inefficiencies and rigidities implicit in it.

Section I describes the data employed by this study. It uses detailed and relatively rare panel data to take a more comprehensive and dynamic approach than has been feasible previously and to weigh the plausibility of these alternative views. The Mexican Urban Employment Survey follows individual workers across 15 months. It thus permits us to examine earnings differentials and mobility patterns as workers transition between formal salaried employment and the three modalities of informal work. This panel analysis has several advantages over traditional cross-sectional work.

Comparisons of sectoral earnings differentials have been the mainstay of segmentation tests to date (see Rosenzweig 1988). However, they suffer from two drawbacks: unobserved individual characteristics may be correlated with both sector choice and earnings, and the value of unobserved job characteristics may not be captured. Section II presents the earnings differentials generated by individuals moving between sectors. These differentials largely ameliorate the correlation problem and constitute more precise and reliable estimates than those previously possible. But because of unobserved job-related characteristics, even these differentials cannot prove or disprove segmentation.

As an alternative approach, I examine patterns of worker mobility. Section III generates what is perhaps the first dynamic overview of movement through a labor market in a developing country. It characterizes the interactions among the four classes of work and compares them with the patterns implicitly predicted by the standard dualistic view. This type of analysis has not been possible with conventional cross-sectional data sets that permit only tabulations of the composition of the workforce and, perhaps, changes in those stocks over time. Although unable to provide conclusive evidence, this dynamic approach offers strong reasons to question the dualistic view as the primary explanation for the existence of and changes within the informal sector.

I. DATA

This study employs two sources of data on men ages 16–65 who have a high school education or less and live in 16 major metropolitan areas in Mexico. The National Urban Employment Survey (*Encuesta Nacional de Empleo Urbano*, ENEU) conducts extensive quarterly interviews and is structured so as to generate panels that track one-fifth of the sample across five quarters. Workers are matched by household, role in the household, gender, level of education, and age to ensure against generating spurious transitions. Although five quarters do not permit a full description of the life cycle of an individual, it is nonetheless possible to sketch patterns of mobility between sectors and to identify worker characteristics that correspond to these patterns. To generate a sufficiently large sample of roughly 15,000 observations, three contiguous ENEU cohorts were combined: 1990:3–1991:3, 1991:1–1992:1, and 1991:2–1992:2.[2]

Another panel was created spanning 1991:1–1992:1 that terminates in the more detailed Micro-Enterprises Survey (*Encuesta Nacional de Microempresas*, ENAMIN). This survey was constructed by identifying 11,000 owners of micro-enterprises, defined as firms with fewer than six individuals, from the 1990:4 ENEU and interviewing them again in more depth about capital structures, costs, and employment patterns in 1991:1. It specifically asked why they left their previous job and why they started their present business. These questions offer a partial alternative to naive corrections for selectivity bias. Thus the reduced sample provides far more information on both earnings differentials and motivations for moving. The panels in the combined ENEU sample were chosen to include and span either side of the ENAMIN panel.

Two popular definitions were used to define the informal sector. The first definition focuses specifically on the issue of protection: owners or workers in firms with fewer than 16 employees who do not have social security or medical benefits (see table 1). The second addresses the issue of the role of the small firm or microenterprise, using the Mexican government's definition: firms with fewer than six workers (see table 2). Because only the latter definition is consistent with the sampling of the ENAMIN, it is the one used for the joint ENEU/ENAMIN panel.

Tables 1 and 2 give the mean values for the sample using the unprotected workers and microenterprise definitions, respectively, for age, work experience, level of schooling, and an index of the real wage in the first period for workers who stay in or change sectors. Although the differences in sample means are sometimes statistically significant, they are never large, confirming that there is great overlap in the composition of the two populations. This is largely because among entrepreneurs in firms with fewer than six workers, only 5 percent are

2. The cohort beginning in the fourth quarter of 1990 was not used because it would incorporate the end-of-year bonuses, the *aguinaldo*. This normally amounts to a thirteenth monthly payment but may vary by year and by firm, imparting an undetermined upward bias to the monthly wage reports. Because we are concerned with expected income differentials, the wages of workers who reported normally receiving the *aguinaldo* are multiplied by 13/12.

Table 1. *Summary Statistics for the Sample Using the Unprotected Workers' Definition of the Informal Sector, Mexico, 1991–92*

Workers' transitions between sectors	Number of observations	Mean: Age (years)	Experience (years)	Schooling (years)	Initial real wage[a] (index)
To the formal salaried sector from					
Formal salaried	4,168	34.1	20.7	7.36	100.0
		(0.19)	(0.21)	(0.04)	(0.10)
Self-employed	264	37.0	24.6	6.39	124.2
		(0.75)	(0.83)	(0.60)	(0.61)
Informal salaried	315	28.7	16.0	6.66	70.3
		(0.70)	(0.79)	(0.16)	(0.23)
Contract	192	34.2	21.9	6.28	148.2
		(0.80)	(0.90)	(0.21)	(1.46)
To the self-employed sector from					
Formal salaried	270	36.2	23.7	6.56	106.8
		(0.68)	(0.78)	(0.18)	(0.52)
Self-employed	1,793	43.1	31.6	5.52	136.4
		(0.28)	(0.32)	(0.08)	(0.59)
Informal salaried	212	34.7	23.3	5.41	74.6
		(0.89)	(1.01)	(0.21)	(0.28)
Contract	125	39.4	27.5	5.91	121.4
		(1.19)	(1.35)	(0.28)	(0.86)
To the informal salaried sector from					
Formal salaried	308	29.6	17.0	6.62	87.2
		(0.67)	(0.77)	(0.16)	(0.26)
Self-employed	242	37.4	26.1	5.31	94.2
		(0.86)	(0.96)	(0.20)	(0.37)
Informal salaried	529	29.0	16.9	6.11	69.8
		(0.54)	(0.61)	(0.13)	(0.19)
Contract	116	34.3	22.7	5.59	101.8
		(1.23)	(1.38)	(0.28)	(0.92)
To the contract sector from					
Formal salaried	212	31.5	18.7	6.87	111.3
		(0.67)	(0.76)	(0.19)	(0.50)
Self-employed	149	37.3	25.2	6.09	134.6
		(0.96)	(1.08)	(0.24)	(0.74)
Informal salaried	93	29.4	17.2	6.15	104.5
		(1.19)	(1.31)	(0.32)	(1.78)
Contract	439	35.2	23.0	6.18	132.7
		(0.59)	(0.66)	(0.15)	(0.45)

Note: Transitions correspond to initial and final positions across a five-quarter period. The informal sector consists of workers in unprotected firms with fewer than 16 total workers who are not covered by medical or social security benefits. Self-employed workers are owners of informal firms. Salaried informal workers are employed by informal firms. Contract workers are unsalaried workers who do piecework or other contact work. Standard errors are in parentheses.

a. The formal sector wage is the base.

Source: Author's calculations based on the Mexican National Urban Employment Survey.

Table 2. *Summary Statistics for the Sample Using the Microenterprise Definition of the Informal Sector, Mexico, 1991–92*

Workers' transitions between sectors	*Number of observations*	*Mean*			
		Age (years)	*Experience (years)*	*Schooling (years)*	*Initial real wage*[a] *(index)*
To the formal salaried sector from					
Formal salaried	4,421	33.8	20.4	7.35	101.2
		(0.18)	(0.20)	(0.04)	(0.10)
Self-employed	301	36.3	23.8	6.47	121.7
		(0.66)	(0.73)	(0.17)	(0.57)
Informal salaried	388	29.9	17.2	6.65	77.8
		(0.63)	(0.72)	(0.15)	(0.24)
Contract	241	34.3	22.1	6.17	146.4
		(0.77)	(0.86)	(0.19)	(1.29)
To the self-employed sector from					
Formal salaried	320	36.4	24.1	6.27	106.4
		(0.64)	(0.72)	(0.17)	(0.50)
Self-employed	1,812	43.1	31.6	5.50	131.0
		(0.28)	(0.32)	(0.07)	(0.56)
Informal salaried	209	35.0	23.2	5.75	78.5
		(0.87)	(0.99)	(0.21)	(0.31)
Contract	133	39.1	27.2	5.93	121.0
		(1.15)	(1.30)	(0.27)	(0.81)
To the informal salaried sector from					
Formal salaried	391	29.2	17.3	6.56	85.8
		(0.62)	(0.70)	(0.15)	(0.27)
Self-employed	244	37.7	26.5	5.23	97.6
		(0.88)	(0.99)	(0.20)	(0.46)
Informal salaried	581	30.4	18.1	6.34	68.7
		(0.51)	(0.58)	(0.12)	(0.16)
Contract	94	34.4	22.9	5.51	102.9
		(1.22)	(1.43)	(0.32)	(1.10)
To the contract sector from					
Formal salaried	245	31.7	19.0	6.64	110.2
		(0.65)	(0.72)	(0.18)	(0.49)
Self-employed	146	37.0	24.9	6.15	133.6
		(0.97)	(1.08)	(0.24)	(0.73)
Informal salaried	88	29.2	16.7	6.53	105.7
		(1.12)	(1.30)	(0.34)	(1.84)
Contract	439	35.3	23.0	6.23	131.1
		(0.59)	(0.66)	(0.15)	(0.45)

Note: Transitions correspond to initial and final positions across a five-quarter period. The informal sector consists of workers in microenterprise firms with fewer than six workers. Self-employed workers are owners of informal firms. Salaried informal workers are employed by informal firms. Contract workers are unsalaried workers who do piecework or other contact work. Standard errors are in parentheses.

a. The formal sector wage is the base.

Source: Author's calculations based on the Mexican National Urban Employment Survey.

covered by benefits, and the vast majority of microenterprises have fewer than three workers. Therefore, the results of the analysis are unlikely to be driven by the particular definition chosen. To conserve space, this article presents results only for the microenterprise definition of the informal sector. For complete results for both definitions, see Maloney (1998a).

II. Wage Differentials as Evidence of Segmentation

Traditional efforts to identify segmentation by comparing conditional means between sectors are unsatisfying for two reasons. First, as is well documented in the literature, unobserved worker characteristics that affect productivity—ability to tolerate authority, punctuality, entrepreneurial ability—may also influence an individual's choice of sector in which to work and bias estimates of the sector differential.

Second, and of greater concern, the specific characteristics of work that pertain to or even define the formal and informal sectors affect the earnings paid in each sector. These characteristics make it unclear what the magnitude or sign of the differential should be, even in an unsegmented market. In a market without distortions wages should move to equalize the utility obtained from working in each sector. Most generally, informal sector earnings should exceed formal sector wages by the expected value of benefits not received, but should fall below them by the amount of taxation that informal workers often evade. More specifically, earnings for both contract and self-employed workers may reflect a premium for risk and a more independent lifestyle. In the case of self-employed workers the premium includes the implicit cost of capital invested and the value of unpaid work by family members. The ENAMIN suggests that unpaid family members make up 34 percent of microenterprise employees. Informal salaried workers are among the youngest workers, and the ENAMIN reveals that roughly 30 percent of informal salaried workers are related to their employer (tables 1 and 2). Their reported earnings may therefore incorporate training costs or unobserved payments in kind, which could imply that they are paid less than they would be in a formal firm, even in the absence of segmentation (see Hemmer and Mannel 1989 and Roberts 1989). The problems that these unobservables pose become apparent in interpreting the earnings differentials.

Estimates of Earnings Differentials

Using the vast but undetailed ENEU, the first columns of table 3 show the percentage change in hourly real earnings generated by individuals moving between sectors. These estimates hold worker characteristics constant and leave the variations in the characteristics of the work itself as the residual explanatory factor. The next columns account for these characteristics more than has been previously possible and present changes in the real wage net of taxes and for the real wage net of taxes per hour. Tax payments are calculated based on the Mexican tax tables, assuming that all informal workers avoid paying taxes and that all formal workers pay taxes.

The non-normality of the distributions makes the sample mean an inadequate measure of the central tendency of the data and its significance. Table 3 presents two alternative measures: the median and the mean, as determined by a robust estimation technique (the robust mean). The median is largely unaffected by the non-normality of the distribution and is robust to outliers. The robust mean

Table 3. *Real Wage Differentials from Transitions between Sectors, Mexico, 1991–92*

		Differential (percent)								
		Real wage			Real wage net of taxes		Net wage per hours worked		Hours worked	
Transition	Number of observations	Mean[a]	Robust mean[b]	Median[c]	Robust mean[b]	Median[c]	Robust mean[b]	Median[c]	Robust mean[b]	Median[c]
To the self-employed sector from										
Formal salaried	320	17.66	8.69	3.96	16.63	11.82	27.61	24.51	–3.55	0.00
			(2.41)	(0.74)	(4.36)	(2.36)	(6.48)	(5.95)	(2.27)	(0.00)
Self-employed	1,812	–5.35	–1.31	–4.21	–1.31	–4.21	0.98	–4.09	–1.63	0.00
			(0.91)	(1.60)	(0.91)	(1.60)	(0.60)	(2.04)	(2.26)	(0.00)
Informal salaried	209	27.05	18.87	16.04	18.87	16.04	23.30	22.28	–3.37	–1.64
			(3.92)	(2.93)	(3.92)	(2.93)	(4.66)	(3.88)	(1.78)	(0.57)
Contract	133	17.00	6.95	8.43	6.95	8.43	14.29	11.92	–2.91	0.00
			(1.45)	(1.18)	(1.45)	(1.18)	(2.44)	(1.94)	1.13	(0.00)
To the informal salaried sector from										
Formal salaried	391	–3.92	–1.91	–2.89	5.06	3.87	6.14	1.68	0.99	0.00
			(0.89)	(1.03)	(2.22)	(1.42)	(2.57)	(0.60)	(1.36)	(0.00)
Self-employed	244	–26.41	–20.53	–22.31	–20.53	–22.31	–20.76	–22.31	–0.89	0.00
			(7.41)	(5.31)	(7.41)	(5.31)	(7.05)	(5.01)	(0.56)	(0.00)
Informal salaried	581	–0.13	2.65	2.72	2.65	2.72	1.87	–0.36	0.83	0.00
			(1.64)	(2.01)	(1.64)	(2.01)	(1.15)	(0.16)	(1.27)	(0.00)
Contract	94	–19.69	–12.68	–13.68	–12.68	–13.68	–6.29	–4.95	–1.82	0.00
			(3.37)	(2.76)	(3.37)	(2.76)	(1.24)	(0.88)	(0.88)	(0.00)

To the formal salaried sector from										
Formal salaried	4,421	6.84	5.28	4.02	7.26	6.35	7.53	6.42	0.42	0.00
			(8.95)	(8.32)	(12.44)	(8.64)	(11.31)	(8.60)	(2.54)	(0.00)
Self-employed	301	−13.47	−0.71	−0.90	−6.15	−7.84	−13.59	−13.02	3.77	0.00
			(0.21)	(0.16)	(1.94)	(1.55)	(4.17)	(2.83)	(1.90)	(0.00)
Informal salaried	388	19.25	14.40	11.28	8.78	5.16	8.71	6.75	0.12	0.00
			(5.74)	(3.11)	(3.70)	(1.70)	(3.52)	(1.71)	(0.19)	(0.00)
Contract	241	−18.66	−7.13	−13.68	−12.87	−17.62	−14.07	−14.19	0.07	0.00
			(2.19)	(3.48)	(4.31)	(4.60)	(4.78)	(4.56)	(0.06)	(0.00)
To the contract sector from										
Formal salaried	245	12.89	18.94	15.92	27.84	24.70	22.60	14.90	3.82	1.64
			(4.78)	(4.01)	(6.69)	(6.26)	(5.33)	(3.28)	(3.73)	(1.05)
Self-employed	146	−5.21	−3.63	−7.28	−3.63	−7.28	−7.57	−13.68	2.72	0.00
			(0.77)	(1.56)	(0.77)	(1.56)	(1.64)	(2.02)	1.00	(0.00)
Informal salaried	88	−3.86	15.08	7.90	15.08	7.90	19.61	15.15	1.88	3.28
			(2.46)	(0.87)	(2.46)	(0.87)	(2.81)	(1.35)	(0.95)	(1.59)
Contract	439	3.09	2.50	−0.13	2.50	−0.13	1.29	−0.20	0.17	0.00
			(1.07)	(0.04)	(1.07)	(0.04)	(0.54)	(0.08)	(0.16)	(0.00)

Note: Transitions correspond to initial and final positions across a five-quarter period. The informal salaried sector consists of workers in microenterprise firms with fewer than six workers. See table 2 for summary statistics and definitions of worker types. *t*-statistics are in parentheses.

a. Mean average differential weighted by initial real wage to give differential between sample.

b. The robust mean is calculated using Huber weights to redress non-normality.

c. The median is from a quantile regression using bootstrapped standard errors.

Source: Author's calculations based on the Mexican National Urban Employment Survey.

attempts to recover the information in the tails, while compensating to some extent for outliers and non-normality. All calculations are done in STATA.

Ideally, we might correct for selectivity bias (see, for example, Marcouiller, Ruiz de Castilla, and Woodruff 1997). However, Heckman (1979) argues that the standard two-stage methods depend on having confidence in the underlying model of how workers choose among sectors. A poor first-stage selection specification may actually induce an unknown bias rather than improve the ordinary least squares estimates. This is likely to be the case, since the premise of this work is that we have little knowledge of the role each sector serves. To check for the robustness of the results, the Heckman procedure was run for all transitions, but in most cases it had little effect on the results (see Maloney 1998a).

For all three transitions into self-employment the ENEU/ENAMIN panel allows further adjustment for the imputed return on the value of capital (tools, inventories, and location, if owned) and for hours worked by unpaid workers (table 4).

The results in table 3 appear to invert the conventional view of the relationship between the formal and informal sectors. Movement from self-employment or contract work into formal salaried employment is associated with a significant decline in remuneration, and movement from formal salaried employment to self-employment or contract work leads to a significant increase. By contrast, movement from informal salaried to formal salaried work increases remuneration, and movement from formal salaried to informal salaried work leads to no significant change.

But here we immediately fall afoul of our inability to capture the value of unmeasured characteristics of different kinds of work, such as benefits, compensation for risk, independence, in-kind payments, or implicit training costs. The magnitude of the distortion-free differential cannot be known ex ante, and hence there is no benchmark against which to compare segmentation. For example, workers might fully value benefits in the formal sector, which are roughly 40 percent of the wage, and other unobservables might be negligible (see Dávila-Capalleja 1994). In this case the differential in total remuneration for movement out of formal employment would be negative for all transitions; for the reverse movement it would be positive, as would be the case in a segmented market. Unfortunately, the analysis cannot assume that workers fully value benefits or that the other unobservables are negligible. For example, the value of not having a boss is difficult to measure. In sum, neither these nor any previously reported sectoral differentials are reliable measures of segmentation.[3]

Presumably, the differentials between informal sectors are less affected by unobservables, such as the loss of formal sector benefits. Table 3 shows that movement into self-employment from every other sector is always associated

3. Table 4 suggests that those who move voluntarily into self-employment do far better than those who move involuntarily. However, the asymmetries in the differentials in table 3 do not necessarily indicate that there is a larger component of voluntary movement into the formal salaried sector. Involuntary separations due to the closure of formal sector firms imply the loss of well-paying jobs and a large differential. The analogous failure of a microenterprise may imply low earnings prior to the transition and a smaller differential.

Table 4. *Real Wage Differentials for Workers Moving into Self-Employment in the* ENEU/ENAMIN *Panel, Mexico, 1991–92*

		Real wage		*Real wage net of taxes*		*Real wage net of capital costs*[a]		*Adjusted for worker hours*[b]		*Hours*[c]	
Indicator	*Number of observations*	*Robust mean*[d]	*Median*[e]	*Robust mean*[d]	*Median*[e]	*Robust mean*[d]	*Median*[e]	*Robust mean*[d]	*Median*[e]	*Robust mean*[d]	*Median*[e]
Sector of origin											
Formal salaried	139	33.17	25.60	30.26	25.60	25.36	22.09	18.30	10.63	5.52	8.33
		(4.54)	(3.59)	(4.27)	(4.15)	(3.57)	(2.70)	(2.58)	(2.26)	(1.43)	(2.78)
Informal salaried	125	9.83	–0.77	7.76	–0.77	–0.77	–10.45	–3.74	–7.63	–0.04	2.27
		(1.57)	(0.08)	(1.27)	(0.06)	(0.13)	(1.13)	(0.60)	(0.70)	(0.00)	(0.66)
Contract	192	10.97	5.35	17.21	11.36	14.85	10.73	12.63	10.01	–1.91	2.00
		(2.20)	(1.34)	(3.31)	(1.51)	(2.86)	(2.15)	(2.16)	(1.96)	(0.52)	(0.88)
Reason for leaving previous formal salaried job (single response)											
More independence	67	20.16	12.88	27.14	17.80	24.57	16.05	21.59	19.64	–4.61	0.00
		(2.22)	(1.60)	(2.90)	(7.14)	(2.66)	(1.65)	(2.16)	(1.66)	(0.88)	(0.00)
Higher pay	62	19.25	13.48	25.54	18.94	24.50	18.92	16.74	12.09	6.15	12.00
		(2.19)	(0.79)	(2.82)	(1.34)	(2.73)	(0.91)	(1.57)	(1.25)	(0.96)	(1.82)
Involuntary	55	–8.76	–5.42	–2.64	0.07	–5.08	–3.83	–2.59	–14.66	–4.09	9.09
		(1.13)	(0.44)	(0.32)	(0.00)	(0.60)	(0.34)	(0.24)	(0.95)	(0.51)	(1.17)
Reason for starting microenterprise (multiple response)											
Independence	120	8.94	8.66	14.61	14.66	14.04	13.89	3.61	12.85	–3.57	2.00
		(1.40)	(3.27)	(2.19)	(3.00)	(2.13)	(1.90)	(0.51)	(1.36)	(0.75)	(0.38)
Higher pay	63	22.45	17.68	30.94	28.11	28.17	27.80	21.25	18.82	3.53	5.00
		(3.06)	(1.65)	(4.10)	(2.88)	(3.79)	(3.70)	(2.82)	(2.53)	(0.66)	(1.53)
Fired or unable to find other work	38	–11.93	–8.83	–6.88	–3.73	–7.72	–7.35	–18.15	–19.94	9.44	11.11
		(1.51)	(0.94)	(0.83)	(0.28)	(0.96)	(0.72)	(1.78)	(1.82)	(1.32)	(1.59)
Tradition	12	32.45	22.50	28.32	24.32	22.21	24.14	9.48	–8.64	18.27	24.07
		(1.40)	(0.56)	(1.42)	(0.62)	(0.82)	(0.49)	(0.28)	(0.22)	(2.11)	(1.82)

Note: Transitions correspond to initial and final positions across a five-quarter period. The informal salaried sector consists of workers in microenterprise firms with fewer than six workers. See table 2 for summary statistics and definitions of worker type. *t*-statistics are in parentheses.

a. The tax-adjusted differential net of capital costs imputed at 10 percent of the sum of the value of tools, inventories, and location if owned.

b. The capital cost and tax-adjusted differential adjusted for total hours worked by all workers including unpaid workers.

c. The differential in hours worked by the principal worker.

d. The robust mean using Huber weights to redress non-normality.

e. The median from a quantile regression using bootstrapped standard errors.

Source: Author's calculations based on the joint ENEU/ENAMIN sample.

with a substantial and significant rise in per-hour after-tax remuneration. The increases reach more than 15 percent for movement from the formal and informal salaried sectors.[4]

In table 3 movement from formal salaried and informal salaried work into self-employment is associated with a substantial and significant increase in wage. However, in moving from contract work to self-employment, the differential is insignificant. Thus contract work appears similar to self-employment and may therefore share the same composition of voluntary and involuntary entrants.

For movements into and out of formal salaried work, salaried informal employment suffers a discount relative to self-employment and contract work. After adjusting the differentials for standard firm size effects observed in the United States, moving to an informal sector firm from a comparably sized formal sector firm yields a 12–15 percent increase in remuneration. The large asymmetries suggest that informal salaried workers may gain far more by entering formal salaried work than they lose by leaving it, as would be predicted if workers were queuing for formal salaried jobs.

But, again, just as the premium earned from entering self-employment or contract work does not imply that these sectors are superior to the others or that workers queue to enter them, the discount in the informal salaried sector cannot be seen as evidence of the generally assumed inferiority of this type of work. The analysis suggests that the composition of the premium in the absence of distortions is distinct for the informal salaried sector. This sector may have a lower premium for risk because the workers are not entrepreneurs, the value of in-kind benefits for the 30 percent of informal salaried workers who are related to the owner may be large, or the implicit training costs may be substantial. Any of these factors could easily account for a 15 percent differential. Further, if deductions for training costs were a substantial fraction of the wage, the asymmetry seen entering and leaving the formal sector would be expected because returning workers might not work as apprentices and hence would not pay for training. The negative premium for informal salaried work reveals little about the relative desirability of the sector.

Are the Differentials Consistent with an Integrated Labor Market?

Although the analysis cannot credibly prove or disprove segmentation based on these differentials, it can ask whether the differentials seem consistent with an

4. Table 4 shows the joint ENEU/ENAMIN panel with modest adjustments for capital costs and much larger ones for unpaid labor. For capital costs a return of 10 percent is imputed. Most microenterprises that save in commercial accounts, or *cajas de ahorro*, receive a real rate of interest of 3 percent. This analysis assumes a rate of 7 percent for depreciation. The low level of capital employed results in the overall differential being relatively insensitive to the cost of capital chosen.

The more detailed treatment of taxation in the ENAMIN allows us to drop the previous assumption of complete avoidance by self-employed workers and slightly moderates the after-tax differential. The adjustment for unpaid labor may be overstated if the unpaid labor is in training or if Balán, Browning, and Jelin (1973: 218) are correct that, "Since in most cases these family members would not otherwise be employed outside the household, their contribution to family finances is a 'net' one."

absence of segmentation. If the responses of those reporting voluntary moves into self-employment are reliable, does the 15–20 percent differential seem large enough to cover the value of benefits, some return to risk, and whatever value is placed on one style of work compared with another? The breakdown by motivation in table 4 could, in theory, offer some measure of the value of different styles of work. But the differential for "independence" does not vary significantly from that for "higher pay"; depending on the question, it implies contradictory signs on the premium. This ambiguity could mean that independence has a small effect. If the risk premium is positive, the value of benefits may be even smaller—half or even a quarter of those on paper.

Three factors make this plausible and suggest why the attraction of formal sector employment may be overstated. First, because the medical benefits program covers a worker's entire family, the marginal value of benefits to the second formal sector worker in a family is zero. This would seem particularly important for informally employed workers in households whose principal breadwinner may be formally employed: there is no reason to pay again the implicit tax for benefits already received. Second, administrative overhead costs are high, and the benefits may have a low value given their cost. In his interviews with workers in Guadalajara, Roberts (1989: 50) finds that, "Many informants cited the deduction made for welfare as a disadvantage of formal employment, particularly since the services they received were poor." Third, rapid rates of turnover mean that leaving does not necessarily imply the loss of nominally generous separation benefits and pensions, since Balán, Browning, and Jelin (1973: 212) find in their extensive surveys of worker career trajectories in Monterrey that, "Many change enterprises quite often, and thus they cannot benefit from the seniority accumulated in each of them." In each case the value of formal sector benefits to workers is below their value on paper and, in a market with reasonably flexible wages, below the taxes workers implicitly pay (see Bell 1997).

In sum, earnings differentials do not offer compelling evidence in favor of the segmentation hypothesis, and given the difficulty of quantifying the unobservables, earnings differentials are unlikely ever to be convincing tests.

III. Patterns of Mobility

By contrast, the patterns of worker transitions can offer additional information on the validity of the dualistic view. Ideally, a model of the behavior of each of the four sectors and workers' choices to enter them would be postulated and held up against the evidence. However, this is a vast research agenda in itself. This article seeks to develop only a few provocative stylized facts about the dynamics of the market and to characterize the natures of and the interactions among the sectors.

The dualistic view predicts that some general patterns should emerge. If formal sector work is preferred to informal work, then workers will queue up for formal sector jobs and relinquish them only under the limited conditions permit-

ted by the constitution—egregious conduct or "acts of god" that induce firms to downsize. (Mexico's constitution conceives of the employment relationship as a lifetime contract, and workers may only be fired under extreme circumstances and at great cost.) This situation should imply the following:

- Very low rates of formal sector turnover.
- A largely unidirectional flow of workers who graduate from the informal sector to the formal sector, where they stay until retirement. Flows in the other direction should be largely involuntary and, in relatively prosperous times, far less. When this survey was conducted, the Mexican economy was growing, and the unemployment rate was at its lowest since 1989, at around 2.6 percent.
- Given a constant probability of being selected from the queue in each time period, the probability of entry into formal salaried work is an increasing function of experience.

This section tests for the presence of these patterns in two ways. First, table 5 provides summary data on transitions between sectors by tabulating the conditional probability of finding a worker in sector j at the end of the period given that the worker began in sector i, P_{ij}.[5]

The row percentages in panel 1 of table 5 sum to 100 percent, and the totals at the bottom represent the share of workers in each category at the end of the period, $P_{.j}$. The first three columns and rows represent individuals who are not working: those out of the labor force, not currently working, and not searching for work; those studying; and those looking for work (the unemployed). The bold rectangle encloses five categories of work, beginning with unpaid labor, and the shaded area includes all paid jobs. Panel 1 describes raw tendencies among sectors, including the percentage of any given group that will move to another sector by the end of a 15-month period. The diagonals reflect the share of workers who do not move. From this number the mean time spent in the sector can be calculated as $1.25/(1 - P_{ii})$. If 50 percent of workers leave the sector in 15 months, then the mean time spent in the sector is 2.5 years.

Panel 1 cannot provide any measure of whether flows into a particular sector are especially high or low. In a random shuffling of workers, P_{ij} would clearly increase with $P_{.j}$, the size of the terminal sector; therefore, panel 2 of table 5 standardizes the transition probabilities by size of the terminal sector, $P_{ij}/P_{.j}$. From any given sector, reading along the row gives a feel for whether flows into any j are high compared with a purely chance distribution. Looking for symmetry in the flows between any two sectors—for example, movement from school to formal work compared with movement from formal work to school—may suggest whether flows tend to be unidirectional or bidirectional.

5. In a spirit similar to this work, Sedlacek, Paes de Barros, and Varandas (1990) study the mobility of Brazilian workers with and without signed work cards, and hence worker protections. They find little evidence of strong barriers to mobility.

Table 5. *Worker Transitions among Sectors of the Labor Market across Five Quarters, Mexico, 1991–92*

	Terminal sector								
Initial sector	*Out of labor force*	*In school*	*Unemployed*	*Unpaid*	*Self-employed*	*Informal salaried*	*Formal salaried*	*Contract*	*Other*
Panel 1. Probability of moving from initial to terminal sector, P_{ij} (percent)									
Out of labor force	73	1	6	1	10	3	5	1	0
In school	2	49	3	6	3	9	25	3	0
Unemployed	20	4	22	3	11	11	24	5	0
Unpaid	3	11	5	27	16	18	13	6	0
Self-employed	2	0	2	1	69	9	10	5	1
Informal salaried	2	2	1	3	16	41	29	7	0
Formal salaried	1	1	1	0	6	7	78	5	0
Contract	1	0	1	0	15	10	26	45	1
Other	0	0	2	0	26	6	34	6	27
Total ($P.j$)	5	4	2	2	22	12	45	8	1
Panel 2. Probability standardized by size of the terminal sector $\frac{P_{ij}}{P_{.j}}$ (percent)									
Out of labor force		24	321	45	43	23	12	13	15
In school	31		155	389	11	77	55	41	13
Unemployed	385	113		161	51	91	53	60	0
Unpaid	61	306	259		72	159	30	77	0
Self-employed	41	9	91	71		77	23	70	128
Informal salaried	31	54	60	180	71		64	83	38
Formal salaried	26	26	38	30	27	63		58	50
Contract	25	12	36	17	68	86	59		72
Other	9	0	97	0	115	49	75	73	

(Table continues on following page.)

Table 5. *(continued)*

Initial sector	Terminal sector: Out of labor force	In school	Unemployed	Unpaid	Self-employed	Informal salaried	Formal salaried	Contract	Other
Panel 3. Disposition to move to a sector, V_{ij}									
Out of labor force		180	1,551	229	517	145	201	89	79
In school	230		394	1,050	71	257	498	146	36
Unemployed	1,856	287		283	209	199	313	140	0
Unpaid	311	825	454		314	369	185	190	0
Self-employed	487	57	374	309		415	335	406	556
Informal salaried	199	182	130	418	383		491	257	88
Formal salaried	449	234	224	187	397	483		483	312
Contract	172	45	85	41	393	266	488		180
Other	45	0	171	0	500	114	468	182	

Note: The sample aggregates three panels—1990:3–91:3, 1991:1–1992:1, 1991:2–92:2—generating roughly 15,000 observations. The sample includes male workers, ages 16–65, with a high school education or less, in 16 metropolitan areas. Informal workers are workers in firms with fewer than six employees.

Source: Author's calculations based on the Mexico National Urban Employment Survey, 1991–92.

This standardized index is still an imperfect measure of fluidity between sectors. Although transitions between any pair of initial and terminal sectors can be compared, it is not possible to compare more generally across the table because there are differences in separation rates from the initial sector and differences in the likelihood that, given the size of the terminal sector, a position will actually open up. A weak desire to leave the initial sector will yield a low value in panel 2, as will a distortion-induced low level of turnover in the terminal sector:

$$P_{ij} / P_{.j} = (1 - P_{ii}) V_{ij} (1 - P_{jj}). \tag{1}$$

V_{ij}, tabulated in panel 3 of table 5, captures the disposition or economic or institutional logic that compels a worker leaving the initial sector to enter an open position in j. For example, although both third- and fourth-grade elementary school classes may fully turn over every year, V would be large for transitions in the ascending direction and zero in the reverse. In the present case the disposition to enter paid employment from school is two to three times that of the reverse transition, as would be expected if workers generally graduate from school to employment. If the dualism hypothesis that workers graduate from informal to formal employment is correct, similar asymmetric Vs would occur between the sectors.

Finally, the appendix presents the multinomial logit analysis of the determinants of transitions between sectors. Table A-1 reports the results of four sets of regressions that correspond to the four initial sectors in the first period of the panel. Although no particular theoretical model of transition is offered, the logits offer a more statistically rigorous way of asking if, given the initial sector, a worker is more or less likely to move to another sector if he has more experience, has more education, or has lost employment. In this way we can crudely trace out possible patterns of movement, including those across the life cycle. Together, the analysis in table 5 and the multinomial logit analysis offer a view of overall labor force dynamics and how the four sectors interact.

Overview of the Labor Market

Three important general findings are immediately apparent. First, table 5 reveals high levels of mobility, with turnover rates (and implicitly the length of tenure at 5.2–5.7 years) in the formal sector similar to those in the United States (5.1 years).[6] In fact, since we cannot observe transitions within a sector, the turnover rate in Mexico is far higher.

Second, the symmetry of Vs across directions of movement in all sectors of paid work seems more consistent with a well-integrated market in which work-

6. The median tenure for all U.S. workers more than 16 years old in 1991 was 5.1 years (U.S. Department of Labor 1992).The implicit tenure based on the July/August 1994 median separation rate of 1.1 percent a month was 7.6 years (U.S. Bureau of National Affairs 1994). The mean tenure is calculated as (span of panel) / $(1 - P_{ii})$. Krebs and Maloney (1999) and Maloney (1999) discuss how turnover rates might vary across the development process and hence not be strictly comparable. There appears to be little evidence of the rigidities that would be expected, given the incentives in the Mexican labor code.

ers search for job opportunities across sectors than with one in which informal workers seek permanent status in the formal sector and stay until they retire. Third, and most striking, the logit results show that in no sector does the probability of moving into the formal sector relative to staying raise overall experience, as would be expected if workers were queuing to enter the sector. This result suggests relatively easy access to formal employment and is consistent with earlier findings on Mexican migrants. Gregory (1986: 267) argues that, "The empirical evidence . . . represents the antithesis of the Todaro [dualist] model. Rather than flowing into a queue to await the opening of improved employment opportunities, migrants moved quickly and easily into employment opportunities in both the formal and informal sectors." Overall, the data suggest an urban labor market that is at once very fluid and integrated.

Finally, every movement out of formal salaried work decreases with education. This is consistent with unskilled workers facing a lower opportunity cost of becoming informal.

Self-Employment as an Alternative to Employment in the Formal Sector

Self-employment constitutes the largest source of employment (25 percent) after formal salaried employment (50 percent). Although it may play the traditionally postulated role of a holding pattern or safety net, the data are consistent with self-employment being a desirable sector in itself. As a first approximation, it may be more correct to assume that small-scale firms in developing countries have origins and dynamics similar to their counterparts in industrial countries, rather than being a distinct phenomenon (see Levenson and Maloney 1998).

The motivational responses from the joint ENEU/ENAMIN panel discussed in table 4 show that at least two-thirds of those entering self-employment from formal salaried employment report moving voluntarily, citing a desire for greater independence or higher pay as the principal motive. This percentage remains relatively unchanged even when the sample is restricted to those previously working in firms with more than 50 employees. The results support Balán, Browning, and Jelin's (1973) finding that being one's own boss is well regarded and that movements into self-employment from salaried positions often represent an improvement in job status. Of the moves from formal positions into self-employment they studied, 57 percent were upward moves in job quality, 30 percent were horizontal (which the authors argue is welfare improving because of the greater independence), and only 11 percent were downward.

These results are also very close to those of Gottschalk and Maloney (1985), who find that roughly 70 percent of U.S. job changes are voluntary. Put differently, if self-employment and contract work, given their common earnings differentials, are close substitutes for formal salaried work, the implied rates of involuntary entry would be normal by the standards of a flexible labor market in an industrial country.

The transitional evidence corroborates the motivational reports. Turnover rates in the self-employment sector (and implicit tenure at 3.7 years) are far closer to

those in the formal sector than to those in the other two informal sectors. As in industrial countries, self-employment is not an entry occupation from school (Aronson 1991), and there is little evidence that the self-employment sector serves as a holding pattern for young workers. The *V* values from school are only one-fifth, and from unemployment about one-half, of those entering formal salaried employment. Transitions into self-employment from the other paid sectors occur four to six years later than transitions into the other alternatives, including formal salaried work (tables 1 and 2). The mean age of workers in the self-employment sector is eight years higher than in the next closest sector. From every paid sector the logit results reveal that the probability of moving into self-employment is associated with greater experience (see table A-1).[7]

These patterns support the recent literature on liquidity constraints in industrial countries, which dictates that there is a threshold level of financial and human capital necessary to start a business. This capital can be accumulated only with time and work as a salaried employee (see Evans and Jovanovic 1989; Aronson 1991; and Aroca and Maloney 1998). This situation is exacerbated in developing countries. Balán, Browning, and Jelin (1973: 217) argue that, "First, the man must accumulate capital. This is no easy matter when he has a manual job and must provide for a large family, so it generally takes years to accumulate enough capital. There must be sufficient funds not only to set up the business, but also to keep it going during the months or years while it runs at a deficit . . . these kinds of capital requirements are modest enough, but the capital is not easy to come by for the working classes of Monterrey or elsewhere in Mexico."

As with Evans and Leighton (1989), Balán, Browning, and Jelin find that the percentage of workers who enter self-employment is roughly constant across age cohorts. Maloney (1998b) finds, again counter to the standard dualistic view, that the share of the working population in self-employment behaves procyclically, suggesting that entrepreneurs wait until better times before opening their businesses. The fact that, regardless of destination, the less experienced are more likely to leave self-employment is in line with the mainstream literature on firm dynamics, which shows that younger firms (and, on average, less experienced workers) have higher failure rates (see Jovanovic 1982 and Evans and Leighton 1989).

But what would compel workers to give up the ostensibly large benefits of the formal sector? First, it may be that the decision process of the self-employed is not fundamentally different from that of their counterparts in industrial countries, who also take on responsibility for medical insurance or saving for retire-

7. The negative quadratic term on experience in table A-1 (with the exception of the contract sector for which it was never significant) is consistent with the findings of Brock, Evans, and Phillips (1986) for the United States and makes transitions into self-employment negative in experience after 28 years. This suggests that the sector is not primarily a haven for older workers who lose their jobs in the formal sector. The likelihood ratios on the transition between the formal sector and self-employment indicate that those who involuntarily left their previous job are more likely to end up self-employed than to stay in the formal sector. The reverse dynamic appears to be just as strong: the significant likelihood and *Z*-statistics on the involuntary interactive terms on the transition between self-employment and formal sector work suggest that the formal sector acts as a safety net for failed entrepreneurs.

ment that was previously covered by their employers. Second, because the cost of benefits to employers reduces the wage component of remuneration in the formal sector, a perceived value below that cost, as suggested in section II, would lead workers to seek out jobs in the unregulated sector where remuneration is entirely monetary. Third, Balán, Browning, and Jelin's (1973) interviews suggest that the very legislation that is thought to induce rigidities into the labor market in fact stimulates such turnover and encourages workers to leave salaried employment. The paucity of openings for promotion on the rigid *escalafón* (career ladder) as well as the ceiling on mobility opportunities for manual workers make self-employment the remaining outlet for further advancement. These last two issues suggest that, in contrast to the usual view, extant labor protections may make formal sector work less desirable rather than less attainable.

This logic, which applies to all three informal sectors, is most compelling where small-scale firms can offer remuneration that is comparable to that earned in the formal sector—among workers with little education who are unlikely to generate much firm-specific capital. The logit results show that workers become less likely to leave formal employment for self-employment, or any other informal sector work, as their level of education increases. The available macro-level evidence suggests that as the opportunity cost of being one's own boss rises with labor productivity in the formal sector, the share of the labor force in self-employment may decline from its present level. Employing cross-sectional data from industrial countries and Latin America, Maloney (1999) finds a strong negative relationship between the share of the workforce in self-employment and industrial productivity.

Contract Work

The data cannot approach the level of institutional detail that many case studies offer on the contract sector, nor can the brief period examined establish whether contract work is the result of a process of "deprotecting" workers who were previously protected. However, contract work accounts for a relatively small share of informal production (20 percent), suggesting that it is probably incorrect to generalize subcontracting relations to the informal sector as a whole. Further, the similarities between contract work and self-employment, in particular the common earnings differentials, suggest that common motivations may underlie a worker's decision to engage in contract work and that the sector may not represent inferior work.

There is no characteristic that raises the probability of leaving contract work for salaried formal work as opposed to staying, and the *V*s are symmetrical: there seems to be little evidence of unidirectional graduation from contract work to formal salaried work.[8] Given that those who move to contract work from formal

8. The logit results do not provide evidence of involuntary entrance. The fact that the joint impact of the experience terms is negative for workers coming from formal salaried employment but is positive for workers entering self-employment may suggest that the accumulation of capital is less necessary where the subcontracting firm provides needed inputs.

salaried employment are those with less education, it is possible that the low-skilled laborer who prefers more independence or who might do better on commission than in a factory would voluntarily move to contract work. The differentials between the cost of benefits to firms and the value to workers offer a benign interpretation of informal subcontracting as a way of reducing firms' costs: it becomes a Pareto-improving trade in which contract workers gain the value of benefits, while firms' costs of nonwage labor fall.

Roberts's (1989) interviews of workers in Guadalajara suggest that, given very weak unions and low wages, informalization is not primarily a strategy for reducing remuneration and worker control over production. "Market uncertainty and the large number of income opportunities in the city mean that it is useful for *both* employees and employers to have flexibility in allocating labor" (Roberts 1989: 48, emphasis added). The probabilities in table 5 suggest that contract workers are less likely than other workers to become unemployed, leave the labor force, go to school, or become unpaid workers. Therefore, turnover seems unrelated to instability in employment itself. Instead, it may occur as workers redefine themselves with rapid shifts in clientele: a self-employed worker who takes on a short-term contract may suddenly appear to shift sectors.

More generally, it is possible that subcontracting may not so much avoid labor legislation as it avoids inefficiencies in the law. Given the political difficulties of taking on the anachronisms in the 70-year-old labor code, contract work may represent less a threat to worker protections than a means to induce the flexibility necessary in a modern open economy that the data suggest is not obviously detrimental to the workers involved.

Informal Salaried Employment and Entry into Work

Even if the self-employed benefit from being their own boss, the mainstream view is that those who work for them are the worst off of the urban workforce: salaried, yet without benefits. However, rather than being a stagnant group of disadvantaged workers, the sector appears to serve primarily as the principal, although not exclusive, port of entry for young, poorly educated workers moving into paid employment. The mean age of 29 is 5 years below that of formal and contract workers, and 14 years below that of the self-employed. Table 5 shows a cluster of high mobility among school attendance, unpaid work, and, to a lesser extent, unemployment, suggesting a pool of workers not yet tracked into regular employment. Those leaving school and those who are unemployed show disproportionate movement into unpaid labor (and, to a lesser extent, informal salaried work). The extremely high *V*s between school and unpaid work suggest intermittent work at home or perhaps an apprenticeship before schooling is completed.

Subsequently, unpaid workers move disproportionately into the informal salaried sector. This suggests that while in school and just after completing school, many individuals help out at the family business and eventually get paid. They spend on average only two years doing this before moving onto other paid work.

The brevity of tenure is the same as that found in Brazil by Sedlacek, Paes de Barros, and Varandas (1995) and similar to the tenure of young workers in the United States, where the median is only 1.4 years for workers 16 to 24 years old and 3.4 years for workers 25 to 34 years old (U.S. Department of Labor 1992). Even if this pattern of graduation from school to unpaid work to informal salaried work to other modes of work represents the queuing that the dualistic literature might predict, the time spent in informal salaried work is not long.

If Hemmer and Mannel (1989) are correct that in many countries small informal enterprises train more apprentices and workers than the formal education system and the mostly government job-training schemes together, these years to a large degree may constitute continued schooling. Further, the symmetry of the flows back into informal salaried employment from all three of the other sectors suggests that the opportunities there are not considered uniformly worse than those in the other sectors, nor is the likelihood on involuntary terms in the logit regressions significant, implying that the sector is not primarily a safety net. The logit results suggest that from every sector workers entering informal salaried employment have less experience and schooling than other workers (see table A-1). Balán, Browning, and Jelin (1973: 132) provide one possible explanation: "The first years in the labor force are ones for learning skills, 'shopping around,' exploring alternatives . . . Very few men . . . held at age 25 the same job they had ten years earlier."

The concentration among the very poor and uneducated again suggests the low opportunity cost of leaving formal sector employment. The better-educated workers push up the mean for schooling and wages in the formal sector (tables 1 and 2) and, from the logit results, are more likely to enter formal sector employment. They may not consider salaried employment in the informal sector comparable, but those working menial or assembly-line jobs at less-well-paying formal sector firms may.

IV. Conclusion

This article offers an alternative to the traditional dualistic view of the interaction between the formal and informal sectors and some supportive evidence from observed patterns of transition between sectors. It argues that good reasons for workers to prefer informal employment arise from the desirable characteristics of the various subsectors, the inefficiencies in present labor codes, and the relatively low levels of formal sector productivity in developing countries. It must be stressed that the last two factors may be arguments for labor market reform to the degree that they are affected by poorly designed institutions. However, the frequent inferences of rigidities, segmentation, and distortion in the labor market because of a large informal sector should probably be reconsidered. Both earnings differentials and patterns of mobility indicate that much of the informal sector is a desirable destination and that the distinct modalities of work are relatively well integrated.

Institutional rigidities may account for some fraction of the informal sector, particularly during cyclical downturns. The period examined was a relatively prosperous one in which minimum wages were not binding and union power was weak. Efficiency wage arguments may accurately describe a subsegment of the formal workforce. It is possible that the market is dualistic; however, the good job/bad job division cuts across lines of formality. This article does not deny the possibility of exploitive relations arising from subcontracting, despite its plausible benefits for both parties. However, the informal sector is likely to persist even in the absence of these effects.

Appendix. The Multinomial Logit Analysis of Movements Between Sectors

The second through sixth columns in table A-1 present the results of the multinomial logit analysis in the standard exponential form

$$\frac{P_{ij}}{P_{ii}} = e^{X\beta_j}$$

where the vector β_j measures the degree to which an increase in worker characteristic X (listed across the top of the table) increases the probability of a worker going to sector j relative to the probability of staying in sector i. The worker characteristics include experience, experience squared, and schooling. Because these are often the factors included in Mincerian earnings equations, the real wage in the initial period is included to ensure that the results do not reflect just the wage effect. The interpretation of the coefficients of the first three variables is therefore the impact of a rise in education or experience for a given level of income. By the same logic, the wage term must be interpreted as the wage given the first three variables, that is, earnings above or below what would be predicted by human capital, rather than as absolute income. Although the impact of this measure of unobserved individual characteristics may have interesting interpretations, the variable is included primarily to isolate the school and experience effects. Because these regressions condition on the initial sector, which is likely to result from a self-selection process on the part of the workers, they must be interpreted as such.

For workers who experience a spell of unemployment during the transition between sectors across the five quarters, the survey tabulates whether they left their previous jobs voluntarily. The binary variable takes the value 1 in the relatively infrequent event of involuntary movement into unemployment and 0 otherwise. This variable is included as well as the involuntary variable multiplied by each of the four explanatory variables in the logit regressions. Columns 7 to 11 in table A-1 present the coefficients on these dummy and interactive effects of involuntary separation. The signs and significance of these interactive coefficients, both individually and as a block, show whether workers who were fired have statistically different patterns of transition. In only four cases does the likelihood

Table A1. *Multinomial Logit Analysis of Worker Characteristics Affecting Transitions between Sectors, Mexico, 1991–92*

Workers' transitions between sectors	*Constant*	*Experience*[a]	*Experience*^{2}	*Schooling*	*Real wage*	*Involuntary interactive terms*[b] *Involuntary job loss*	*Experience*	*Experience*^{2}	*Schooling*	*Real wage*	*Likelihood ratio* $\chi^2(5)$
From self-employment to[c]											
Informal salaried	1.27	–0.12	1.34E–03	–0.13	–12.44	628	9.54	–1.24E+00	–115	5,869	24.11
	(3.45)	(6.36)	(4.40)	(4.45)	(3.31)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	[0.00]
Formal salaried	–0.77	–0.03	–1.91E–04	0.00	–1.68	7.18	–0.23	2.01E–03	–0.40	–0.98	21.62
	(2.13)	(1.47)	(0.54)	(0.07)	(0.79)	(2.38)	(2.00)	(1.01)	(1.71)	(0.04)	[0.00]
Contract	–0.97	–0.05	1.49E–04	–0.04	–0.09	–0.92	–0.17	3.39E–03	0.20	9.21	1.55
	(2.06)	(1.96)	(0.33)	(1.12)	(0.08)	(0.08)	(0.97)	(1.10)	(0.47)	(0.21)	[0.91]
From informal salaried to[d]											
Self-employed	–2.19	0.06	–6.42E–04	0.01	17.77	–4.82	0.28	–3.56E–03	0.25	–11.83	4.80
	(5.45)	(2.48)	(1.59)	(0.15)	(2.39)	(1.58)	(1.55)	(1.07)	(1.06)	(0.27)	[0.44]
Formal salaried	–0.70	–0.04	6.87E–04	0.02	23.53	5.04	–0.04	–7.97E–05	–0.39	–166	10.51
	(2.18)	(2.11)	(2.07)	(0.73)	(3.77)	(1.54)	(0.25)	(0.03)	(1.42)	(1.67)	[0.62]
Contract	–2.39	0.01	–5.00E–04	–0.01	34.48	6.24	–0.16	6.26E–04	–0.33	–114	5.25
	(4.63)	(0.38)	(0.74)	(0.18)	(4.49)	(1.75)	(0.75)	(0.10)	(1.06)	(1.25)	[0.39]
From formal salaried to[e]											
Self-employed	–2.66	0.07	–1.27E–03	–0.12	5.10	1.52	–0.02	1.28E–03	0.02	–21.23	29.15
	(8.64)	(3.79)	(3.84)	(4.90)	(1.89)	(1.20)	(0.37)	(1.09)	(0.16)	(0.98)	[0.00]
Informal salaried	0.44	–0.09	9.29E–04	–0.22	–6.34	–0.46	0.05	–4.75E–04	0.05	5.91	8.86
	1.70	(6.21)	(3.48)	(8.92)	(1.51)	(0.37)	(0.75)	(0.33)	(0.37)	(0.30)	[0.12]
Contract	–1.66	0.02	–1.22E–03	–0.19	8.98	–1.48	0.20	–2.85E–03	0.12	–56.78	1.86
	(5.14)	(1.11)	(2.86)	(6.53)	(3.46)	(0.68)	(1.49)	(0.95)	(0.65)	(1.22)	[0.87]

From contract to[f]									
Self-employed	−1.97	0.02	0.05	−5.05	−6.03	0.17	0.54	−29.37	8.29
	(1.47)	(2.78)	(1.29)	(1.17)	(1.29)	(1.90)	(1.39)	(0.62)	[.08]
Informal salaried	−0.01	−0.01	−0.10	−19.34	−6.23	0.16	0.57	−17.49	4.84
	(0.01)	(1.43)	(2.05)	(2.63)	(1.28)	(1.75)	(1.39)	(0.34)	[.30]
Formal salaried	−0.33	−0.01	−0.03	2.77	−0.32	0.05	0.23	−72.01	3.72
	(0.97)	(1.21)	(0.92)	(1.36)	(0.09)	(0.71)	(0.71)	(1.52)	[.44]

Note: The coefficients reflect how experience, experience^2, schooling, and the initial real wage affect the probability of moving from the initial sector to the terminal sector relative to the probability of staying in the initial sector. Columns 7 to 11 present the dummy and interactive effect of involuntary separation from the previous job (tabulated only if unemployed between jobs). The informal sector consists of workers in microenterprises with fewer than six workers. See table 2 for summary statistics and definitions of the four sectors. *Z*-statistics are in parentheses. *P*-values are in square brackets.

a. Sign of the value of $\beta_{exp} + 2\beta_{exp^2}\,\overline{EXP}$ is that of experience with the exception of the transition from formal salaried to contract work.

b. Involuntary interactive dummies = 1 if involuntarily separated and unemployed. The likelihood ratio tests joint significance of interactive effects.

c. Number of observations is 2,503, pseudo R^2 is 0.0540, $\chi^2(27)$ is 238.341[00], and log likelihood is 2,086.

d. Number of observations is 1,266, pseudo R^2 is 0.0296, $\chi^2(27)$ is 90.26[00], and log likelihood is 1,477.

e. Number of observations is 5,377, pseudo R^2 is 0.0397, $\chi^2(27)$ is 281.92[00], and log likelihood is 3,409.

f. Experience^2 is never significant (at 10 percent) in the contract worker regressions and is dropped. Number of observations is 907, pseudo R^2 is 0.0219, $\chi^2(27)$ is 48.42[00], and log likelihood is 1,082.

Source: Author's calculations based on the Mexican National Urban Employment Survey.

ratio confirm the significance of these terms as a block (at the 10 percent level), but in no case does their inclusion substantially alter the initial parameters.

REFERENCES

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

Aroca, Patricio, and William F. Maloney. 1999. "Logit Analysis in a Rotating Panel Context and an Application to Self-Employment Decisions." IBRD Working Paper 2069. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed.

Aronson, Robert L. 1991. *Self-Employment: A Labor Market Perspective.* Ithaca, N.Y.: ILR Press.

Balán, Jorge, Harley L. Browning, and Elizabeth Jelin. 1973. *Men in a Developing Society.* Austin: University of Texas at Austin, Institute of Latin American Studies.

Bell, Linda. 1997. "The Impact of Minimum Wages in Mexico and Colombia." *Journal of Labor Economics* 15(3):S102–35.

Brock, William A., David S. Evans, and Bruce D. Phillips. 1986. *The Economics of Small Businesses: Their Role and Regulation in the U.S. Economy.* New York: Holenamin and Meier.

Dávila Capalleja, Enrique. 1994. "Regulaciones laborales y mercado de trabajo en México." In Gustavo Márquez, ed., *Regulación del mercado de trabajo en América Latina.* San Francisco: Centro Internacional para el Desarrollo Económico.

Esfahani, Hadi, and Djavad Salehi-Isfahani. 1989. "Effort Observability and Worker Productivity: Towards an Explanation of Economic Dualism." *Economic Journal* 99(397):818–36.

Evans, David. S., and Boyan Jovanovic. 1989. "Estimates of a Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97(4):808–27.

Evans, David S., and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79(3):519–35.

Fields, Gary S. 1990. "Labor Market Modelling and the Urban Informal Sector: Theory and Evidence." In David Turnham, Bernard Salomé, and Antoine Schwarz, eds., *The Informal Sector Revisited.* Paris: Organisation of Economic Co-operation and Development.

Gottschalk, Peter, and Tim Maloney. 1985. "Involuntary Terminations, Unemployment, and Job Matching: A Test of Job Search Theory." *Journal of Labor Economics* 3(2):109–23.

Gregory, Peter. 1986. *The Myth of Market Failure: Employment and the Labor Market in Mexico.* Baltimore, Md.: Johns Hopkins University Press for the World Bank.

Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment, and Development: A Two-Sector Analysis." *American Economic Review* 60(1):126–42.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–61.

Hemmer, Hans R., and C. Mannel. 1989. "On the Economic Analysis of the Urban Informal Sector." *World Development* 17(10):1543–52.

Jovanovic, Boyan. 1982. "Selection and Evolution of Industry." *Econometrica* 50(3):649–70.

Krebs, Tom, and William F. Maloney. 1999. "Quitting and Labor Turnover: Microeconomic Evidence and Macroeconomic Consequences." IBRD Working Paper 2068. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed.

Levenson, Alec, and William F. Maloney. 1998. "The Informal Sector, Firm Dynamics, and Institutional Participation." IBRD Working Paper 1988. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed.

Maloney, William F. 1998a. "Are Labor Markets in Developing Countries Dualistic?" IBRD Working Paper 1941. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed.

———. 1998b. "The Structure of Labor Markets in Developing Countries: Time-Series Evidence on Competing Views." IBRD Working Paper 1940. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed.

———. 1999. "Self-Employment and Labor Turnover: Cross-Country Evidence." IBRD Working Paper 2102. Latin America and the Caribbean Region, Poverty Reduction and Economic Management Unit, World Bank, Washington, D.C. Processed. (Forthcoming.)

Marcouiller, Douglas, Veronica Ruiz de Castilla, and Christopher Woodruff. 1997. "Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador, and Peru." *Economic Development and Cultural Change* 45(2):367–92.

Mazumdar, Dipak. 1983. "Segmented Labor Markets in LDCs." *American Economic Review* 73(2):254–59.

Ozorio de Almeida, Anna Luiza, Leandro Alves, and Scott E. M. Graham. 1995. "Poverty, Deregulation, and Employment in the Informal Sector of Mexico." ESP Discussion Paper 54. Education and Social Policy Department, World Bank, Washington, D.C. March. Processed.

Portes, Alejandro. 1994. "The Informal Economy and Its Paradoxes." In Neil J. Smelser and Richard Swedberg, eds., *The Handbook of Economic Sociology*. Princeton, N.J.: Princeton University Press.

Portes, Alejandro, Manuel Castells, and Lauren A. Benton, eds. 1989. *The Informal Economy*. Baltimore, Md.: Johns Hopkins University Press.

Roberts, Bryan. 1989. "Employment Structure, Life Cycle, and Life Chances: Formal and Informal Sectors in Guadalajara." In Alejandro Portes, Manuel Castells, and Lauren A. Benton, eds., *The Informal Economy*. Baltimore, Md.: Johns Hopkins University Press.

Rosenzweig, Mark R. 1988. "Labor Markets in Low-Income Countries." In Hollis Chenery and T. N. Srinivasan, eds., *Handbook of Development Economics*. Volume 1. New York: North-Holland.

Sabot, R. H. 1977. "The Meaning and Measurement of Urban Surplus Labor." *Oxford Economic Papers* 29(3):389–411.

Sedlacek, Guilherme L., Ricardo Paes de Barros, and Simone Varandas. 1990. "Segmentação e mobilidade no mercado de trabalho: A carteira de trabalho em São Paulo." *Pesquisa e Planejamento Economico* 20(1): 87–103.

Stiglitz, Joseph E. 1974. "Alternative Theories of Wage Determination and Unemployment in LDCs: The Labor Turnover Model." *Quarterly Journal of Economics* 88(2):194–227

Thomas, J. J. 1992. *Informal Economic Activity*. Ann Arbor: University of Michigan Press.

Tokman, Victor E. 1992. *Beyond Regulation: The Informal Economy in Latin America*. Boulder, Col.: Lynne Rienner.

U.S. Bureau of National Affairs. 1994. "Bulletin to Management." Washington, D.C. December 8. Processed.

U.S. Department of Labor. 1992. *The Informal Sector in Mexico*. Washington, D.C.

———. 1993. *Work without Protections: Case Studies of the Informal Sector in Developing Countries*. Washington, D.C.

World Bank. 1995. "Labor and Economic Reforms in Latin America and the Caribbean." *Regional Perspectives on World Development Report*. Washington, D.C.

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2 303-31

Tax Incidence in Madagascar: An Analysis Using Household Data

Stephen D. Younger, David E. Sahn, Steven Haggblade, and Paul A. Dorosh

This article discusses tax incidence in Madagascar and asks who pays the taxes that finance government spending. Its main concern is to identify the progressivity of different taxes levied in Madagascar, based on the consumption and income patterns found in the 1994 Enquête Permanente auprès des Ménages, *a nationally representative survey. The results suggest that most taxes are progressive, meaning that wealthy households pay proportionately more of these taxes relative to their expenditures than do poor households. Two notable exceptions are taxes on kerosene and export duties on vanilla, both of which are regressive. These results are consistent with those of a study of Ghana, the only other comparable research on tax incidence in Africa. That study found taxes on kerosene and cocoa exports to be the most regressive taxes in Ghana.*

Making firm policy recommendations for tax reform would require an analysis of the economic efficiency and administrative efficacy of different taxes to complement this article's work on their equity implications. Nevertheless, the results suggest that the movement away from trade taxes, especially export duties, and toward broadly based value added or income taxes would be more equitable and more economically efficient. The only legitimate impediment to such reforms in Madagascar is administrative, that is, the government's ability to collect different taxes effectively. Although administrative efficiency may be a problem for value added or income taxes, taxes on petroleum products (except kerosene) are highly progressive and provide a good tax handle.

Maintaining fiscal balance is central to any adjustment effort. The most successful adjustment programs in Africa and elsewhere have quickly and permanently eliminated fiscal deficits, while many other attempts to reestablish macroeconomic control have foundered on an inability to match expenditures to revenues. Despite this generally accepted maxim, many critics of adjustment programs have expressed concern that fiscal stabilization will hurt the poor. Most of that literature has focused on expenditure reductions and their likely impact on poor households. The report produced by Cornia, Jolly, and Stewart (1987) for the United Nations Children's Fund is the most famous example (Sahn, Dorosh, and Younger

Stephen D. Younger, David E. Sahn, Steven Haggblade, and Paul A. Dorosh are with the Cornell Food and Nutrition Policy Program at Cornell University. The authors conducted part of this research while Stephen D. Younger was a visiting Research Fellow at the Free University in Amsterdam. They gratefully acknowledge that university's financial support and hospitality.

1996 contest many of their arguments). Even though tax increases have received much less attention (Younger 1996 is an exception), similar concerns about how tax reform affects poverty and inequality are relevant. When governments change the structure of taxation, who suffers the incidence of those taxes?

This question is especially pertinent in Africa, where interest in reforming taxes takes two directions. First, the ratio of tax revenue to gross domestic product (GDP) is low in Africa (except for countries with large mineral royalties), often no more than 10 or 12 percent of GDP. Thus there is a need to raise overall revenue. Second, many tax systems in Africa are very distortionary, concentrating on trade taxes while neglecting direct taxes and broad-based indirect taxes, such as a value added tax (VAT). To some extent, both of these problems reflect deeper structural issues. Tax handles for broad-based taxes are limited in economies with small formal sectors that are dominated by the public sector. Further, formal sector enterprises and workers are often politically powerful and thus able to lobby against their own taxation.

The situation in Madagascar is typical of these general patterns. Tax revenues fell steadily during the past 20 years, while expenditures remained constant (with the exception of a sharp decline in 1981), leaving a substantial and growing fiscal deficit. At the same time, trade taxes continue to account for more than half of Madagascar's tax revenue, significantly distorting incentives. Thus, like many African countries, Madagascar needs to increase its tax revenue and do so in a less distortionary way.

Policymakers interested in tax reform must consider tax incidence. This article begins to address that issue for Madagascar. We use household income and expenditure data from the *Enquête Permanente auprès des Ménages* (EPM; Government of Madagascar 1994), a nationally representative survey of 4,500 households conducted from April 1993 to April 1994 by the *Institut National de la Statistique* (INSTAT 1995). We use these data to determine which households are likely to pay certain taxes. We then judge the progressivity of the tax based on whether paying households are from the lower or upper ends of the expenditure distribution.

I. Tax Reform in Madagascar

During most of the 1980s Madagascar maintained relatively low budget deficits as part of its stabilization and structural adjustment efforts. But since 1988, and especially after 1990, the fiscal deficit has widened considerably, threatening macroeconomic stability and prospects for economic growth. The share of government expenditures in GDP rose from 16.7 percent in 1990 to 19.7 percent in 1993, before falling back to 17.6 percent in 1995. Over the same period government revenues fell from 11.8 percent to only 8.3 percent of GDP (table 1). Thus by 1995 Madagascar's fiscal deficit stood at 9.3 percent of GDP.

Trade taxes account for more than half of Madagascar's government revenue, and fluctuations in trade tax revenues explain most of the variation in total gov-

Table 1. *Central Government Budgetary Revenue in Madagascar, 1978 and 1988–95*
(percentage of GDP)

Type of revenue	*1978*	*1988*	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*
Tax revenue	12.7	10.5	8.8	9.4	6.8	8.7	8.2	7.7	8.1
Taxes on net income and profits	2.6	1.5	1.1	1.5	1.3	1.3	1.6	1.7	1.2
Taxes on property	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Taxes on domestic goods and services	3.5	2.8	2.5	2.2	1.7	2.7	2.2	2.2	2.1
Taxes on foreign trade	6.3	5.8	5.0	5.5	3.8	4.5	4.5	3.7	4.6
Import duties	4.4	4.6	3.9	4.7	3.1	4.0	4.0	3.5	4.3
Import duties on petroleum	—	0.3	0.2	0.1	0.2	0.6	0.9	0.2	0.6
Export duties	1.9	1.3	1.2	0.8	0.7	0.5	0.2	0.2	0.3
Other taxes	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Nontax revenue	0.7	0.3	0.5	0.7	0.4	0.3	0.4	0.5	0.2
Total budgetary revenue	13.4	10.7	9.3	10.1	7.2	9.0	8.6	8.2	8.3
Nonbudgetary revenue	3.0	2.3	1.9	1.7	1.3	0.8	1.1	0.1	0.0
FNUP[a]	3.0	2.0	1.3	0.8	0.5	0.8	0.2	0.1	0.0
Export duties including FNUP[a]	4.9	3.3	2.5	1.7	1.2	1.3	0.4	0.3	0.3
Total trade taxes including FNUP[a]	9.3	7.8	6.4	6.4	4.3	5.2	4.7	3.8	4.6
Nontrade taxes	7.1	5.2	4.9	5.4	4.2	4.6	5.1	4.5	3.7
Total revenue	16.4	13.1	11.3	11.8	8.5	9.8	9.7	8.3	8.3

— Not available.
a. FNUP is the *Fonds National Unique de Péréquation*, the export price stabilization fund.
Source: World Bank (1984, 1986, 1991); International Monetary Fund data.

ernment revenue in the past two decades. In the late 1970s, when world prices of Madagascar's major exports—coffee, vanilla, and cloves—were high, export tax revenues, including revenues of the export price stabilization fund, *Fonds National Unique de Péréquation* (FNUP), surged, helping to finance an ill-fated public investment push (the *investir à outrance* policy). Import tax revenues also increased as imports of capital and intermediate goods rose, so that trade taxes reached 9.3 percent of GDP in 1978. World prices of Madagascar's exports fell in the early 1980s, however, and trade tax revenues also declined, necessitating large cuts in government expenditures to reduce budget deficits. Later in the decade the government liberalized marketing and eliminated most export taxes on all crops, except vanilla, as part of an overall trade reform designed to promote exports. These lower tax rates together with a decline in world prices of exports reduced export tax revenues from 5.0 percent of GDP in 1987 to 0.3 percent in 1995 (table 1).

Explicit export taxes represented only part of the total taxation on export crop producers. In most years prior to 1994 Madagascar maintained foreign exchange controls that forced exporters to surrender their foreign exchange earnings at the official exchange rate, while limiting importers' access to foreign exchange. Excess demand for foreign exchange at the official exchange rate led to a parallel market. Exporters were thus implicitly taxed through the overvalued exchange rate, with the premium in the parallel market approximating this implicit tax.[1] These implicit taxes on coffee producers significantly raised the total rate of taxation. The system of fixed producer prices and government-controlled marketing resulted in generally large, but variable, rates of effective taxation on exports and wide fluctuations in FNUP revenues.

Government efforts to raise tax revenues through other channels have not been effective. Nontrade tax revenues have declined along with export tax revenues, from 5.4 percent of GDP in 1990 to only 3.7 percent in 1995. Recent efforts to reform the tax system have shifted focus from sales taxes and turnover taxes to a VAT, which was instituted in 1994. In theory, the advantage of a VAT is that it is less distortionary than turnover taxes, which essentially tax intermediate purchased inputs twice: first, when the firm purchases intermediate goods and, second, when the firm sells final goods. Likewise, the VAT is in theory more efficient than import tariffs, which raise the domestic price of imported goods, thus protecting import-substituting industries. In contrast, a uniform VAT taxes both the value added content of imports and the value added in domestically produced goods at the same rate.

Madagascar's VAT differs from a uniform nondistortionary VAT in two respects: many import tariffs remain, and most domestic production escapes the VAT. Thus

1. A better measure of the indirect tax rate on exporters is the difference between an equilibrium real exchange rate and the actual official real exchange rate (Krueger, Schiff, and Valdes 1988). Dorosh, Bernier, and Sarris (1990) use this methodology to calculate that Madagascar's total tax on coffee producers in 1981–87 averaged 77.5 percent of the border price measured at the equilibrium exchange rate, with the indirect tax from exchange rate distortions equaling 49.3 percent.

in 1995 total revenues from the VAT on domestic value added equaled only 232.6 billion Madagascar francs (FMG)—only about 8 percent of value added in Madagascar's formal sector. Officially, however, the VAT rate was 25 percent. In contrast, the average import tariff (calculated as total import tariffs divided by the c.i.f. value of imports) was 23 percent, whereas the VAT on imports equaled 10.9 percent of the c.i.f. value of imports.

II. Methodology

In general terms, a tax transfers real purchasing power from households to the government. The incidence of the tax refers to who pays the tax in real terms, that is, whose real purchasing power falls when the government imposes the tax. Taxes are said to be progressive if poorer households pay a proportionately smaller share of the tax than do wealthy households, relative to some measure of overall welfare, usually income or expenditures. Taxes are regressive if the opposite is true and neutral if tax shares are equal to overall income or expenditure shares.

In this study we use household expenditures (per capita) rather than income as our welfare measure. We do this for two reasons. Practically, households tend to report their expenditures more accurately than they report their incomes. They have less incentive to hide expenditures than income (from the enumerator and from family members). Theoretically, the life-cycle/permanent-income hypothesis suggests that expenditures are a more stable representation of a household's long-term welfare than is income, because households try to smooth their expenditures given income fluctuations over time. As a result, expenditures reflect households' own estimates of their permanent income over time and are thus a better proxy for their long-term welfare.

Statutory and Economic Incidence

Since the work of David Ricardo economists have understood that the entities that are legally required to pay a tax are not necessarily those that suffer a reduction in real purchasing power from imposition of the tax. They may successfully shift the tax onto other households. Governments in developing countries collect most taxes from firms, but firms do not suffer reductions in purchasing power. Either the households that own them do, or firms shift the taxes onto their customers or suppliers by changing prices. For example, it is standard to assume that if an industry is competitive, then a tax on its product will be passed onto consumers through a price increase equal to the tax rate. In contrast, a tax on a firm's profits probably falls mostly on the firm's owners. The other common example is avoiding a tax by changing one's pattern of consumption or income. For instance, households that have high elasticities of demand for gasoline can avoid paying a tax on gasoline consumption by switching to substitutes with little loss in welfare, while those with an inelastic demand cannot do the same so easily.

The *economic incidence* of a tax refers to where the reduction in real purchasing power falls, while the *statutory incidence* refers to who is legally required to pay the tax. Clearly, it is the economic incidence that is of interest in any analysis of how taxes affect poverty and inequality. Unfortunately, it is often much easier to identify the statutory incidence. Here, we examine economic incidence, but in doing so we must rely on strong assumptions. For direct taxes we assume that the factors producing the associated incomes pay the taxes. Thus wage workers pay the withholding tax on wage income, business owners pay the tax on their firm's profits, and so on. This assumption is equivalent to assuming that households supply the associated factors completely inelastically so that they cannot shift the tax. Selden and Wasylenko (1992) defend this assumption on the grounds that, while restrictive, it often produces results similar to those of more sophisticated models, but at a substantially lower cost in terms of time and effort.

For indirect taxes we assume that households that consume the taxed items pay the associated taxes. Thus smokers pay taxes on tobacco, households that use kerosene for lamps pay taxes on kerosene, and so on. There are, however, two exceptions to this general rule, made largely because of the controversy that surrounds two types of taxes. For gasoline taxes no one doubts that direct consumption of gasoline is highly concentrated in the upper end of the expenditure distribution. Yet critics of gasoline taxes argue that the secondary impact of such taxes is regressive because an increase in gasoline prices gives rise to increases in other prices, especially transport, on which poor people depend more than rich people. To include this effect, we assume that the gasoline tax falls on both direct consumers of gasoline and on consumers of public transport. Still, this adjustment is only partial, because it does not include the effects of the tax when transport is an intermediate product.

Import duties are the other type of tax that is difficult to manage. Household surveys do not ask whether goods consumed are imported or not, so we cannot directly identify consumers of imports. Rather, we assume that the prices of all goods for which imports are a large share of the market go up by the amount of the tariff when it is imposed. Thus those who pay the tax are consumers of the good, whether it is imported or produced domestically. But the full payment does not go to the government. A share of the benefits from the import duty goes to protected local producers who can charge a higher price for their output. Thus the cost to consumers is not equal to the government's revenue.

We describe our tax calculations in greater detail in appendix A. Data problems call for some caution in interpreting results. Total household consumption is substantially underreported in the EPM survey, which may be due to underreporting (by households) or to undervaluing consumption from own production. In addition, the total value of taxes paid by households in the survey is lower than revenues reported by the Ministry of Finance. For example, the value of the VAT, income tax, and vanilla tax reported in the survey is three-quarters or more of government revenue. In contrast, this figure is only 56 percent for excises on alcohol and tobacco, probably reflecting underreporting of consump-

tion. Import taxes on food and consumer products, excluding petroleum, are roughly equal in the survey and in the national accounts. But despite the good correspondence of consumer products, we do not capture the import taxes paid on intermediate inputs and raw materials.

Conceptual Methods for Evaluating Tax Incidence

A general method for studying tax incidence is to test for "welfare dominance" (Yitzhaki and Slemrod 1991). Yitzhaki and Slemrod construct concentration curves—diagrams that are similar to Lorenz curves in that they align households from the poorest to the wealthiest along the horizontal axis and the cumulative proportion of taxes paid along the vertical axis. Yitzhaki and Slemrod then prove that for any social welfare function that favors an equitable distribution of income, changing the tax structure by slightly reducing taxes on good x and increasing those on good y by just enough to keep total revenues constant will improve social welfare when x's concentration curve is everywhere above y's.[2] In this case we say that taxing x dominates taxing y (or x dominates y). The intuition is straightforward. If poorer households tend to consume less of a particular good, say gasoline, and more of another, say food, then reducing taxes on food and raising taxes on gasoline will improve the distribution of welfare. Yitzhaki and Slemrod refer to this as "welfare dominance," making an analogy to the concept of second-order stochastic dominance in the finance literature. The concentration curve for food is above that for gasoline because poorer households account for a larger share of total food consumption than gasoline consumption.

In addition to comparing the concentration curves for different taxes, it is also insightful to compare each tax's concentration curve to two benchmarks: the Lorenz curve for expenditures and the 45-degree line. A tax whose concentration curve is below the Lorenz curve for expenditures is progressive, and a tax whose concentration curve is above the Lorenz curve is regressive.[3] As the tax's concentration curve approaches the 45-degree line, it becomes extremely regressive, as in a head tax.

Statistical Tests

Unlike many other works in the field (including Yitzhaki and Slemrod 1991), we use statistical tests to determine whether one concentration curve is everywhere above another. In particular, we use Davidson and Duclos's (1997) variance-covariance estimator for the ordinates of two possibly dependent con-

2. Technically, the argument also requires that the efficiency consequences of the tax change be at least neutral, that is, that the efficiency of the allocation of resources not worsen with the change. This condition is more difficult to identify in practice, but we will assume that it is satisfied in our discussion.

3. A referee pointed out that this definition of progressivity is less stringent than the one usually found in public finance textbooks: that the marginal tax rate exceed the average tax rate everywhere. Because we are working with cumulative tax payments, it is possible for the marginal rate to fall below the average rate in some ranges and still be progressive by our definition.

centration curves to test for differences in these ordinates (see appendix B). This procedure involves testing for differences at a finite number of ordinates, thus restricting the range of the dominance test. Here, we use 20 evenly spaced ordinates, so that our conclusions are valid for the range 0.05–0.95.

Typically, researchers who apply statistical methods to test for differences between the ordinates of two concentration curves at several abscissa use t-tests. They reject the null hypothesis of nondominance when one of the ordinates differs statistically in the direction of dominance and none of the other pairs differs statistically in the opposite direction (see, for example, the recent article by Gouveia and Tavares 1995). Howes (1996) uses both theoretical and simulation arguments to show that this procedure probably rejects the null too frequently, especially when the concentration curves cross. He argues that we can only be sure of the probability of a type I error (that is, the size of the test) if we reject the null hypothesis when the difference in the ordinates of the two curves is nonzero for *every* ordinate tested (and when the difference is of the same sign).

By using more careful statistical procedures, we reduce the power of the test so that we often do not reject the null. When the dominance tests are inconclusive, we can draw conclusions only by being more specific about the importance of each household in the social welfare function. To do this, we rely on cardinal measures of inequality. (Recall that if one distribution is welfare dominant over another, then the first will be preferred to the second under any social welfare function that favors progressivity.) For example, if we rank taxes by the Gini coefficients for their concentration curves, we will always have an ordering, but it comes at a price: by comparing Gini coefficients for different concentration curves, we implicitly accept the social welfare function of the Gini formula. Another welfare function might yield a different ordering. Yitzhaki (1983) provides a middle ground between the normative generality (and consequent indeterminacy) of the welfare-dominance approach and the precision (and lack of normative generality) of the Gini coefficient. He shows that an extended Gini coefficient can adjust the weight given to poorer households and thus better depict how more progressive social welfare functions would rank different taxes. The coefficient is defined as:

$$G(v) = -v\,[\mathrm{cov}\{e,[1 - F(y)]^{(v-1)}\} / \bar{e}],\ v > 1 \tag{1}$$

where e measures households' payment of a tax, $F(y)$ is the cumulative distribution of all households ranked from the poorest to the richest, $\bar{e}$ is the mean of e, and v is a parameter that affects the weighting of each point on the Lorenz curve. In particular, $G(2)$ yields the traditional Gini coefficient, while values of v greater than 2 yield measures that give greater weight to poorer households. Thus by calculating the extended Gini coefficient for increasing values of v, we can gain a sense of how a wide range of increasingly progressive social welfare functions rank the value of a given tax. We calculate Ginis for values of v from 1.01 to 10, in increments of 0.5. If all pairs of extended Ginis are significantly different in the same direction, then we conclude that one tax "Gini dominates" the other. This concept

clearly is not as general as the ordinal measure, but the implied policy conclusion is similar, even if based on cardinal measures.

Finally, we examine the sensitivity of our results to the choice of household equivalence scale. Much recent poverty literature argues that this choice is arbitrary and that any method of adjusting household incomes for differences in household size and composition reflects the researcher's value judgment more than an empirically testable scale (Deaton and Muellbauer 1980, 1986; Lanjouw and Ravallion 1995; Browning 1992; Blundell and Lewbel 1991; and Coulter, Cowell, and Jenkins 1992). Previous empirical research has shown the importance of the judgments made regarding equivalence on measures of inequality, particularly the range of cardinal measures, but also ordinal measures represented by the Lorenz curve (see, for example, Cutler and Katz 1992; Sahn, Younger, and Simler 1997). In this article we add to those findings by testing the sensitivity of the dominance tests and the extended Ginis to two single-parameter characterizations of the money metric of equivalent income: we set the elasticity with respect to household size initially to 1.0 (the per capita measure) and then to 0.5.

Sensitivity to Assumptions about Tax Incidence

Because concentration curves are based on cumulative shares of the consumption of a particular commodity, they are not sensitive to errors in the assumptions we make about the amount of tax paid per unit of the good consumed as long as the error is the same across households. Such errors are proportional to consumption and thus cancel the share of consumption from both the numerator and the denominator. The same is true of the actual tax rate that applies. If we want to consider the incidence of taxes on several different goods at once, however, then errors in the assumptions about taxes matter, because we add the actual taxes (not a ratio) across commodities. For example, suppose that we undervalue by 50 percent the taxes that households pay for alcohol consumption. The concentration curve for taxes on alcohol is not affected because each household's share of total consumption (and therefore total calculated taxes) remains the same. But problems would arise if we added the tax on alcohol to the tax on tobacco and checked the incidence of the two together. Because the estimated alcohol tax is too low, its weight in the composite commodity comprising alcohol and tobacco is also too low, and the concentration curve for the two together, which is a weighted average of the individual curves, will be weighted too little by the concentration curve for alcohol and too much by the curve for tobacco.

In the case of taxes that fall on many goods, such as the VAT, such problems are unavoidable. Still, we prefer to keep the taxes as disaggregated as possible to avoid the potential errors that composite goods present. In particular, we make no attempt to judge the overall progressivity of the entire tax system. Instead, we stick to individual taxes.

Another problem we face is that households misreport their consumption, especially of alcohol and tobacco. If underreporting is correlated with household

Table 2. *Dominance Results for Taxes in Madagascar, 1994*
($\theta = 1$)

Variable	*45-degree line*	*Vanilla*	*Kerosene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Auto-mobile*	*Transport and gasoline*[b]	*Gaso-line*[c]
45-degree line			D	D	D	D	D	D	D	D	D	D	D	D	D
Vanilla			X	X	X	X	X								
Kerosene		X		D	D	D	D	D	D	D	D	D	D	D	D
Expenditures		X					D	D	D	D	D	D	D	D	D
Imports		X						D	D	D	D	D	D	D	D
Petroleum		X										D	D	D	D
Value added tax		X										D	D	D	D
Alcohol												D	D	D	D
Gasoline via transport[a]												D	D	D	D
Excises												D	D	D	D
Tobacco												D		D	D
Wages															
Automobile															
Transport and gasoline[b]															
Gasoline[c]															

Note: D indicates that we reject the null hypothesis in favor of dominance, that is, the item in the column dominates the item in the row. *X* indicates that we reject the null in favor of crossing. The elasticity with respect to household size, θ, is set at 1.0.

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.

b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.

c. Gasoline refers to the direct tax on gasoline.

Source: Calculated from Government of Madagascar (1994).

expenditures (over the welfare measure), then our results will be biased. For example, if wealthier households consume more alcohol than poorer households, but they do not report their consumption, perhaps because consuming more than a small amount is frowned on, then our estimate of the tax incidence will be too regressive. A similar problem occurs with tax evasion. We are applying statutory tax rates to observed income and expenditure patterns. But wealthier people may be better at avoiding taxes, especially direct taxes, so that our estimated incidence will be too progressive.

III. Results

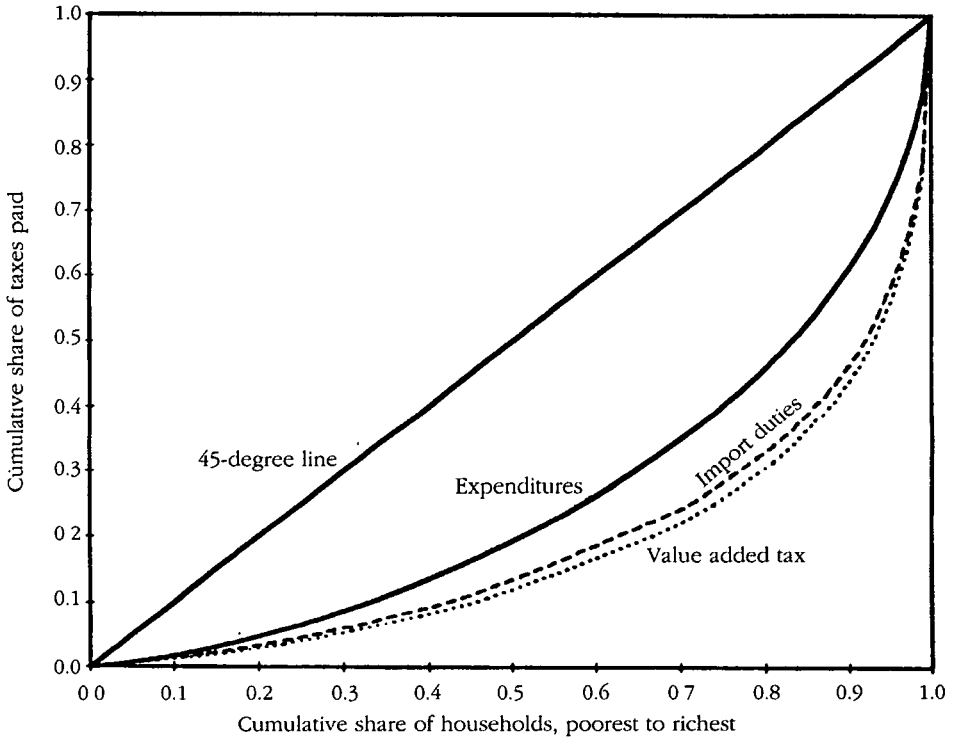
For the most part the results based on the dominance testing and extended Ginis give similar results.[4] Table 2 summarizes the results of statistical tests of the null hypothesis that the concentration curves for two taxes are equal and that they are equal to the expenditure Lorenz curve. Recall that welfare dominance means that a slight increase in the dominant tax, offset by a slight reduction in the dominated tax, will improve social welfare for any welfare function that favors a more equitable distribution of expenditures. Put another way, the dominant tax is more progressive than the other.

The two most important taxes in Madagascar are import taxes (duties and tariffs) and the VAT. The VAT is progressive, while import duties are not, although we cannot reject the null hypothesis that they have the same concentration curves (figure 1 and table 2). However, we cannot reject the null because of one insignificant *t*-statistic (at 0.05). Tests based on the extended Ginis do reject the null (extended Ginis are listed in table 3). Thus, at first appearance, reforming the tax structure by moving away from import duties and toward the VAT may have a positive, albeit small, impact on the after-tax distribution of welfare in Madagascar. But our analysis has not captured an important difference between these two taxes. We have assumed that import duties increase the price of all goods of the same type, whether imported or not, so that buyers of those goods suffer the incidence of the tax. The offsetting increase in real purchasing power does not go entirely to the government, however. Local producers receive part of the increase because the import protection allows them to charge higher prices. Under the VAT, however, all of the benefits go to the government. While we cannot be sure of the incidence of all the benefits of import taxes, it is reasonable to assume that the benefiting firms are owned by households at the upper end of the expenditure distribution. So, the net effect of the import duties is even less progressive than this analysis indicates, meaning that a shift from import taxes to the VAT is likely to be even more progressive than figure 1 suggests.

The other indirect taxes in Madagascar are excise duties on specific products: alcohol, tobacco, automobiles, and petroleum products (figures 2 and 3). The

4. Ordinate estimates and standard errors for tax concentration curves are given in appendix C. Ordinate estimates, their standard errors, and *t*-statistics for differences between each transfer and the Lorenz curve are available from the authors.

Figure 1. *Concentration Curves for Import Duties and the Value Added Tax in Madagascar, 1994*



Note: Expenditure and tax data are on a weighted and per capita basis.
Source: Calculated from Government of Madagascar (1994)

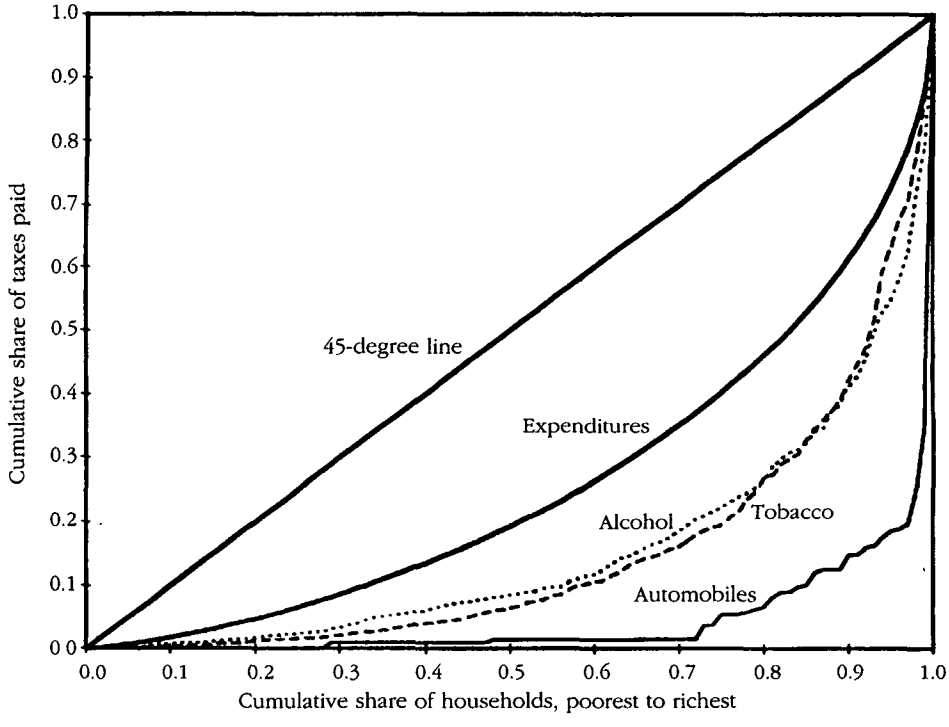
concentration curves for excise taxes on tobacco and alcohol are statistically indistinguishable from the concentration curve of the VAT but dominate that of import taxes. Taxes on automobiles are more progressive than import duties, the VAT, and other excise taxes (table 2). Taxes on direct consumption of gasoline are also more progressive than all other taxes except automobile duties. Nevertheless, most gasoline is consumed as an intermediate input to other services, predominantly transport. We try to capture at least part of this indirect impact by assuming that gasoline accounts for 20 percent of the cost of intracity and intercity transport. (That is the input-output coefficient for petroleum in the transport sector.) We then assume that part of the gasoline tax falls on users of public transport, and we construct a concentration curve for that part of the tax, as well as one for the combined effect of direct purchases of gasoline and indirect purchases through public transport (figure 3). The results show that even the tax on public transport is progressive (although less so than the direct consumption of gasoline), mostly because it is concentrated among urban households. The combined tax (transport and gasoline) dominates both the VAT and import taxes.

Table 3. *Extended Gini Coefficients for Taxes in Madagascar, 1994*

v[a]	*45-degree line*	*Vanilla*	*Kero-sene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[b]	*Excises*	*Tobacco*	*Wages*	*Auto-mobile*	*Transport and gasoline*[c]	*Gasoline*[d]
1.5	0.000	0.0711	0.1915	0.3283	0.3429	0.4138	0.4538	0.4911	0.4813	0.4972	0.5017	0.6698	0.7646	0.8559	0.8700
2.0	0.000	0.1551	0.2838	0.4569	0.4795	0.5092	0.5911	0.6496	0.6561	0.6707	0.6861	0.8482	0.8923	0.9638	0.9701
4.0	0.000	0.4353	0.4369	0.6347	0.6624	0.6230	0.7426	0.8132	0.8437	0.8502	0.8770	0.9679	0.9755	0.9956	0.9973
10.0	0.000	0.6650	0.5759	0.7589	0.7801	0.7157	0.8278	0.8967	0.9273	0.9267	0.9485	0.9929	0.9942	0.9963	0.9969

a. A parameter that affects the weighting of each point on the Lorenz curve.
b. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.
c. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.
d. Gasoline refers to the direct tax on gasoline.
Source: Calculated from Government of Madagascar (1994).

Figure 2. *Concentration Curves for Alcohol, Tobacco, and Automobile Excises in Madagascar, 1994*

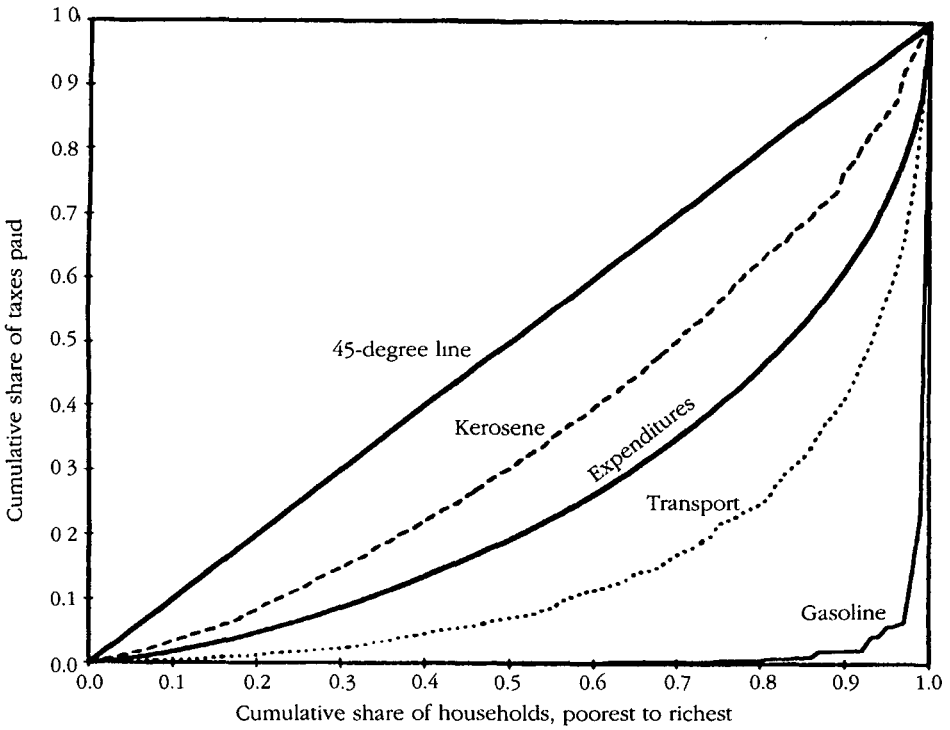


Note· Expenditure and tax data are on a weighted and per capita basis.
Source: Calculated from Government of Madagascar (1994).

Thus policymakers can rest assured that taxes on gasoline are not falling disproportionately on the poor. Rather, they are quite progressive. In contrast, taxes on kerosene, widely used as a fuel for lighting and cooking, are regressive. To the extent that it is both technically and practically feasible, it would be preferable to concentrate duties on gasoline and to reduce duties on kerosene. This result is similar to those in Younger (1996) and Yitzhaki and Lewis (1996).

The last two taxes we examine are the direct tax on wage earnings and the export duty on vanilla (figure 4). The tax on wages, which makes up a little less than half of direct taxes (the rest comes from corporations), is highly progressive, as we would expect, because it falls entirely on workers in the formal sector. For vanilla duties, even though the concentration curve is well above all others, we find no statistical difference between it and the concentration curves for other taxes and between it and the Lorenz curve. This finding is due to the small number of vanilla producers in the sample, 103, which makes rejection of the null difficult. Statistical comparisons of the extended Ginis support this view (table 4). By this criterion vanilla taxes are more regressive than many other types of taxes, particularly excises and taxes on wages and salaries. But the statistical

Figure 3. *Concentration Curves for Petroleum Excises in Madagascar, 1994*



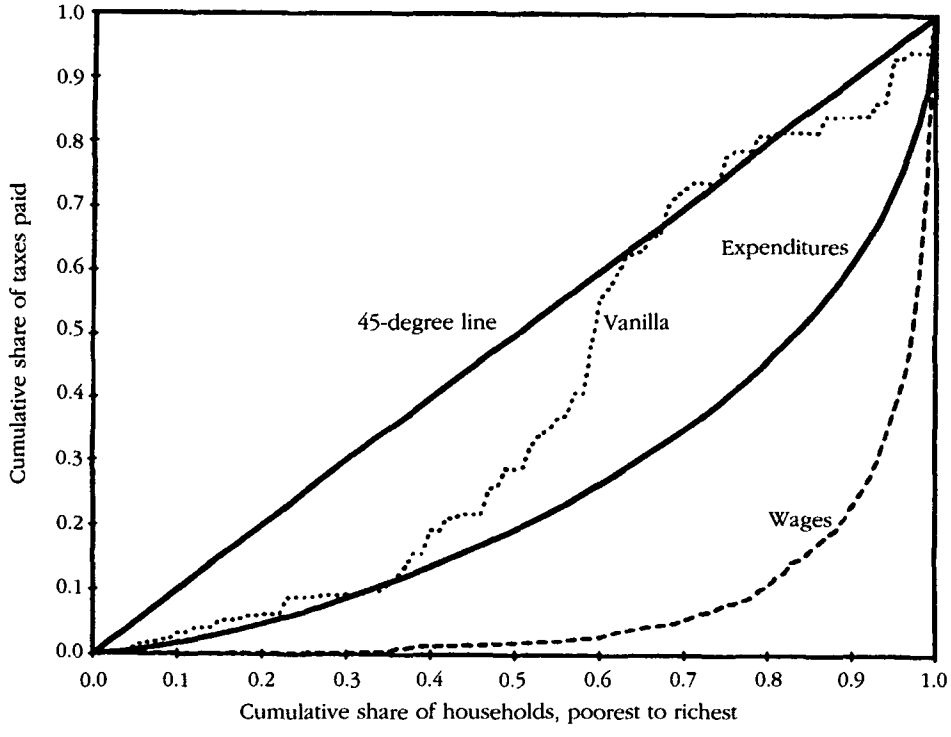
Note: Expenditure and tax data are on a weighted and per capita basis.
Source: Calculated from Government of Madagascar (1994).

comparisons still fail to show that vanilla taxes are regressive or that they are more regressive than the VAT or import duties. If we had a larger sample, we would expect to see that vanilla producers, who are rural farmers, are not as wealthy as the population in general. Thus the government's movement away from export duties probably has a positive distributional impact.

Next we examine the extent to which the results presented above are sensitive to the scale factor used in adjusting household incomes. The results continue to show that most of the taxes we examine are progressive (tables 5 and 6). Likewise, the general ordering of the taxes does not depart dramatically from that generated when we used the per capita equivalence unit, although there are fewer cases of dominance. This lack of sensitivity to household scale is unusual. As discussed, most other studies found that their results were not at all robust to size elasticity. One possible explanation is that the taxes we study in this article are not designed to be correlated with household size, while transfer payments studied in other articles usually are. Thus our results are less sensitive to household scaling than others found in the literature.

Finally, the extended Gini comparisons of progressivity are even more robust to assumptions about adjusting income for household size. In fact, there are only

Figure 4. *Concentration Curves for Vanilla Duties and Wages in Madagascar, 1994*



Note: Expenditure and tax data are on a weighted and per capita basis.
Source: Calculated from Government of Madagascar (1994).

four differences in the pair-wise comparisons of progressivity when using the different equivalence units. Not only do the Ginis imply a clearer ordering of the progressivity of different types of taxes, but the results are less sensitive to different equivalence scales.

IV. Conclusions

The progressivity of the taxes we examine is striking. In Ghana, the only other African country for which we have a comparable analysis, the broad-based consumption taxes are neutral, and the income tax, while progressive, is less so than in Madagascar. This contrast is all the more impressive given the high concentration of per capita consumption in Madagascar. (Madagascar's Gini for per capita expenditures is 0.48 compared with 0.36 for Ghana.) It is interesting to note that the only regressive taxes in Madagascar, those on kerosene and (probably) agricultural exports, are also the regressive taxes in Ghana.

Our analysis informs the debate on tax reform in Madagascar in several ways. First, economists usually argue that tax reform should shift the tax structure

Table 4. *Gini Dominance Results for Taxes in Madagascar, 1994*
($\theta = 1$)

Variable	*45-degree line*	*Vanilla*	*Kero-sene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Auto-mobile*	*Transport and gasoline*[b]	*Gasoline*[c]
45-degree line			D	D	D	D	D	D	D	D	D	D	D	D	D
Vanilla								D	D	D	D	D	D	D	D
Kerosene				D	D	D	D	D	D	D	D	D	D	D	D
Expenditures							D	D	D	D	D	D	D	D	D
Imports							D	D	D	D	D	D	D	D	D
Petroleum												D	D	D	D
Value added tax											D	D	D	D	D
Alcohol												D	D	D	D
Gasoline via transport[a]												D	D	D	D
Excises												D	D	D	D
Tobacco												D	D	D	D
Wages															D
Automobile															
Transport and gasoline[b]															
Gasoline[c]															

Note: D indicates that we reject the null hypothesis in favor of dominance, that is, the item in the column dominates the item in the row. X indicates that we reject the null in favor of crossing. The elasticity with respect to household size (θ) is set at 1.0.

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.

b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.

c. Gasoline refers to the direct tax on gasoline.

Source: Calculated from Government of Madagascar (1994).

Table 5. *Dominance Results for Taxes in Madagascar, 1994*
($\theta = 0.5$)

Variable	*45-degree line*	*Vanilla*	*Kero-sene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Auto-mobile*	*Transport and gasoline*[b]	*Gasoline*[c]
45-degree line		X	D	D	D	D	D	D	D	D	D	D	D	D	D
Vanilla	X		X			X									
Kerosene		X		D	D	D	D	D	D	D	D	D	D	D	D
Expenditures						X	D	D	D	D	D	D	D	D	D
Imports						X		D	D	D	D	D	D	D	D
Petroleum		X		X	X							D	D	D	D
Value added tax												D	D	D	D
Alcohol															
Excises												D	D	D	D
Gasoline via transport[a]												D	D	D	D
Tobacco												D	D	D	D
Wages															
Automobile															
Transport and gasoline[b]															
Gasoline[c]															

Note: D indicates that we reject the null in favor of dominance, that is, the item in the column dominates the item in the row. X indicates that we reject the null in favor of crossing. The elasticity with respect to household size (θ) is set at 0.5, which is near the lower end of scales found in the literature.

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.

b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.

c. Gasoline refers to the direct tax on gasoline.

Source: Calculated from Government of Madagascar (1994).

Table 6. *Gini Dominance Results for Taxes in Madagascar, 1994*
($\theta = 0.5$)

Variable	*45-degree line*	*Vanilla*	*Kerosene*	*Expenditures*	*Imports*	*Petroleum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Automobile*	*Transport and gasoline*[b]	*Gasoline*[c]
45-degree line			D	D	D	D	D	D	D	D	D	D	D	D	D
Vanilla			X						D	D	D	D	D	D	D
Kerosene		X		D	D	D	D	D	D	D	D	D	D	D	D
Expenditures							D	D	D	D	D	D	D	D	D
Imports							D	D	D	D	D	D	D	D	D
Petroleum												D	D	D	D
Value added tax												D	D	D	D
Alcohol												D	D	D	D
Excises												D	D	D	D
Gasoline via transport[a]												D	D	D	D
Tobacco												D	D	D	D
Wages															
Automobile															
Transport and gasoline[b]															
Gasoline[c]															

Note: *D* indicates that we reject the null in favor of dominance, that is, the item in the column dominates the item in the row. *X* indicates that we reject the null in favor of crossing. The elasticity with respect to household size (θ) is set at 0.5, which is near the lower end of scales found in the literature.

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.

b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.

c. Gasoline refers to the direct tax on gasoline.

Source: Calculated from Government of Madagascar (1994).

toward broad-based taxes, such as a VAT or income tax, on economic efficiency grounds. In industrial countries such taxes tend to be less progressive than other more specific taxes concentrated on luxury goods, because their breadth of coverage, which makes them less distortionary, also brings their distribution close to the overall income or expenditure distribution. In Madagascar, however, that is not the case. Both the VAT and the tax on wage and salary income are progressive. At first sight this result is comforting for those who favor a move to the traditional broad-based taxes. But it also reveals that these taxes are not nearly as broad-based as they are in an industrial economy, because the formal sector, to which they apply, is relatively small in Madagascar. Although the taxes are progressive, they may be more distortionary than is typically supposed.

Import duties are less progressive than the VAT, although not dramatically so, probably because the formal sector produces few goods that are not also imported, so the tax bases for the two are not as different as one might expect. This does not mean that the two can be substituted, however. Import duties still distort incentives, favoring a closed economy. Also, some of the benefits of import duties go to local firms, or more precisely, their owners, who are almost surely from the top end of the income distribution, whereas all of the benefits of a VAT go to the government, whose expenditures are probably more progressive.

The export duty on vanilla has the highest concentration curve, yet we cannot reject the null of equality between it and all other taxes or between it and the expenditure distribution. This indeterminate result is likely due to the small number of producers in our sample. Yet the striking shape of the concentration curve offers tentative support for reductions in export duties on equity grounds. Of course, to the extent that Madagascar enjoys substantial market power in the world market for vanilla, there may be an optimum export duty on vanilla that helps to keep world prices high by restricting Madagascar's supply to the market. But the loss of market share to other exporters, especially Indonesia, casts doubt on the degree of market power that Madagascar really has.

Finally, taxes with narrow bases are concentrated in three areas in Madagascar: petroleum duties, export duties, and "sin" taxes on alcohol and tobacco. The major rationale for these taxes is that they provide good tax handles, even if they are distortionary. There is also an argument in favor of using these taxes to discourage consumption of alcohol and tobacco for health reasons, and petroleum products for environmental reasons. Because taxes on alcohol and tobacco are progressive, it is difficult to criticize them, and there is not much policy interest in reducing them. The petroleum taxes are much more controversial, with the controversy focusing on their purported adverse effect on distribution. Our results suggest, however, that taxes on gasoline are highly progressive, even after including the indirect impact on public transport. For these products, then, the critics are mistaken, and the government should consider gasoline taxes as an attractive possibility for further revenue increases. The same cannot be said of kerosene, however. Used primarily for cooking and as a source of light in house-

holds that do not have electricity, a duty on kerosene is the one clearly regressive tax in Madagascar.

It is important to remember that progressivity is not the only measure of a good tax. Policymakers must also consider the tax's impact on economic efficiency (distortions to the allocation of resources), its administrative efficacy (whether it is a good tax handle), and, of course, the utility of corresponding government expenditures. Nevertheless, progressivity does matter, not least at a political level, where arguments that a tax hurts the poor are often more persuasive than considerations of economic and administrative efficiency. Our results should contribute to the debate over policy reform in Madagascar.

Appendix A. Calculation of Households' Tax Payments

This appendix describes our calculations of indirect and direct taxes in greater detail.

Indirect Taxes on Expenditure Items

Table A-1 lists the expenditure items included in the EPM survey and indicates which indirect taxes we have assumed are included in purchasing these items. To estimate the tax base for each tax, we first assume that:

- The VAT is levied on the c.i.f. value of imports plus all import duties paid.
- Import duties and tariffs are levied on the c.i.f. value of imports.
- Commodity-specific excises are levied on the c.i.f. value of imports (if the goods are imported) or the value of domestic sales.
- There is an untaxed retail markup on all expenditure items.

We then apply the rate shown in the table to the calculated base. We calculate the retail markup from a 1995 input-output table for Madagascar, using the ratio of retail and wholesale services to the sum of domestic sales plus imports. We then use the appropriate industry's ratio for each product in the expenditure survey.

Results for single-item taxes are not sensitive to errors in the percentages we use because the concentration curves are calculated as ratios. But for multi-item taxes (import duties and the VAT), errors across items with different tax rates could change the incidence calculations.

Other Indirect Taxes

The only other indirect tax we examine is the export duty on vanilla. We apply the duty rate (25 percent) to each household's sales of vanilla (in Madagascar francs) to estimate its tax payments for vanilla exports. This assumes that farmers pay only a part of the total duty, equal to their share in the total price of vanilla, that is, they share the tax incidence with middlemen who buy their vanilla output and sell it to the Vanilla Marketing Board. As with expenditure

Table A-1. *Assumed Taxes on Expenditure Items in Madagascar*
(percent)

Expenditure item	*Import duties*	*Value added tax*	*Others*
Adults' clothing	50	20	
Children's clothing	50	20	
Underwear	50	20	
Cloth for clothing	50	20	
Accessories	50	20	
Other clothing	50	20	
Sewing materials	50	20	
Adults' shoes	50	20	
Children's shoes	50	20	
Electricity	0	20	
Candles	0	20	
Water	0	20	
Kerosene[a]	0	0	FMG133 per liter
Natural gas	0	0	FMG50 per kilo
Candles	0	20	
Water	0	20	
Furniture	0	20	
Household accessories	0	20	
Household linen	0	20	
House furnishings	0	20	
Household appliances	40	20	
Kitchen appliances	40	20	
Cooking appliances	40	20	
Glassware	40	20	
Kitchen utensils	40	20	
Household utensils	40	20	
Home maintenance products	0	20	
Home maintenance tools	0	20	
Other home maintenance	0	20	
Sports and cultural events	0	20	
Hotels, vacations	0	20	
Radios and videocassette recorders	40	20	
Cameras	40	20	
Sports equipment	40	20	
Other durable equipment	40	20	
Books, magazines, and newspapers	0	20	
Leisure accessories	0	20	
Medicine	10	0	
Personal care articles	50	20	
Automobiles	50	50	15
Motorcycles	40	50	
Bicycles	20	20	
Gasoline and lubricants	0	0	FMG480 per liter
Transportation in cities	0	20	FMG20 of FMG480 per liter[b]
Intercity transportation	0	20	FMG20 of FMG480 per liter[b]
Mail and telecommunications	0	20	
Watches	40	20	

Table A-1. *(continued)*

Expenditure item	*Import duties*	*Value added tax*	*Others*
Jewelry	0	20	
Education and training fees	0	20	
All foods except those listed below	0	0	
Milled rice	30	0	
Rice flour	30	0	
Wheat[c]	20	0	
Other cereals	20	0	
Cheese	30	0	
Other dairy products	30	0	
Peanut oil	20	0	
Coconut oil	20	0	
Soybean oil	20	0	
Butter	20	0	
Margarine	20	0	
Lard	20	0	
Marinated or salted vegetables	40	20	
Other canned vegetables	40	20	
Jams and jellies	40	20	
Canned fruits	40	20	
Canned meats	40	20	
Canned fish	40	20	
Other canned food	40	20	
Condensed or powdered milk	40	20	
Baby food	0	20	
Fruit juice	50	20	
Syrup and soda	50	20	
Bottled water	50	20	
Meals in restaurants	0	20	
Rum	50	20	170
Beer	50	20	70
Wine and liquor	50	20	120
Cigarettes	50	25	60
Parakay (chewing tobacco)	0	15	
Chairs[d]	0	20	
Tables[d]	0	20	
Beds[d]	0	20	
Other furniture[d]	0	20	
Sewing machine[d]	40	20	
Gas store[d]	40	20	
Refrigerator[d]	40	20	
Television[d]	40	20	

a. There was no excise tax on kerosene in 1994. In order to say something about the incidence of the kerosene duty that came later, we have used the 1996 duty per liter, deflated by the proportion that the gasoline duty increased from 1994 to 1996.

b. For transport we assume that 20 percent of the cost is due to taxes on petroleum products.

c. Includes wheat in bread.

d. For durable items we use 10 percent of the value of the items owned, found in section 11, part B, rather than the expenditure information in section 8.

Source: Calculated from Government of Madagascar (1994).

taxes, this assumption does not affect the concentration curve for vanilla duties, but it probably underestimates the nominal amount that farmers pay.

Direct Taxes

The only direct tax included in this article is the income tax on wages. We have assumed that only workers who are employed by the public sector or formal enterprises pay income taxes on their wages and benefits (question 13, section 4, part B responses 1 or 2, and analogously for other jobs). We use the 1994 tax tables to be consistent with the nominal value of salaries earned in those years.

APPENDIX B. CALCULATION OF THE ESTIMATOR

This appendix presents the estimator that we use for the covariance matrix of the ordinates of two concentration curves that may be dependent. Davidson and Duclos (1997) develop the estimator, and our exposition depends almost entirely on their work.

Let X and Y be two jointly distributed random variables, and let F be the marginal distribution of Y. For our purposes we can think of Y as the variable that measures household welfare (such as per capita income) and X as a tax. Let $p = [p(1), p(2), \ldots p(k)]$ be a vector of abscissa on the x-axis of a Lorenz or concentration curve, and define $\gamma_{p(j)}$ as the expected value of X given that Y is in the lower $p(j)$ quantile of its distribution. Then, an estimator of the ordinate for a concentration curve at $p(j)$ is $p(j)[\gamma_{p(j)} / \gamma_1]$, where $\gamma_{p(j)} = E[X \mid F(Y) \leq p(j)]$, that is, the expected value of X (the tax) conditional on the household being found in the lowest $p(j)$ quantile of the income distribution. Note that γ_1 is just the mean value of the tax for all households. If we repeat the same argument for another tax, say W, and another welfare variable, Z, and we define $\delta_{p(j)} = E[W \mid F(Z) \leq p(j)]$, then we have $p(j)[\delta_{p(j)} / \delta_1]$ as an estimator for the ordinate of W's concentration curve, and $\lambda_j = p(j)[(\gamma_{p(j)} / \gamma_1) - (\delta_{p(j)} / \delta_1)]$ is the difference between the two at abscissa $p(j)$. For our work Y and Z are always the same variable, such as per capita income.

Both the standard errors for each ordinate and the difference between them depend on the joint distribution of:

$$\text{(B-1)} \quad \omega = [\, p(1)\gamma_{p(1)}(1), \ldots, p(k)\gamma_{p(k)}(k), \gamma_1, p(1)\delta_{p(1)}(1), \ldots, p(k)\delta_{p(k)}(k), \delta_1]^T.$$

Using Gaussian kernel estimates (see, for example, Silverman 1986, ch. 3) for the conditional means $\gamma_{p(k)}$ and $\delta_{p(k)}$, Davidson and Duclos prove that $N^{-0.5}(\hat{\omega} - \omega)$ is asymptotically normal with mean zero and an asymptotic covariance matrix that can be estimated without knowledge of the population distribution. We have used the same estimators here.

Finally, Davidson and Duclos note that, by a result in Rao (1973: 388–89), we can generate the covariance matrix for λ by pre- and post-multiplying ω's covariance matrix with the Jacobian for λ with respect to the vectors γ and δ. Formally,

let $\theta = [\gamma_{p(1)}, \ldots, \gamma_1, \delta_{p(1)}, \ldots, \delta_1]^T$, and let $\frac{\partial \lambda_i}{\partial \theta_j} = [S(\gamma) \mid S(\delta)]$ be the Jacobian for ω's covariance matrix, where

$$S(\gamma) = \begin{pmatrix} \frac{1}{\gamma_1} & -p(1)\frac{\gamma_{p(1)}}{\gamma_1^2} \\ \vdots & \vdots \\ \frac{1}{\gamma_1} & -p(k)\frac{\gamma_{p(k)}}{\gamma_1^2} \end{pmatrix} \tag{B-2}$$

and $S(\delta)$ is defined similarly. Then the k by k matrix $\left(\frac{\partial \lambda_i}{\partial \theta_j}\right)\omega\left(\frac{\partial \lambda_i}{\partial \theta_j}\right)^T$ is the covariance matrix for λ.

Appendix C. Ordinate Estimates and Standard Errors for Tax Concentration Curves in Madagascar, 1994

This appendix presents estimates of the ordinates and standard errors of the concentration curves used in this paper.

(Tables begin on the following page)

Table C-1. *Ordinate Estimates*

Ordinate	*45-degree line*	*Vanilla*	*Kero-sene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Auto*	*Transport and gasoline*[b]	*Gasoline*[c]
0.05	0.050	0.000	0.018	0.007	0.006	0.012	0.006	0.001	0.002	0.001	0.002	0.000	0.000	0.000	0.000
0.10	0.100	0.032	0.037	0.019	0.018	0.025	0.014	0.009	0.004	0.006	0.003	0.000	0.000	0.000	0.000
0.15	0.150	0.054	0.059	0.034	0.030	0.040	0.024	0.016	0.008	0.010	0.006	0.000	0.000	0.000	0.000
0.20	0.200	0.071	0.085	0.052	0.046	0.056	0.035	0.022	0.016	0.016	0.011	0.000	0.000	0.000	0.000
0.25	0.250	0.085	0.123	0.072	0.066	0.082	0.049	0.032	0.026	0.023	0.017	0.003	0.000	0.000	0.000
0.30	0.300	0.151	0.156	0.095	0.085	0.104	0.065	0.041	0.034	0.030	0.022	0.003	0.000	0.000	0.000
0.35	0.350	0.155	0.194	0.119	0.109	0.129	0.083	0.053	0.042	0.039	0.030	0.003	0.015	0.000	0.000
0.40	0.400	0.164	0.231	0.146	0.135	0.154	0.101	0.077	0.057	0.056	0.040	0.006	0.015	0.000	0.000
0.45	0.450	0.234	0.271	0.176	0.163	0.181	0.122	0.094	0.073	0.071	0.054	0.016	0.015	0.000	0.000
0.50	0.500	0.353	0.314	0.208	0.193	0.210	0.145	0.115	0.090	0.087	0.067	0.020	0.015	0.000	0.000
0.55	0.550	0.415	0.360	0.242	0.224	0.240	0.171	0.129	0.112	0.107	0.091	0.024	0.022	0.002	0.001
0.60	0.600	0.502	0.406	0.280	0.259	0.271	0.200	0.147	0.134	0.129	0.115	0.030	0.022	0.002	0.001
0.65	0.650	0.594	0.455	0.323	0.301	0.305	0.232	0.173	0.175	0.157	0.146	0.038	0.022	0.003	0.001
0.70	0.700	0.713	0.504	0.369	0.350	0.340	0.265	0.210	0.205	0.193	0.180	0.055	0.022	0.014	0.007
0.75	0.750	0.842	0.567	0.422	0.400	0.382	0.302	0.248	0.240	0.231	0.219	0.071	0.025	0.014	0.007
0.80	0.800	0.854	0.625	0.481	0.463	0.421	0.350	0.305	0.297	0.289	0.277	0.103	0.059	0.014	0.007
0.85	0.850	0.906	0.686	0.550	0.538	0.465	0.409	0.361	0.353	0.357	0.355	0.163	0.114	0.025	0.014
0.90	0.900	0.910	0.749	0.635	0.627	0.512	0.495	0.459	0.492	0.455	0.452	0.235	0.202	0.044	0.022
0.95	0.950	0.936	0.838	0.747	0.741	0.602	0.611	0.599	0.629	0.597	0.596	0.376	0.295	0.062	0.085
1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.
b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.
c. Gasoline refers to the direct tax on gasoline.
Source: Calculated from Government of Madagascar (1994).

Table C-2. *Standard Errors*

Ordinate	*45-degree line*	*Vanilla*	*Kero-sene*	*Expen-ditures*	*Imports*	*Petro-leum*	*Value added tax*	*Alcohol*	*Gasoline via transport*[a]	*Excises*	*Tobacco*	*Wages*	*Auto*	*Transport and gasoline*[b]	*Gasoline*[c]
0.05	0.000	0.001	0.002	0.000	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000
0.10	0.000	0.017	0.003	0.001	0.001	0.004	0.001	0.003	0.001	0.001	0.001	0.000	0.000	0.000	0.000
0.15	0.000	0.020	0.004	0.001	0.001	0.006	0.002	0.003	0.001	0.002	0.001	0.000	0.000	0.000	0.000
0.20	0.000	0.023	0.005	0.001	0.002	0.008	0.002	0.004	0.002	0.002	0.002	0.000	0.000	0.000	0.000
0.25	0.000	0.024	0.006	0.002	0.002	0.012	0.003	0.005	0.003	0.003	0.003	0.001	0.000	0.000	0.000
0.30	0.000	0.035	0.007	0.002	0.003	0.015	0.004	0.006	0.004	0.003	0.003	0.001	0.000	0.000	0.000
0.35	0.000	0.035	0.009	0.003	0.003	0.018	0.005	0.007	0.005	0.004	0.004	0.001	0.008	0.000	0.000
0.40	0.000	0.035	0.010	0.004	0.004	0.022	0.006	0.010	0.006	0.006	0.005	0.003	0.008	0.000	0.000
0.45	0.000	0.042	0.012	0.004	0.005	0.025	0.007	0.012	0.008	0.006	0.006	0.007	0.008	0.000	0.000
0.50	0.000	0.052	0.013	0.005	0.005	0.029	0.008	0.014	0.009	0.008	0.007	0.007	0.008	0.000	0.000
0.55	0.000	0.057	0.015	0.006	0.006	0.034	0.009	0.015	0.011	0.009	0.009	0.007	0.011	0.002	0.001
0.60	0.000	0.061	0.016	0.006	0.007	0.038	0.011	0.017	0.012	0.010	0.011	0.007	0.011	0.002	0.001
0.65	0.000	0.064	0.018	0.007	0.008	0.043	0.012	0.019	0.016	0.012	0.013	0.008	0.011	0.002	0.001
0.70	0.000	0.060	0.020	0.008	0.009	0.048	0.014	0.024	0.018	0.014	0.015	0.010	0.011	0.008	0.004
0.75	0.000	0.055	0.022	0.009	0.010	0.053	0.016	0.027	0.021	0.017	0.018	0.011	0.012	0.008	0.004
0.80	0.000	0.054	0.024	0.010	0.011	0.059	0.018	0.032	0.026	0.020	0.021	0.015	0.029	0.008	0.004
0.85	0.000	0.047	0.026	0.012	0.012	0.065	0.021	0.038	0.030	0.024	0.025	0.023	0.044	0.015	0.009
0.90	0.000	0.047	0.028	0.013	0.013	0.071	0.024	0.046	0.040	0.029	0.029	0.029	0.067	0.023	0.012
0.95	0.000	0.043	0.030	0.014	0.014	0.083	0.029	0.056	0.049	0.034	0.034	0.041	0.094	0.031	0.047
1.00	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

a. Gasoline via transport refers to the part of the direct tax on gasoline that falls on users of public transport.
b. Transport and gasoline refers to the combined impact of the direct tax on gasoline and the indirect tax on users of public transport.
c. Gasoline refers to the direct tax on gasoline.
Source: Calculated from Government of Madagascar (1994).

REFERENCES

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

Blundell, Richard, and Arthur Lewbel. 1991. "The Information Content of Equivalence Scales." *Journal of Econometrics* 50(1):49–68.

Browning, Martin. 1992. "Children and Household Economic Behavior." *Journal of Economic Literature* 30(3):1435–75.

Cornia, G. A., Richard Jolly, and Frances Stewart. 1987. *Adjustment with a Human Face* 2. Oxford: Clarendon Press.

Coulter, Fiona, Frank Cowell, and Stephen Jenkins. 1992. "Differences in Needs and Assessment of Income Distributions." *Bulletin of Economic Research* 44(2):77–124.

Cutler, David M., and Lawrence F. Katz. 1992. "Rising Inequality? Changes in Distribution of Income and Consumption in the 1980s." *American Economic Review Papers and Proceedings* 82(May):546–51.

Davidson, Russell, and Jean-Yves Duclos. 1997. "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers." *Econometrica* 65(6):1453–66.

Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior.* Cambridge, U.K.: Cambridge University Press.

———. 1986. "On Measuring Child Costs: With Applications to Poor Countries." *Journal of Political Economy* 94(4):720–44.

Dorosh, Paul, René Bernier, and Alexander Sarris. 1990. *Macroeconomic Adjustment and the Poor: The Case of Madagascar.* Monograph 9. Ithaca, N.Y.: Cornell University, Cornell Food and Nutrition Policy Program.

Gouveia, Miguel, and José Tavares. 1995. "The Distribution of Household Income and Expenditure in Portugal, 1980 and 1990." *Review of Income and Wealth* 41(1):1–17.

Government of Madagascar. 1994. *Enquête Permanente auprès des Ménages 1994.* Antananarivo, Madagascar.

Howes, Stephen. 1996. "The Influence of Aggregation on the Ordering of Distributions." *Economica* 63(250):253–72.

INSTAT (Institute National de la Statistique). 1995. "1995 Madagascar SAM." Antananarivo, Madagascar. Processed.

Krueger, Anne O., Maurice Schiff, and Alberto Valdes. 1988. "Agricultural Incentives in Developing Countries: Measuring the Effect of Sectoral and Economywide Policies." *The World Bank Economic Review* 2(3):255–71.

Lanjouw, Peter, and Martin Ravallion. 1995. "Poverty and Household Size." *Economic Journal* 105(433):1415–34.

Rao, R. C. 1973. *Linear Statistical Inference and Its Applications.* New York: John Wiley and Sons.

Sahn, David E., Paul A. Dorosh, and Stephen D. Younger. 1996. "Exchange Rate, Fiscal, and Agricultural Policies in Africa: Does Adjustment Hurt the Poor?" *World Development* 26(4):719–47.

Sahn, David E., Stephen D. Younger, and Kenneth R. Simler. 1997. "Dominance Testing of Transfers in Romania." Food and Nutrition Policy Program, Cornell University, Ithaca, N.Y. Processed.

Selden, Thomas, and Michael J. Wasylenko. 1992. "Benefit Incidence Analysis in Developing Countries." Policy Research Working Paper 1015. World Bank, Policy Research Department, Washington, D.C. Processed.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis.* New York: Chapman and Hall.

World Bank. 1984. "Madagascar: Current Economic Situation and Prospects." Report 5154-MAG. Washington, D.C. Processed.

———. 1986. "The Democratic Republic of Madagascar: Country Economic Memorandum." Report 5996-MAG. Washington, D.C. Processed.

———. 1991. "Madagascar: Beyond Stabilization to Sustainable Growth." Report 9101-MAG. Washington, D.C. Processed.

Yitzhaki, Shlomo. 1983. "On an Extension of the Gini Inequality Index." *International Economic Review* 24(3):617–28.

Yitzhaki, Shlomo, and Jeffrey Lewis. 1996. "Guidelines on Searching for a Dalton-Improving Tax Reform: An Illustration with Data from Indonesia." *The World Bank Economic Review* 10(3):541–62.

Yitzhaki, Shlomo, and Joel Slemrod. 1991. "Welfare Dominance: An Application to Commodity Taxation." *American Economic Review* 81(3):480–96.

Younger, Stephen. 1996. "Estimating Tax Incidence in Ghana: An Exercise Using Household Data." In David E. Sahn, ed., *Economic Reform and the Poor in Africa.* Oxford: Clarendon Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2. 333-56

Genuine Savings Rates in Developing Countries

Kirk Hamilton and Michael Clemens

Augmented measures of savings and wealth in the national accounts are critical to conceptualizing and achieving sustainable development. After developing the theory of genuine savings—traditional net savings less the value of resource depletion and environmental degradation plus the value of investment in human capital—this article presents empirical estimates for developing countries. These calculations account for resource depletion and carbon dioxide emissions, using consistent time series data for 1970–93. The empirical evidence shows that levels of genuine savings are negative in a wide range of countries, particularly in Sub-Saharan Africa, and that these countries are being progressively impoverished. Increasing the coverage of natural resources and pollutants in our calculations would reduce the estimated levels of genuine savings overall. The use of genuine savings measures suggests a series of policy questions that are key to sustaining development. These are also explored, specifically the extent to which monetary and fiscal policies, exports of exhaustible resources, stronger resource policies, and pollution abatement measures boost genuine savings rates. For policymakers, linking sustainable development to genuine savings rates means that there are many possible interventions to increase sustainability, from the macroeconomic to the purely environmental.

However defined, achieving sustainable development necessarily entails creating and maintaining wealth. Given the centrality of savings and investment in economic theory, it is surprising that the effects of depleting natural resources and degrading the environment have not, until recently, been considered in measurements of national savings. Augmented measures of savings and wealth in the national accounts are critical to conceptualizing and achieving sustainable development, which was a prime motivation for publishing *Expanding the Measure of Wealth* (World Bank 1997).

Valuing depletion and degradation within a national accounting framework is an increasingly viable proposition, both as a result of the significant progress made in techniques of valuing environmental resources (for a recent example, see Freeman 1994) and as a result of the expanding foundation that theoretical developments are placing under the methods of "green" national accounting (Weitzman 1976, Hartwick 1990, Mäler 1991, and Hamilton 1994, 1996). The first cross-country application of these greener accounting methods to the measurement of

Kirk Hamilton is with the Environment Department at the World Bank, and Michael Clemens is with the Department of Economics at Harvard University. The authors gratefully acknowledge the comments of Danny McCoy, David Pearce, and three reviewers.

net savings appears in Pearce and Atkinson (1993), who combine published estimates of depletion and degradation with standard national accounting data to calculate true savings for 20 countries. According to this measure many countries appear to be on unsustainable paths because their gross savings are less than the sum of conventional capital depreciation and natural resource depletion.

Enlarging the concept of net savings to include the depletion of natural resources is a reasonable way to extend traditional savings concepts. The depletion of a natural resource is, in effect, the liquidation of an asset and therefore should not appear as a positive contribution to net income or net savings. Although minor technical issues remain, the methods of valuing the discovery, depletion, and growth of commercial natural resources in the context of the System of National Accounts (SNA) are now well developed (Hamilton 1994 and Hill and Harrison 1994).

More problematic is the valuation of environmental degradation. While United Nations guidelines for environmental accounting favor valuation according to the cost of restoring the environment to its state at the beginning of the accounting period—the maintenance-cost method—theoretical approaches suggest that the marginal social costs of pollution are a more correct basis for valuing emissions into the environment (United Nations 1993 and Hamilton 1996).

After developing the theory of genuine savings, this article presents empirical estimates for developing countries. These calculations account for resource depletion and carbon dioxide emissions using consistent time series data for 1970–93. The article concludes with a discussion of the policy issues raised by greener national accounting.

I. Genuine Savings: A Formal Model

The notion of genuine savings is presented briefly and informally in Hamilton (1994) and Pearce, Hamilton, and Atkinson (1996). This section provides a more rigorous development using a model that, although extremely simple, identifies the adjustments to savings measures that must be made to account for natural resources, pollutants, and human capital. Sustainable development is defined to be, as in Pezzey (1989), nondeclining utility.

We assume a simple closed economy with a single resource used as an input to the production of a composite good that may be consumed, invested in produced assets or human capital, or used to abate pollution, so that $F(K,R,N) = C + \dot{K} + a + m$, where K is physical capital, R is resource use, N is human capital, C is consumption, $\dot{K}$ is investment in physical capital, a is pollution abatement expenditures, and m is investment in human capital (current education expenditures). The function $q(m)$ transforms education expenditures into human capital that does not depreciate (and can be considered as a form of disembodied knowledge), so that $\dot{N} = q(m)$.[1] Labor is fixed and therefore factored out of the production function.

1. Investing in human capital is a type of endogenous technical progress. See Weitzman and Löfgren (1997), who deal with exogenous technical change.

Pollution emissions are a function of production and abatement, $e = e(F, a)$, and pollutants accumulate in a stock X such that $\dot{X} = e - d(X)$, where d is the quantity of natural dissipation of the pollution stock. The flow of environmental services B is negatively related to the size of the pollution stock, so that $B = \alpha(X)$, $\alpha_X < 0$. Resource stocks S grow by an amount g, are depleted by extraction R so that $\dot{S} = g(S) - R$, and are assumed to be costless to produce. The utility of consumers is assumed to be a function of consumption and environmental services, $U = U(C,B)$. There is a fixed pure rate of time preference r.

Following Hartwick (1990), new green national accounting aggregates are defined on the basis of an intertemporal optimization problem. Wealth, W, is defined to be the present value of utility on the optimal path. It is assumed that a social planner wishes to maximize wealth as follows:

$$\begin{aligned} \max W &= \int_t^\infty U(C,B)e^{-rs}ds \quad \text{subject to:} \\ \dot{K} &= F - C - a - m \\ \dot{X} &= e - d \\ \dot{S} &= -R + g \\ \dot{N} &= q(m). \end{aligned}$$

The current-value Hamiltonian function, which is maximized at each point in time, is given by,

$$H = U + \gamma_K\dot{K} + \gamma_X\dot{X} + \gamma_S\dot{S} + \gamma_N\dot{N} \tag{1}$$

where γ_K, γ_X, γ_S, and γ_N are, respectively, the shadow prices (in utils) of capital, pollution, resources, and human capital. Deriving the first-order conditions for a maximum, the Hamiltonian function may be written as

$$H = U(C,B) + U_C[\dot{K} - (1 - be_F)F_R(R - g) - b(e - d) + q/q']$$

where b is the marginal cost of pollution abatement. These Hamiltonian equalities hold because we assume the economy is on an optimal trajectory. Hamilton (1996) shows that b is precisely equal to the marginal social cost of pollution emissions and is therefore also equal to the level of the Pigovian tax on emissions required to maximize welfare. The term be_F can thus be interpreted as the effective tax rate on production as a result of the emissions tax. So, although we started with an optimal growth problem, the prices that result are those that would prevail in a competitive economy with a Pigovian tax on pollution. Note as well that $1/q'$ is the marginal cost of creating a unit of human capital.

Since $\dot{S} = g - R$, $\dot{X} = e - d$, and $\dot{N} = q$, the bracketed expression in the second term of the right side of the Hamiltonian is equal to the change in the real value of all assets in this simple economy, where human capital is valued at its marginal creation cost, pollution stocks are valued at marginal abatement costs, and natural resources are valued at the resource rental rate, F_R, net of the effective tax

rate on production associated with pollution emissions. Therefore the bracketed expression also serves to define genuine savings, G,

$$G \equiv \dot{K} - (1 - be_F)F_R(R - g) - b(e - d) + q/q'. \tag{2}$$

For nonliving resources g is equal to zero, and for "pure" cumulative pollutants d is also zero.

Genuine savings consists therefore of investment in produced assets and human capital, less the value of depletion of natural resources and the value of accumulated pollutants. As shown in Hamilton (1997),

$$U_C G = \dot{W} = rW - U. \tag{3}$$

Expression 3 implies the following: negative genuine savings at a point in time means that future utility must be less than current utility over some period on the optimal path. In other words, negative genuine savings serves as an indicator of nonsustainability. This expression also implies that Hicksian income, the maximum amount of produced output that could be consumed while leaving total wealth instantaneously constant, is given by

$$NNP = C + \dot{K} - (1 - be_F)F_R(R - g) - b(e - d) + q/q'. \tag{4}$$

Asheim (1994) points out that prices observed in the marketplace will generally differ from those required to support a sustainable (constant-utility) path. However, Asheim is working with a model in which the optimal path is not sustainable. He uses the model of Dasgupta and Heal (1979), in which fixed technology, exhaustible resources, and a fixed pure rate of time preference lead to an economy where welfare declines asymptotically to zero along the optimal path. For this economy the maximal sustainable path is supported by the Hartwick rule (Hartwick 1977): genuine savings are set to be identically zero at each point in time, and as a result utility is constant. Within the confines of the fixed-technology model, Asheim's point stands.

For a suitable specification of the human capital accumulation function, the model used in this article is one in which the optimal path is sustainable. Even if resources are exhaustible and pollution stocks do not dissipate, the endogenous technical progress inherent in human capital formation will yield an optimal path along which utility rises asymptotically to a positive maximum value. The interesting question for this model concerns divergences between observed "real world" prices and theoretical shadow prices on the optimal path. Hamilton, Atkinson, and Pearce (1998) argue that policy distortions in a typical economy lead to overextraction of natural resources and excess pollution emissions. Under these conditions it can be shown that real world resource rents exceed their optimal levels, as do marginal pollution damages. More efficient resource and environmental policies will reduce this bias and also increase genuine savings.

The current model can easily be extended to include foreign trade and depreciation of produced assets. If produced capital depreciates at a percentage rate equal to δ, then the accounting identity for these assets becomes

$$\dot{K} = F - C - a - m - \delta K$$

where $\dot{K}$ now measures *net* investment.

Turning to foreign trade, net foreign assets, A, accumulate as a result of exports, E, and decumulate with imports, M. For a fixed international rate of return, i, therefore, the asset accounting identity is

$$\dot{A} = iA + E - M.$$

Under these assumptions the measure of Hicksian income, NNP, for an open economy is given by

$$NNP = C + \dot{K} + E - M + iA - (1 - be_F)F_R(R - g) - b(e - d) + q/q'.$$

The first six terms in this expression are precisely the standard measure of NNP. An expanded conception of the asset base implies that the standard NNP should be adjusted by deducting net depletion of natural resources and the marginal damages from net accumulation of pollution and by adding investments in human capital.

Vincent, Panayotou, and Hartwick (1997) show that for a small resource-exporting economy, taking international prices as given, NNP should include the present value of future capital gains on resource exports. Hamilton (1997) argues that given the long-term flat-to-declining trend in real resource prices (at least for subsoil resources), this potential annuity can be considered to be zero. It was precisely the unwarranted presumption that prices of resource exports would continue to rise that led many natural resource exporters into difficulty in the 1980s.

The treatment of education and pollution abatement expenditures requires more elaboration. Hamilton (1994) essentially argues that current education expenditures are not consumption and therefore should be included in savings. Defining net marginal resource rents as $n \equiv (1 - be_F)F_R$, NNP can be defined as

$$NNP = GNP - \delta K - a - n(R - g) - b(e - d) + q/q' - m \tag{5}$$

where GNP is gross national product.

But $$q/q' - m = \left(\frac{q}{mq'} - 1\right)m = \left(\frac{1/q'}{m/q} - 1\right)m.$$

Since $1/q'$ is the *marginal* cost of creating a unit of human capital, and m/q is the *average* cost, expression 5 suggests that the value of investments in human capital should be greater than current education expenditures (under the usual assumptions about convexity) and that these current expenditures can therefore serve as a lower-bound estimate of the investment in human capital.

Expression 5 says that pollution abatement expenditures, a, are essentially intermediate in character and should be deducted in measuring genuine savings. In practice, most current abatement expenditures are already treated as intermediate inputs in standard national accounting.

Finally, it is important to present the formula for calculating genuine savings from real data. For produced asset depreciation δK, net marginal resource rental rate n, and marginal social cost of pollution σ, this is given by,

$$G = GNP - C - \delta K - n(R - g) - \sigma(e - d) + m. \tag{6}$$

Here, $GNP - C$ is traditional gross savings, which includes foreign savings, while $GNP - C - \delta K$ is traditional net savings. Since carbon dioxide is the only pollutant considered in what follows, be_F can be assumed to be close to zero, so the adjustment to net resource rents $(1 - be_F)$ can safely be assumed to be near 1, while dissipation, d, is assumed to be small relative to emissions, e.[2]

Net natural growth of living resources $(R - g)$ is not added to genuine savings when it is positive, but net depletion (that is, when $R > g$) is deducted. Although this will bias the results against sustainability, Vitousek and others (1986) estimate that less than 33 percent of standing forests are merchantable. Empirical examination of regions where growth exceeds harvest reveals a number of heavily forested countries (including Bolivia, Central African Republic, Republic of Congo, and Guyana), where valuing net growth at current unit rents would equal 20–50 percent of GNP. A clear dichotomy is at work: all of the timber in net depletion countries is merchantable, by definition, while probably less than one-third is merchantable in net growth countries. It is likely that mechanically adding net forest growth to GNP and savings would implicitly include the growth of many uneconomic trees (those with zero rental value).

II. Measuring Resource Depletion and Environmental Degradation

Building on the theory of green national accounting, this article provides a first set of calculations of genuine savings from a consistently derived and reasonably comprehensive time series data set on resource depletion and carbon dioxide emissions. Previous studies, such as Repetto and others (1989), Sadoff (1992), and Kellenberg (1995), deal with particular countries in depth. The calculations presented here necessarily trade off some amount of accuracy against wider coverage.

Data availability limits the adjustments to savings measures to the following: valuing resource rents for nonrenewable resources, valuing depletion of forests beyond replacement levels, and valuing the marginal social costs of carbon dioxide (CO_2) emissions.

2. Carbon dioxide has an atmospheric residency time of 200 years, or a dissipation rate of roughly 0.5 percent a year. For an average growth rate in emissions of 5 percent a year, therefore, the equilibrium ratio of d to e is 1/11.

The basic approach to calculating resource rents for nonrenewable resources is to subtract country- or region-specific average costs of extraction from the world price for the resource in question, all expressed in current U.S. dollars. Many world prices were derived from World Bank (1993): where multiple markets—for example, London and New York—are reported, a simple average of these market prices serves as the world price. So, for minerals the total resource rents are calculated as the world price minus mining costs minus milling and beneficiation costs minus smelting costs minus transport to port minus "normal" return to capital. Cost data are derived from U.S. Bureau of Mines (1987).

For crude oil, unit rents are calculated as the world price minus lifting costs. These lifting costs were estimated based on data from the Inter-American Development Bank (IDB 1981), International Energy Agency (IEA 1994b, 1995c, 1995d, 1996), Jenkins (1989), Sagers, Kryukov, and Shmat (1995), and Smith (1992).

Natural gas, although its international trade has soared in recent years, cannot yet be said to possess a single world price. A world price was estimated by averaging free-on-board prices from several points of export worldwide, after which the unit rents were calculated like those for oil. Production costs were taken from Adelman (1991), Cornot-Gandolphe (1994), IEA (1995d), Julius and Mashayekhi (1990), Khan (1986), Liefert (1988), Mashayekhi (1983), and Meyer (1994).

For hard coal a world price was calculated by combining data on steam and coking coals after adjusting for differences in heat content and quality. A world price for lignite was obtained by analyzing national-level differences in prices between hard coal and lignite in various countries and estimating a similar proportion of values to hold true with respect to the world price for hard coal. Unit rents for both hard coal and lignite were then calculated as for oil. Coal production costs were taken from Bhattacharya (1995), Doyle (1987), IEA (1994a, 1995b, 1995d, 1995e,), Tretyakova and Heinemeier (1986), and World Bank data.

For forest resources only rent on the portion of wood production that exceeds the country's mean annual increment in commercial wood mass was subtracted from savings. A price for the wood of each country was calculated based on the proportions of fuelwood, coniferous softwood, nonconiferous softwood, and tropical hardwood found in total annual production. Representative world prices were used for each type of wood, and a price for fuelwood was estimated using World Bank data. Unit rents were calculated by subtracting average unit harvest costs from the world price.

There are several further points to note about this methodology:

- From a theoretical viewpoint depletion estimates depend on scarcity rents, which should be measured as price minus *marginal* cost of extraction (including a normal return to capital). In practice, data on marginal production costs are almost never available, and practitioners (as evidenced by the green national accounting literature) fall back on using average extraction costs. This tends to overstate calculated resource rents and hence to understate genuine savings.

- Countries may or may not be selling their natural resources for internal consumption at world market prices, although they have good incentives to do so. Moreover, the use of uniform world prices tends to overstate rents for countries with lower-grade resources.
- Extraction costs are measured at a fixed point in time, which differs from country to country and resource to resource according to the availability of data. Extraction costs are held constant (in real terms) during 1970–94. World prices vary over time, leading to corresponding variations in calculated rental rates.
- If the extraction cost data are region- rather than country-specific, the regional cost structure is applied to all of the producing countries in the region.
- Rents on minerals are generally viewed as accruing to the resource owner for the production of the crude form of the material in question, typically an ore. In practice, most mineral operations are vertically integrated to a considerable extent, and the only price and cost data are for refined forms of the materials. Measuring resource rents as described above for these vertically integrated mineral operations therefore implicitly ascribes to the resource rent any excess returns to capital for the milling and refining stages.

Table 1 presents the calculated average rental rates for several resources. The table also shows which cost components, subject to data availability, went into the calculation of rental rates. In most (but not all) cases an explicit rate of return on capital appears as a cost component. Missing cost components lead, of course, to overestimates of resource rents. In line with the formal green national accounting methods of the preceding section, the country-specific unit resource rents in each year are multiplied by the quantities of resource extraction for each of the resources in table 1 to arrive at the total value of resource depletion.

For tropical forest resources, valuing depletion is much more complicated. Where deforestation is occurring, the issue is essentially one of land use, with standing forests being one use among many for a particular land area. This suggests that the correct way to value deforestation is to measure the change in land value (which should represent the present value of the net returns under the chosen use for land); this is essentially the result in Hartwick (1992). The formal model suggests that, where deforestation is not occurring but harvest exceeds growth, it is the net depletion of the resource that should be valued.

Because data on the value of forested land before and after clearance are not widely available, deforestation is not treated explicitly, and forest depletion is simply valued as the stumpage value (price minus average logging cost) of the volume of commercial timber and fuelwood harvested in excess of natural growth in commercially valuable wood mass for that year. Harvest rates by country are as given in FAO (1994). The annual increment is estimated using World Bank data and Duvigneaud (1971), Lamprecht (1989), FAO/UNECE (1992), and Kanowski and others (1992). Stumpage rates come from World Bank data, Openshaw and

Table 1. *Rental Rates for Natural Resources*
(share of world price)

Natural resource	*Mean*	*Standard deviation*	*Cost components*
Bauxite	0.61	0.13	Mining, milling
Copper	0.49	0.20	Mining, milling, smelting, 15 percent return on capital
Crude oil	0.65	0.26	Production costs
Forestry	0.45	0.13	Logging and transport
Gold	0.29	0.15	Mining, milling, transport, 15 percent return on capital
Hard coal	0.28	0.17	Mining and transport
Iron ore	0.58	0.24	Mining, beneficiation, transport
Lead	0.23	0.12	Mining, milling, smelting, transport, 15 percent return on capital
Lignite	0.38	0.17	Mining and transport
Natural gas	0.59	0.24	Production costs
Nickel	0.35	0.21	Mining, milling, smelting
Phosphate rock	0.26	0.14	Mining, milling, transport, 15 percent return on capital
Silver	0.31	0.22	Mining, milling, transport, 15 percent return on capital
Tin	0.35	0.16	Mining, milling, 15 percent return on capital
Zinc	0.21	0.13	Mining, milling, smelting, transport, 15 percent return on capital

Note: Values are for unweighted pooled data from 1985–94 (except silver and tin, 1975–94), excluding negative values.
Source: Authors' estimates.

Feinstein (1989), Kellenberg (1995), and others; market prices are from FAO (1983, 1995), Openshaw and Feinstein (1989), van Buren (1990), Barnes (1992), and World Bank (1993).

The foregoing description of the valuation of forest depletion suggests that the calculations are quite rough. It should also be obvious that the values calculated pertain only to commercial exploitation, so that the values of biodiversity, carbon sequestration, and other uses are not captured.

Pollution damages can enter green national accounts in different ways. Although damage to produced assets (the damage to building materials caused by acid rain, for example) is in principle included in depreciation figures, in practice most statistical systems are not detailed enough to pick this up. The effects of pollution on output (damaged crops, lost production owing to morbidity) are usually not broken out explicitly, but because they are reflected implicitly in the standard national accounts, there is no need to adjust savings measures in this regard.[3] Traditionally, the key pollution adjustment is for welfare effects, valuing the willingness-to-pay to avoid excess mortality and the pain and suffering from pollution-linked morbidity. Because these marginal damage figures are locale-specific, no general treatment of pollution emissions is attempted in the following discussion.

3. However, if the productive capacity of an asset, such as soil fertility, is damaged by pollution, then the loss in asset value should be deducted from savings.

The only social costs considered here, therefore, are for carbon dioxide. Data are readily available for this pollutant, and damages are global rather than strictly local. The basic emissions data employed are from the Carbon Dioxide Information and Analysis Center (CDIAC 1994), covering fossil fuel combustion and cement manufacture. The global marginal social cost of a metric ton of carbon emitted is assumed to be $20 in 1990, taken from Fankhauser (1994). Global damages are charged to emitting countries on the assumption that the property right to a clean environment lies with the pollutee—for example, we are assuming that the Comoros Islands have the right not be inundated as a result of CO_2 emissions elsewhere.

A key element missing in the calculations is any valuation of soil erosion, owing to the lack of comprehensive data sets on either physical erosion or its value. This is an important gap considering the significance of agriculture in most developing countries—erosion is considered to be a major problem in Sub-Saharan African countries in particular. A second missing element is fish stocks, where data problems, questions of ownership, and near-zero rental values resulting from overfishing all militate against including values of depletion.

III. Empirical Estimates of Genuine Savings

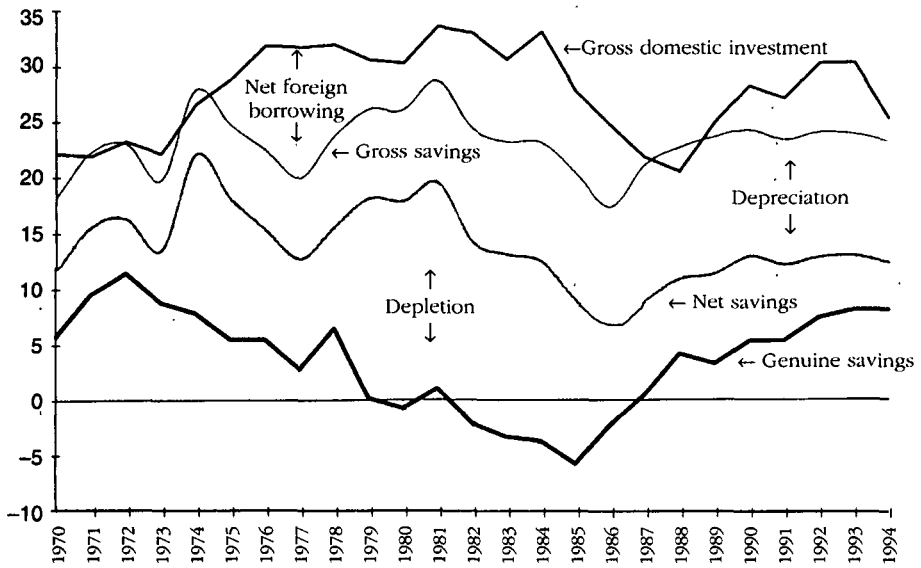
The traditional measure of a nation's rate of accumulation of wealth, as reported in the World Bank's *World Development Indicators,* for instance, is *gross* savings (World Bank, various years). This is calculated as a residual: GNP minus public and private consumption. *Net* savings, gross savings less the value of depreciation of produced assets, is a first step toward a sustainability indicator. Measures of *genuine* savings address a much broader conception of sustainability than net savings by valuing changes in the natural resource base and environmental quality in addition to produced assets. Figure 1 presents the components of genuine savings as the share of GNP for Tunisia. Note that this calculation omits, for the moment, the effects of human capital investment.

The starting point in the calculation of genuine savings is just standard national accounting. The top curve in figure 1 is gross domestic investment: total investment in structures, machinery and equipment, and inventory accumulation. Net foreign borrowing, including net official transfers, is then subtracted from this top curve to give gross savings: the difference between production and consumption over the year. Next, the depreciation of produced assets is deducted, yielding the curve for net savings. Finally, the bottom line is genuine savings, which is obtained by subtracting the value of resource depletion and pollution damages from net savings.

The basic national accounts data used to derive genuine savings rates are taken from the World Bank's *World Tables* (World Bank 1995). However, these data do not include the value of depreciation of produced assets. Unofficial estimates of depreciation, as calculated from perpetual inventory models, are taken from Nehru and Dhareshwar (1993). All of the data sets employed in this article—the

Figure 1. *Genuine Savings in Tunisia, 1970–94*

Percentage of GNP



Source: Authors' estimates.

World Tables data, the depreciation estimates, and the resource depletion and degradation calculations—have gaps in their coverage.[4]

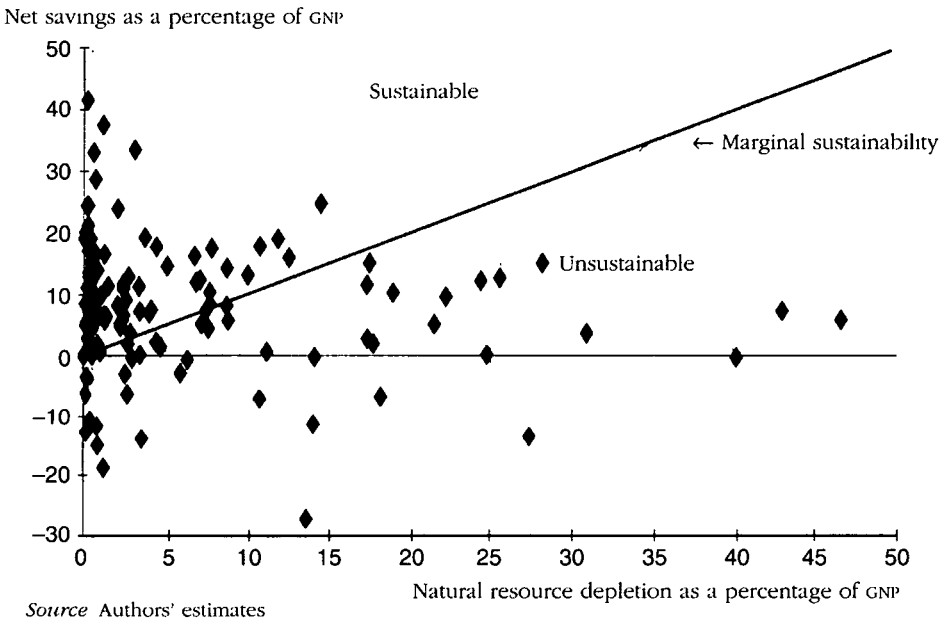
The critical elements added by the green national accounting literature are recognition of natural resources as factors of production and of environmental amenities as sources of welfare. A first question to be answered, therefore, is whether the calculation of depletion and degradation adds substantially to the picture of whether countries are on a sustainable path. This reduces the question to whether there are countries whose net savings rates are positive but whose genuine savings rates are negative.

In figure 2 the net savings rate for the industrial and developing countries in our sample is scatter-plotted against the value of depletion and CO_2 emissions, using average figures for 1988–92. The line labeled "marginal sustainability" is the 45° line—countries falling above this line have genuine savings rates that are positive, while those falling below have negative genuine savings rates. Although several countries have negative net savings rates, and so are unsustainable even by conventional national accounting measures, clearly a number have positive net savings but negative genuine savings. Measuring genuine savings, therefore provides useful new information.

A calculation of genuine savings as a percentage of GNP reveals striking differences across the regions of the world. In many developing areas decisive mo-

4. Resource extraction data in physical quantity are taken from the World Bank's Economic and Social Database.

Figure 2. *Net Savings and Natural Resource Depletion as Shares of GNP, Average 1988–92*



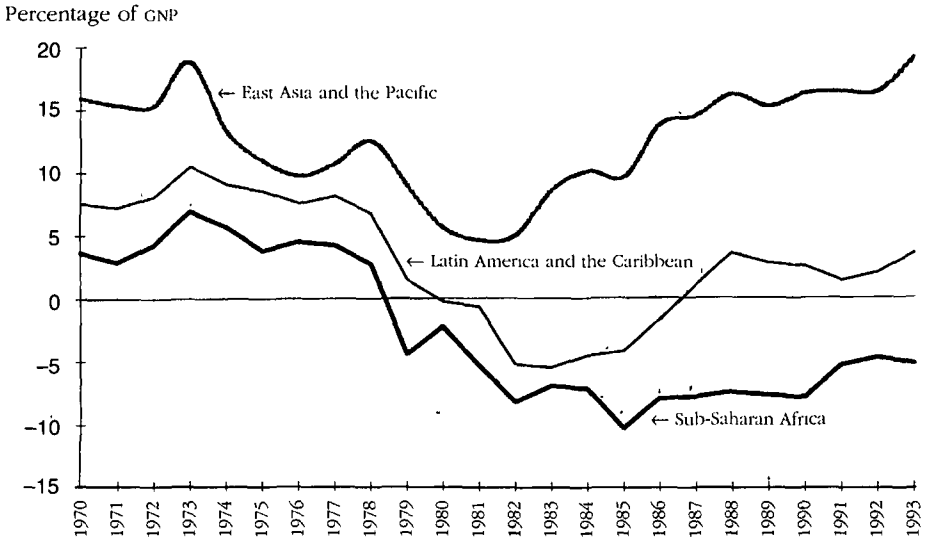
Source Authors' estimates

ments in economic performance are reflected in large movements in the genuine savings rate, shown in figures 3 and 4. These figures omit, for the time being, the effects of human capital investment.

The comparison of genuine savings rates reveals a disappointing trend for the countries of Sub-Saharan Africa (figure 3). Here average genuine savings rates rarely exceeded 5 percent during the 1970s, followed by a sharp negative turn at the end of that decade from which they have never recovered. Despite a slight improvement in the early 1990s, regional genuine *dissaving* has recently been near 7 percent. Equally important, negative genuine savings rates have been accompanied by persistently low regional indicators of human welfare, including education, nutrition, and medical care (World Bank 1996b). The savings analysis highlights the fact that the situation *with regard to future well-being* is worse than might otherwise be thought: not only has Sub-Saharan Africa performed badly by conventional measures, it is clear that the wealth inherent in the resource stocks of these countries is being liquidated and dissipated.

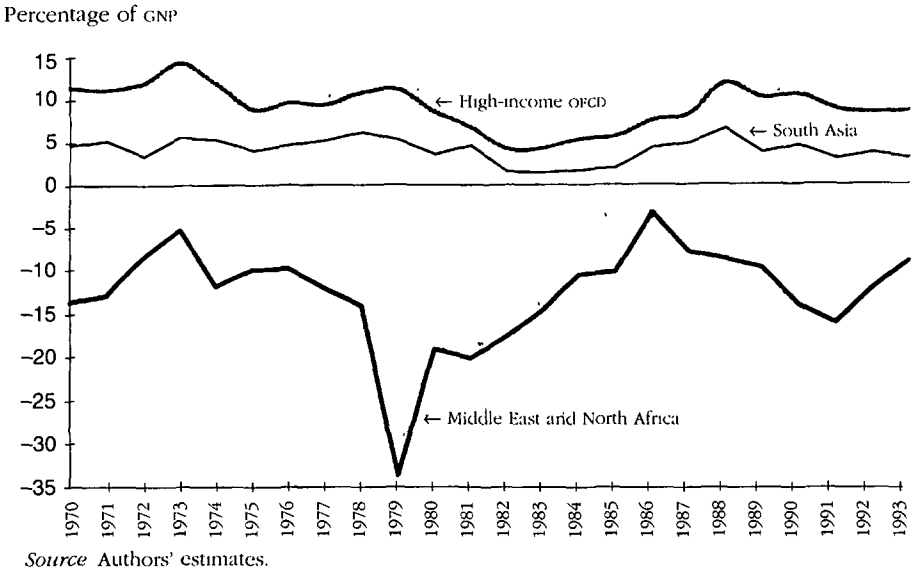
The oil crisis also coincided with a period of decline in genuine savings throughout Latin America and the Caribbean, where figures had previously remained near 8 to 9 percent of GNP. In 1982, the year of Mexico's debt crisis, regional genuine savings rates dropped to negative 5 percent. As the region has emerged from the debt crisis, returned to democratic rule, and spurred the vigorous growth of the "jaguars," genuine savings rates have shown a consistently positive trend. They remain, however, well below 5 percent.

Figure 3. *Genuine Savings as a Percentage of* GNP *for Selected Regions, 1970–93*



Source· Authors' estimates.

Figure 4. *Genuine Savings as a Percentage of* GNP *for Selected Regions, 1970–93*



Source Authors' estimates.

In stark contrast to the situation in Latin America and the Caribbean, the East Asia and the Pacific region has genuine savings rates sometimes topping 15 percent. However, the effects of important local pollutants, such as particulate matter in air, are not included in this calculation. Moreover, the 1997 financial crisis in this region shows that a robust savings effort is a necessary but not a sufficient condition for strong and smooth economic growth.

Consistently negative genuine savings in the Middle East and North Africa region stands out in figure 4. Regional total consumption as a share of GNP rose from around 50 percent in the 1970s to more than 70 percent by the end of the 1980s, and imports of food and manufactured goods flowed into the region as many current account surpluses of the 1970s turned into deficits in the 1980s (World Bank 1996a). However, the caveats about upward biases in the depletion estimates need to be considered when judging these figures: as the most resource-dependent economies, these countries exhibit the highest downward bias in estimated genuine savings rates.

South Asia exhibits moderately positive rates of genuine savings over the period. This is consistent with the moderate rates of economic growth that have characterized the countries in the region.

Finally, genuine savings rates in the high-income industrial countries (members of the Organisation for Economic Co-operation and Development—OECD), pushed upward by high investment, lack of dependence on natural resource depletion, and strong exports of high value-added goods and services, are near 10 percent for much of the period depicted. The recessions in 1982–83 and 1990 coincided with downward turns in genuine savings rates, but the volatility and high rates of genuine dissaving seen in other areas are consistently absent.

This picture of genuine savings rates is not complete, however. Just as standard measures of savings ignore asset consumption in natural resource capital such as forests, so can it be argued that they ignore investment in one of a nation's most valuable assets: its people.

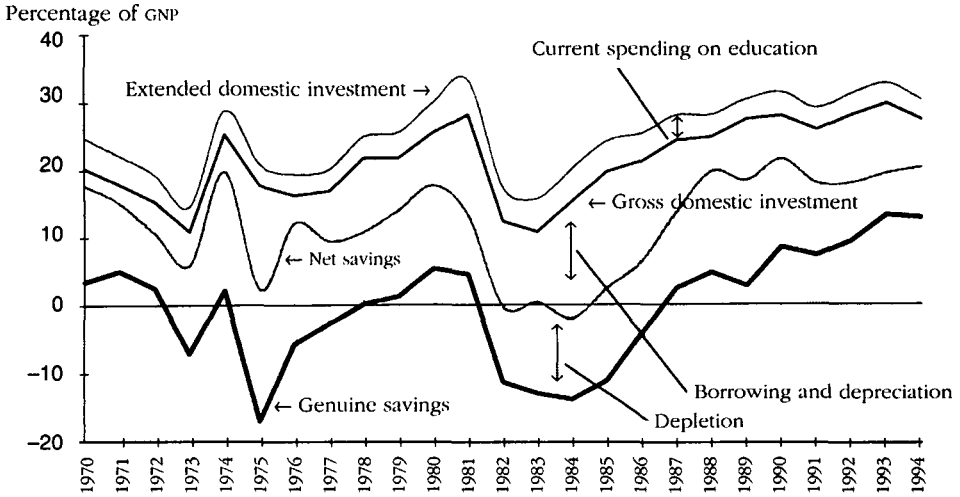
IV. Investing in Human Capital

The process of calculating genuine savings is, in essence, one of broadening the traditional definition of what constitutes an asset. Perhaps the most important of the additions to the asset base is the knowledge, experience, and skills embodied in a nation's populace, its human capital.

The world's nations augment the stock of human capital in large part through their educational systems, into which they collectively pour trillions of dollars each year. Standard national accounts label as an *investment* less than 10 percent of this amount, the portion that is spent on fixed capital, such as school buildings. Current (as opposed to capital) expenditures on education include teachers' salaries and the purchase of books, and are treated strictly as consumption. As the formal accounting model suggests, however, this is clearly incorrect. If a country's human capital is to be regarded as a valuable asset, expenditures on its formation must be seen as an investment.

The effects of including human capital investment in the genuine savings calculation can be significant. In Chile, for example, current educational expenditures represented approximately 3.1 percent of GNP in the early 1990s (figure 5). This level of investment helped to make genuine savings rates positive after the

Figure 5. *Human Capital and Genuine Savings in Chile, 1970–93*



Source: Authors' estimates.

late 1980s. In 1993 and 1994 nearly one-half of the rents from natural resource depletion were, notionally at least, being reinvested in human capital.

Adjusting rates of genuine savings to embrace changes in human capital assets shifts regional genuine savings rates markedly upward (table 2). In Sub-Saharan Africa accounting for education investment brings recent genuine savings rates close to zero. In the Middle East and North Africa genuine savings rates are consistently negative even after adjusting for education. Finally, high rates of education investment in high-income OECD countries and in East Asia and the

Table 2. *Genuine Savings by Region and Income, 1970–93*
(percentage of GNP)

	Average					
Classification	*1970–79*	*1980–89*	*1990*	*1991*	*1992*	*1993*
Region						
East Asia and the Pacific	15.1	12.6	18.6	18.7	18.7	21.3
High-income industrial countries	15.7	12.4	15.7	14.5	14.0	13.9
Latin America and the Caribbean	10.4	1.9	5.5	4.1	4.7	6.1
Middle East and North Africa	−8.9	−7.7	−8.8	−10.8	−6.6	−1.8
South Asia	7.2	6.5	7.6	6.3	7.1	6.4
Sub-Saharan Africa	7.3	−3.2	−3.8	−1.2	−0.6	−1.1
Income category						
Low	9.8	3.3	5.7	7.5	9.0	10.5
Middle	7.2	2.9	10.0	9.7	7.8	8.1
High	15.2	12.3	15.9	14.6	14.1	14.1

Note: Values include an adjustment for spending on education.
Source: Authors' estimates.

Pacific sharpen the contrast between the genuine savings effort in these areas and across the rest of the globe.

Table 3 presents country-level data for genuine savings rates for the 1970s, 1980s, and early 1990s that include current educational expenditures for all countries for which data are available. The pattern of savings shown appears to reflect "the curse of the mineral-rich"—the greater the mineral endowment, the more likely a country is to have a low or negative savings rate (Gelb 1988 and Sachs and Warner 1995).

V. Policy Issues

Many people would argue that obtaining measures of a green *NNP* is intrinsically important. However, by measuring income rather than changes in wealth, green *NNP* will have few direct uses with regard to policies for sustainable development. In contrast, genuine savings measures suggest a series of policy questions that are key to sustaining development.

It is abundantly clear that monetary and fiscal policies are the biggest levers for boosting savings rates—for example, in figure 5 shifts in gross savings in Chile (gross domestic investment less net foreign borrowing) would move all of the curves up or down. The first policy issue is therefore the extent to which monetary and fiscal policies encourage strong domestic savings.

While natural resource exports boost foreign savings and therefore the overall savings effort, the analysis of genuine savings suggests a further question: to what

Table 3. *Genuine Savings by Region, 1970–93*
(percentage of GNP)

	Average					
Economy	*1970s*	*1980s*	*1990*	*1991*	*1992*	*1993*
Latin America and the Caribbean						
Antigua and Barbuda	—	—	13.7	10.6	6.8	12.9
Argentina	17.6	3.8	8.2	2.9	2.8	5.1
Barbados	7.8	12.2	11.5	9.7	12.7	11.9
Belize	—	15.9	25.2	17.3	16.1	16.7
Bolivia	-3.8	-35.6	-19.5	-12.4	-12.6	-15.4
Brazil	12.6	9.4	11.7	9.3	11.0	11.8
Chile	-1.8	-3.4	8.5	7.3	9.4	13.2
Colombia	6.7	4.2	2.9	7.9	4.6	5.2
Costa Rica	13.0	12.2	12.8	15.2	14.7	—
Dominican Republic	13.2	9.7	9.0	3.8	2.5	10.3
Ecuador	0.7	-12.6	-21.3	-7.2	-3.1	-4.1
El Salvador	11.4	1.8	-0.1	-1.1	3.7	5.9
Grenada	—	22.0	16.7	15.6	16.9	9.9
Guatemala	9.2	-0.1	-2.6	0.6	-0.1	1.2
Haiti	0.3	-2.0	-1.1	-7.1	-16.3	-19.1
Jamaica	-0.6	-9.4	-4.5	-1.8	5.4	-2.4
Mexico	9.1	-3.0	0.9	2.0	1.9	3.6

Table 3. *(continued)*

Economy	*Average* 1970s	*Average* 1980s	*1990*	*1991*	*1992*	*1993*
Paraguay	14.9	13.2	10.4	6.6	1.0	1.0
Peru	5.8	-0.8	4.1	4.8	10.0	6.6
Suriname	15.6	7.3	3.8	-15.0	-2.4	15.4
Trinidad and Tobago	-5.8	-20.6	-19.9	-15.4	-11.4	-7.9
Uruguay	13.2	4.1	4.7	4.0	4.3	3.9
Venezuela	1.9	-17.6	-29.2	-17.6	-16.1	-14.5
East Asia and the Pacific						
China	14.1	6.6	10.5	15.5	16.4	21.5
Hong Kong, China	25.5	22.0	23.8	21.7	—	—
Indonesia	3.1	2.2	8.2	5.3	6.9	13.8
Korea, Rep. of	17.0	25.5	33.5	28.2	27.4	29.6
Malaysia	17.7	9.6	9.9	9.9	15.2	18.6
Myanmar	3.6	-0.4	1.4	4.7	3.6	2.4
Papua New Guinea	—	-9.4	-2.9	-1.1	-0.8	6.4
Philippines	14.1	6.3	5.7	5.8	7.8	8.0
Singapore	17.8	30.3	37.1	39.0	39.5	37.9
Taiwan (China)	23.9	22.3	18.6	19.2	17.8	17.2
Thailand	16.4	17.7	27.0	25.7	25.6	28.1
Middle East and North Africa						
Algeria	-1.8	-9.0	-11.7	-2.2	2.1	6.7
Bahrain	—	0.8	-4.3	-26.7	-11.2	-2.7
Egypt	10.5	0.5	2.6	9.4	15.2	13.6
Iran, Islamic Rep. of	—	-2.7	-11.3	-5.8	-3.3	—
Israel	15.9	14.4	18.4	18.5	18.9	16.7
Jordan	—	—	-2.9	-0.2	12.9	13.5
Morocco	12.3	11.8	19.8	15.6	15.5	15.2
Saudi Arabia	-27.6	-25.5	-27.3	-41.8	-33.0	-20.2
Syrian Arab Republic	10.3	-2.3	-9.0	-10.0	—	—
Tunisia	11.6	3.8	10.1	10.2	12.2	12.8
South Asia						
Bangladesh	-2.3	-0.9	-0.1	-2.7	-0.1	2.4
India	8.4	7.1	8.8	8.1	8.2	7.2
Nepal	-0.6	-1.5	-9.5	-8.8	-9.9	-12.3
Pakistan	3.5	5.4	3.8	1.6	5.0	4.7
Sri Lanka	12.3	8.9	7.8	5.5	9.6	12.4
Sub-Saharan Africa						
Benin	8.1	-3.9	1.9	2.0	0.1	0.8
Burkina Faso	14.8	9.1	7.0	6.7	7.4	8.6
Burundi	—	—	-4.4	-1.5	-1.2	0.0
Cameroon	15.8	7.7	-7.0	-2.5	-3.6	-0.6
Central African Republic	8.4	-0.2	-4.0	-0.3	0.2	-7.7
Chad	—	—	-5.6	-7.7	-9.1	-8.4
Congo, Rep. of	-5.3	-16.8	-28.0	-22.0	-18.7	-28.6
Côte d'Ivoire	16.5	-0.3	-9.8	-13.6	-10.7	-12.3
Gabon	19.7	-7.3	-2.7	4.5	-3.0	2.0

(Table continues on the following page.)

Table 3. *(continued)*

Economy	Average 1970s	Average 1980s	1990	1991	1992	1993
Gambia, The	0.3	3.5	15.4	12.8	20.2	7.7
Ghana	4.1	-6.0	1.4	1.8	-3.2	-4.6
Guinea	—	—	-11.6	-13.5	-9.4	0.1
Guinea-Bissau	—	—	8.8	0.1	-9.4	6.4
Kenya	5.8	5.1	2.8	7.0	3.9	1.4
Madagascar	6.0	-1.3	1.1	-8.7	-3.9	-0.2
Malawi	10.3	-2.9	3.5	-2.2	-9.8	-7.9
Mali	7.3	3.5	11.1	12.5	9.5	9.7
Mauritania	-18.4	-15.8	-10.2	-12.0	-11.4	-14.9
Mauritius	13.3	11.3	19.5	18.7	19.8	18.3
Namibia	—	-5.4	-2.0	16.1	16.2	11.4
Niger	8.5	1.9	-5.8	-1.7	-7.5	-6.5
Nigeria	3.3	-25.3	-46.4	-33.9	-30.2	-37.1
Rwanda	4.1	5.6	2.2	-3.2	-4.0	-1.4
Senegal	6.5	-4.1	8.0	2.3	1.6	3.2
Sierra Leone	-2.5	0.5	-10.4	-2.7	—	—
South Africa	10.4	5.4	5.5	6.2	4.7	5.2
Togo	13.9	13.0	12.4	5.6	2.3	-12.5
Uganda	—	-23.2	-13.7	-0.6	-0.2	-8.6
Zambia	-5.7	-27.3	-32.0	-14.5	-5.3	-16.1
Zimbabwe	9.1	7.4	15.6	6.1	-0.8	8.7
High-income industrial countries						
Australia	11.8	7.5	6.1	6.9	7.2	5.5
Austria	18.3	13.4	18.8	20.2	19.2	16.2
Belgium	17.9	9.3	18.3	15.6	15.9	16.7
Canada	16.1	9.8	10.3	6.7	6.2	7.4
Denmark	16.0	8.4	17.4	14.9	15.5	14.2
Finland	15.4	13.9	18.5	9.4	7.0	5.5
France	19.3	12.8	18.6	16.0	15.4	13.5
Germany	—	—	—	11.1	10.8	10.0
Germany, former Federal Rep.	19.9	13.5	14.3	15.0	13.8	12.2
Greece	15.8	4.6	7.6	9.1	8.0	8.0
Ireland	9.5	9.3	15.9	18.5	16.9	17.4
Italy	17.3	13.3	16.9	12.9	11.8	12.3
Japan	26.5	21.7	31.2	28.7	28.3	26.2
Luxembourg	10.1	9.7	15.8	14.6	12.8	13.9
Netherlands	21.0	13.8	20.8	17.8	16.9	15.6
New Zealand	16.3	9.2	9.1	7.4	9.7	13.6
Norway	14.5	8.7	11.1	15.6	14.2	7.0
Portugal	15.2	12.6	23.4	19.7	20.4	18.1
Spain	16.2	11.0	18.8	14.3	12.5	12.3
Sweden	18.3	12.1	16.1	10.8	7.5	5.6
Switzerland	17.1	19.1	25.7	24.0	22.0	19.9
Turkey	16.3	12.3	18.9	13.1	11.9	15.4
United Kingdom	11.0	8.4	11.4	7.0	6.7	6.6
United States	11.0	9.0	8.2	8.8	8.3	9.6

— Not available.
Note: Values include current education expenditures.
Source: Authors' estimates.

extent do exports of *exhaustible* resources boost the rate of genuine savings?[5] The answer to this lies in netting out the value of resource depletion from the value of gross exports.

More optimal natural resource extraction paths will, other things being equal, boost the value of genuine savings. The policy question for natural resource management is therefore the extent to which stronger resource policies (royalty regimes, tenure) can boost the rate of genuine savings. Similarly, reducing pollution emissions to socially optimal levels will boost the value of genuine savings. The policy issue with respect to pollution is the extent to which more optimal pollution control policies can increase the rate of genuine savings.

The policy prescriptions for boosting genuine savings should never be to stop extracting resources or emitting pollutants altogether. Rather, pricing resources and pollutants correctly and enforcing property rights will lead to efficient levels of exploitation of the environment, reducing incentives to "high-grade" resources (meaning, to take the cheapest and most accessible resources first, without considering dynamic efficiency) or pollute indiscriminately. Optimal resource and environmental policies will maximize genuine savings, subject to the macroeconomic policy regime in place. However, the sorts of issues raised by Gelb (1988) about the nature and effects of oil windfalls in developing countries are particularly relevant to the policy issues just addressed: without sound macroeconomic policies and prudent allocation of public resources, the effects of reliance on large resource endowments can be negative for many countries.

VI. Conclusions

Savings rules have been criticized because they are concerned only with *weak* sustainability (see, for example, Martínez-Alier 1995).[6] One response to this criticism is to suggest that countries that fail the savings rule—exhibit persistently negative genuine savings—probably also fail to meet the criteria for *strong* sustainability, in the sense that critical natural assets are being depleted. It would be surprising if this were not the case. And as Pearce, Hamilton, and Atkinson (1996) argue, even if some amount of a critical resource must be preserved to meet the criteria of strong sustainability, savings rules are still required for the remaining resources if sustainability is to be achieved.

Thinking about sustainable development and its measurement leads naturally to a conception of the process of development as one of portfolio management. Prudent governments will not only consider natural resources as assets and pollution stocks as liabilities in the national balance sheet, they will also be concerned with the appropriate mix of produced assets and human capital.

5. The question is also germane for unsustainable forest harvest programs.

6. Pearce, Markandya, and Barbier (1989) define weak sustainability to mean that natural and produced assets are fully substitutable. Strong sustainability implies that at least some natural assets have no substitutes and so may need to be conserved if development is to be sustainable.

Questions of the "appropriate mix" of assets are inherently questions about returns on the marginal investment. This marginal investment may be in better resource management, boosting the value of natural resources in the national balance sheet; it may be in pollution control, decreasing the size of the pollution liability to its efficient level; it may be in infrastructure, as has traditionally been the case; and it may be in primary education, as an essential building block in increasing human capital.

The policy implications of measuring genuine savings are quite direct: persistently negative rates of genuine savings must lead, eventually, to declining well-being. For policymakers, linking sustainable development to rates of genuine savings means that there are many possible interventions to increase sustainability, from the macroeconomic to the purely environmental.

There are several omissions in the foregoing empirical analysis: soils, fish, water resources, water pollutants, and air pollutants (other than CO_2), to name a few. Notwithstanding these omissions, the empirical evidence is that genuine savings is negative in a wide range of countries when the environment and natural resources are included in the savings measure. Negative genuine savings is more than a theoretical possibility, therefore, and the evidence is that many countries, particularly in Sub-Saharan Africa, are being progressively impoverished. Increasing the coverage of natural resources and pollutants in our calculations would lower the estimated levels of genuine savings overall.

In terms of further research there are obvious refinements that can be envisioned, including treating soil degradation and expanding the country coverage of data on the marginal social costs of pollution emissions. The latter is particularly important for rapidly growing countries: as countries grow, they tend to urbanize, and these urban areas tend to develop problem levels of pollution.

A particularly appealing topic for further research would be an exploration of the "resource curse" hypothesis. Sachs and Warner (1995) find that resource dependence (under a variety of definitions) is negatively correlated with economic growth, based on an analysis of industrial and developing countries since the 1970s. The data developed for the genuine savings analysis should allow more precise definitions of resource dependence, permitting the hypothesis to be retested. In addition, the estimated genuine savings rates may turn out to have explanatory power in the resource curse model.

REFERENCES

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

Adelman, M. A. 1991. "U.S. Oil/Gas Production Cost: Recent Changes." *Energy Economics* 13(October):235–37.

Asheim, Geir. 1994. "Net National Product as an Indicator of Sustainability." *Scandinavian Journal of Economics* 96(2):257–65.

Barnes, Douglas F. 1992. "Understanding Fuelwood Prices in Developing Nations." World Bank, Energy, Mining, and Telecommunications Department, Washington, D.C. Processed.

Bhattacharya, Subhes C. 1995. "Estimation of Subsidies on Coal in India." *Natural Resources Forum* 19(2):135–42.

CDIAC (Carbon Dioxide Information and Analysis Center). 1994. *Trends '93: A Compendium of Data on Global Change.* With accompanying magnetic media. Oak Ridge, Tenn.: Oak Ridge National Laboratory.

Cornot-Gandolphe, Sylvie. 1994. *Adequate and Secure Gas Supply for Europe: How to Meet the Growing Demand for Natural Gas?* Paris: International Energy Agency, International Center for Gas Technology Information.

Dasgupta, Partha, and Geoffrey Heal. 1979. *Economic Theory and Exhaustible Resources.* Cambridge, U.K.: Cambridge University Press.

Doyle, Guy. 1987. *China's Potential in International Coal Trade.* London: International Energy Agency Coal Research.

Duvigneaud, Paul, ed. 1971. *Productivity of Forest Ecosystems.* Paris: United Nations Educational, Scientific, and Cultural Organization.

Fankhauser, Samuel. 1994. "Evaluating the Social Costs of Greenhouse Gas Emissions." *The Energy Journal* 15(2):157–84.

FAO (Food and Agriculture Organization of the United Nations). 1983. *Forest Products Prices, 1963–1982.* Forestry Paper 46. Rome.

———. 1994. *Forest Products Yearbook 1993.* Rome.

———. 1995. *Forest Products Prices, 1973–1992.* Forestry Paper 125. Rome.

FAO/UNECE (Food and Agriculture Organization of the United Nations/United Nations Economic Commission for Europe). 1992. *The Forest Resources of the Temperate Zones.* Vol. 1: *General Forest Information.* New York: United Nations.

Freeman, A. M. 1994. *The Measurement of Environmental and Resource Values: Theory and Methods.* Washington, D.C.: Resources for the Future.

Gelb, A. H. 1988. *Oil Windfalls: Blessing or Curse?* New York: Oxford University Press.

Hamilton, Kirk. 1994. "Green Adjustments to GDP." *Resources Policy* 20(3):155–68.

———. 1996. "Pollution and Pollution Abatement in the National Accounts." *Review of Income and Wealth* 42(1, March):13–33.

———. 1997. "Defining Income and Assessing Sustainability." World Bank, Environment Department, Washington, D.C. Processed.

Hamilton, Kirk, Giles Atkinson, and David W. Pearce. 1998. "Savings Rules and Sustainability: Selected Extensions." Paper presented to the World Congress on Environment and Resource Economics, Venice, June 25–27. Processed.

Hartwick, John M. 1977. "Intergenerational Equity and the Investing of Rents from Exhaustible Resources." *American Economic Review* 67(5):972–74.

———. 1990. "Natural Resources, National Accounting and Economic Depreciation." *Journal of Public Economics* 43(3):291–304.

———. 1992. "Deforestation and National Accounting." *Environmental and Resource Economics* 2(5):513–21.

Hill, Peter, and Anne Harrison. 1994. "Accounting for Subsoil Assets in the 1993 SNA." Paper presented to the London Group on National Accounts and the Environment, London, March 15–18. Processed.

IDB (Inter-American Development Bank). 1981. *Investment and Financing Requirements for Energy and Minerals in Latin America.* Washington, D.C.

IEA (International Energy Agency). 1994a. *Energy Policies of the Czech Republic.* Paris: Organisation for Economic Co-operation and Development.

———. 1994b. *Russian Energy Prices, Taxes, and Costs.* Paris: Organisation for Economic Co-operation and Development.

———. 1995a. *Coal Information 1994.* Paris: Organisation for Economic Co-operation and Development.

———. 1995b. *Energy Policies of Poland.* Paris: Organisation for Economic Co-operation and Development.

———. 1995c. *Energy Policies of the Russian Federation, 1995 Survey.* Paris: Organisation for Economic Co-operation and Development.

———. 1995d. *Middle East Oil and Gas.* Paris: Organisation for Economic Co-operation and Development.

———. 1995e. *Oil, Gas, and Coal: Supply Outlook.* Paris: Organisation for Economic Co-operation and Development.

———. 1996. *Energy Policies of Ukraine, 1996 Survey.* Paris: Organisation for Economic Co-operation and Development.

Jenkins, Gilbert. 1989. *Oil Economist's Handbook.* 2 vols., 5th ed. New York: Elsevier.

Julius, DeAnn, and Afsaneh Mashayekhi. 1990. *The Economics of Natural Gas: Pricing, Planning, and Policy.* London: Oxford University Press, Oxford Institute for Energy Studies.

Kanowski, P. J., and others. 1992. "Plantation Forestry." In Narendra P. Sharma, ed., *Managing the World's Forests: Looking for a Balance Between Conservation and Development.* Dubuque, Iowa: Kendall/Hunt Publishing Co.

Kellenberg, John. 1995. "Accounting for Natural Resources: Ecuador, 1971–1990." Ph.D. diss. Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, Md.

Khan, Amar R. 1986. "The Prospects for Gas in Developing Countries." In Robert Mabro, ed., *Natural Gas: an International Perspective.* London: Oxford University Press, Oxford Institute of Energy Studies.

Lamprecht, Hans. 1989. *Silviculture in the Tropics.* Eschborn, Germany: Deutsche Gesellschaft für Technische Zusammanarbeit.

Liefert, William Mark. 1988. "The Full Cost of Soviet Oil and Natural Gas Production." *Comparative Economic Studies* 30(summer):1–20.

Mäler, K.-G. 1991. "National Accounts and Environmental Resources." *Environmental and Resource Economics* 1(1):1–15.

Martínez-Alier, Juan. 1995. "The Environment as a Luxury Good or 'Too Poor to be Green'?" *Ecological Economics* 13(1):1–10.

Mashayekhi, Afsaneh. 1983. "Marginal Cost of Natural Gas in Developing Countries: Concepts and Applications." World Bank, Energy Department, Washington, D.C. Processed.

Meyer, Sandra. 1994. *Natural Gas Statistics Sourcebook.* Tulsa, Oklahoma: PennWell Publishing Co., Oil and Gas Journal Energy Database Publications.

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

———. 1994. "New Estimates of Total Factor Productivity Growth for Developing and Industrial Countries." Policy Research Working Paper 1313. World Bank, Policy Research Department, Washington, D.C. Processed.

Openshaw, Keith, and Charles Feinstein. 1989. "Fuelwood Stumpage: Financing Renewable Energy for the World's Other Half." Policy Research Working Paper 270. World Bank, Policy Research Department, Washington, D.C. Processed.

Pearce, David W., and Giles Atkinson. 1993. "Capital Theory and the Measurement of Sustainable Development: An Indicator of Weak Sustainability." *Ecological Economics* 8(2):103–08.

Pearce, David W., Kirk Hamilton, and Giles Atkinson. 1996. "Measuring Sustainable Development: Progress on Indicators." *Environment and Development Economics* 1(1):85–101.

Pearce, David, Anil Markandya, and Edward Barbier. 1989. *Blueprint for a Green Economy*. London: Earthscan.

Pezzey, John. 1989. "Economic Analysis of Sustainable Growth and Sustainable Development." Working Paper 15. World Bank, Environment Department, Washington, D.C. Processed.

Repetto, Robert, William Magrath, Michael Wells, Christine Beer, and F. Rossini. 1989. *Wasting Assets: Natural Resources in the National Accounts*. Washington, D.C.: World Resources Institute.

Sachs, Jeffrey, and Andrew Warner. 1995. "Natural Resource Abundance and Economic Growth." Development Discussion Paper 517a. Harvard Institute for International Development, Harvard University, Cambridge, Mass. Processed.

Sadoff, C. W. 1992. *The Effects of Thailand's Logging Ban: Overview and Preliminary Results*. Bangkok: Thailand Development Research Institute.

Sagers, Matthew, Valeriy Kryukov, and Vladimir Shmat. 1995. "Resource Rent from the Oil and Gas Sector and the Russian Economy." *Post-Soviet Geography* 36(7):389–425.

Smith, Roger. 1992. "Income Growth, Government Spending, and Wasting Assets: Alberta's Oil and Gas." *Canadian Public Policy* 18(4):387–412.

Tretyakova, Albina, and Meredith Heinemeier. 1986. *Cost Estimates for the Soviet Coal Industry: 1970 to 1990*. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, Center for International Research.

United Nations. 1993. *Integrated Environmental and Economic Accounting*. Series F 61. New York.

U.S. Bureau of Mines. 1987. "An Appraisal of Minerals Availability for 34 Commodities." Bulletin 692. Department of the Interior, Washington, D.C. Processed.

van Buren, Ariane. 1990. *The Woodfuel Market in Nicaragua: The Economics, Sociology, and Management of a Natural Energy Resource*. Amsterdam: Centrum voor Studie en Documentatie van Latijns Amerika.

Vincent, Jeffrey, Theodore Panayotou, and John Hartwick. 1997. "Resource Depletion and Sustainability in Small Open Economies." *Journal of Environmental Economics and Management* 33(3):274–86.

Vitousek, Peter, Paul Ehrlich, Ann Ehrlich, and Pamela Matson. 1986. "Human Appropriation of the Products of Photosynthesis." *BioScience* 36(6, June):368–73.

Weitzman, M. L. 1976. "On the Welfare Significance of National Product in a Dynamic Economy." *Quarterly Journal of Economics* 90(1):156–62.

Weitzman, M. L, and K.-G. Löfgren. 1997. "On the Welfare Significance of Green Accounting as Taught by Parable." *Journal of Environmental Economics and Management* 32(2):139–53.

World Bank. 1993. *Commodity Trade and Price Trends, 1989–91*. Baltimore, Md.: Johns Hopkins University Press.

———. 1995. *World Tables 1995*. Washington, D.C.

———. 1996a. International Economics Department data. Washington, D.C.

———. 1996b. *World Development Report 1996: From Plan to Market*. New York: Oxford University Press.

———. 1997. *Expanding the Measure of Wealth*. ESD Studies and Monographs Series 17. Washington, D.C.

———. Various years. *World Development Indicators*. New York: Oxford University Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2: 357-78

Winners and Losers from the Privatization and Regulation of Utilities: Lessons from a General Equilibrium Model of Argentina

Omar Chisari, Antonio Estache, and Carlos Romero

A computable general equilibrium (CGE) model is used to estimate the macroeconomic and distributional effects of the privatization and regulation of utilities in Argentina, begun in 1989. Based on data available after the privatization that indicate different kinds of efficiency gains in electricity, gas, water, and telecommunications, both the privatization and effective regulation are estimated to yield significant macroeconomic benefits. Gains from the privatization accrue mainly to high-income classes, while gains from the effective regulation of newly privatized utilities accrue mainly to low-income classes. CGE estimates of overall employment effects suggest that privatization was not a major contributor to the dramatic rise in unemployment in Argentina between 1993 and 1995. This rise was more likely due to the "Tequila Effect" of an interest rate shock.

In 1989 Argentina initiated a path-breaking process of privatizing its infrastructure services. The reforms are not yet concluded, and many provincial water and electricity companies remain in the hands of the public sector. But the estimated effects of the initial reforms will probably generalize because the patterns of reform across the country are similar. The reforms are driven primarily by the need to alleviate the fiscal burden imposed by public utilities in every province and by a desire to involve the private sector in financing the expansion of these sectors. The privatization has been praised by some and criticized by others.

This article provides an early assessment of both the macroeconomic and distributional impacts of the private operation of electricity, gas, water and sanitation, and telecommunications services, and indicates the value of effective regulation to the various income classes. The most important conceptual contribution is the use of a computable general equilibrium (CGE) model to estimate the general equilibrium and distributional effects of privatization. The model follows

Omar Chisari and Carlos Romero are affiliated with the Universidad Argentina de la Empresa in Buenos Aires, and Antonio Estache is with the World Bank Institute (formerly the Economic Development Institute). The authors thank Daniel Benitez for extremely competent research assistance and Shanta Devarajan, Mathias Dewatripont, Francisco Ferreira, Cheikh Kane, Ioannis Kessides, Abel Mejia, Martín Rodríguez-Pardina, Chantal Roucolle, Suzanne Smith, participants in an Inter-American Development Bank seminar, and the referees for very useful comments and suggestions. A longer version of the article was issued as Chisari, Estache, and Romero (1997).

the approach described in Shoven and Whalley (1992), in which relative prices adjust to clear all markets. However, unemployment arises because of some inflexibility in foreign exchange markets.

In spite of its well-known limitations, this approach is particularly useful for the following reasons. First, it allows calibration of the key technological parameters based on information requirements that are much less demanding than those of econometric models. Second, it allows comparative static simulations of the impact of changes within the sector, or across the economy, either one at a time or simultaneously (see Bergman 1990). This feature is useful because it assesses the direct and indirect impacts of all the changes in one utility or the impact of a similar change across utilities. Third, the approach allows an assessment of the interactions between privatization and other significant macroeconomic changes, such as the "Tequila Effect."

Galal and Shirley (1994) recently published the results of a detailed World Bank study that focused on the efficiency aspects of privatization in the United Kingdom, Chile, Mexico, and Malaysia, but their methodology does not address the general equilibrium or distributional aspects of privatization. Their methodology also requires more detailed data on the performance of public utilities before privatization than were available in Argentina and does not permit as broad a scope for policy simulations as the approach adopted here. Recent work by Burns and Weyman-Jones (1994) and Button and Weyman-Jones (1994) also deals with the gains from privatization but focuses only on a specific industry. For an overview of the effects of deregulation on the U.S. economy, see Winston (1993).

It is assumed that the changes observed in the privatization already implemented will be duplicated when provincial services are privatized. About 33 percent of industrial production, almost 50 percent of services, and more than 40 percent of the population are concentrated around Buenos Aires, where most of the initial reforms were introduced. Moreover, large electricity users throughout the country can bypass local distribution companies and access the wholesale electricity market, implying that privatizing the remaining provincial public distribution companies will produce modest macroeconomic effects. The only sector significantly affected by the assumption is the water sector, where privatization has been limited so far to a few provinces in addition to Buenos Aires.

Section I discusses the major reforms in the delivery of infrastructure services in Argentina since 1989 and their impact on the performance of utilities. Section II presents the model. Section III explains how the effects of the private operation of utilities (former public enterprises) and their regulation are modeled. Section IV discusses the macroeconomic effects of the reforms. Section V discusses the distributional effects. Section VI summarizes major findings.

I. Privatization of Argentina's Utilities

Some restructuring took place before utilities were transferred to private operators. Restructuring and privatizing electricity began in 1991. The three stages

of production in the sector—generation, transmission, and distribution—were vertically disintegrated, and different regulatory criteria were adopted for each activity. Generation became competitive, while transmission and distribution became regulated private monopolies. About one-third of all distribution companies have now been concessioned. These cover more than 60 percent of the population of the country. Gas was restructured at the end of 1991 when transport and distribution were separated into two transporters and eight regional distribution concessions. These activities are now controlled by local monopolies. The transfer of the telecommunications company to private operators was concluded in November 1990. The service is now provided by two private monopolies instead of a single public monopoly. In the water sector the bulk of reforms are more recent, and competition is being introduced through a bidding process. Concessionary contracts are the main regulatory instrument. About one-third of the states have privatized their water and sanitation in this way, but the affected population represents more than two-thirds of the nation's population.[1]

Ideally, to assess the impact of privatization, the performance of utilities under private operation should be compared with their performance under public management. However, the necessary data were not collected by the public managers of these utilities. Most of the efficiency and quality indicators are available only for the period since private operators took charge, so only progress made during the period of private operation can be followed. It is relatively easy to assess the changes that private operation has brought because a law requires each privatized firm to publish the composition of its costs. This information provides a good indication of the changes that are taking place in each sector and is the basis of the discussion presented here to ensure comparability across sectors.

For the purposes of calibrating the model, the base year is 1993, the first year in which the private sector essentially controlled all sectors. Table 1 shows total changes in performance between 1993 and 1995. Although there had already been improvements (since the date of privatization), the reported gains were sufficient to imply a significant impact on the rest of the economy.

II. The Model

To assess the impact of privatization on the rest of the economy, we need a macroeconomic model accounting for interactions among sectors. Our model is built around a social accounting matrix constructed for 1993 that isolates every utility from the other accounts. (See Chisari and Romero 1996 for a similar model.) It is consistent with national accounts for 1993, which is also the first year in which private operators managed all national utilities. Its basic structure is pro-

1. A brief description of the privatization process in Argentina is available in Shaikh (1996). For a discussion of key regulatory issues in Argentina, see Estache and Rodríguez-Pardina (1997) and Crampes and Estache (1997). A useful complement focusing on electricity is provided by Spiller and Viana Mantorell (1996).

Table 1. *Changes in Performance in Argentina's Utilities, 1993–95*
(percent)

Change	*Electricity*		*Gas distribution*	*Water distribution*	*Telecommunications*
	Generation	*Distribution*			
First year of private operation	1992	1992	1992	1993	1990
Efficiency gains (measured as reductions in intermediate inputs purchased as a share of total sales value)	19.5	6.3	8.8	4.9	11.3
Labor productivity gains (measured as gigawatt-hours per staff for electricity, thousands of cubic meters per staff for gas, population served per staff for water, lines in service per staff for phones)	23.1	17.6	4.8	–27.6	21.3
Improvements in quality (measured as reductions in losses, net of consumption by transmission, per production for electricity and gas; water unaccounted for per production for water; lines in repair per lines in service for phones)	—	10.0	27.8	6.1	4.6
Changes in legal weighted average tariffs deflated by the retail price index (weights are given by sales to each customer group: residential, commercial, industrial)	n.a.	–9.5	–0.5	5.5	–4.9

— Not available.
n.a. Not applicable.
Note: The table reflects the changes achieved under private management of the services. 1993 is the first year in which all sectors had benefited from some initial adjustment by the private operator. 1995 is the last year for which data were available at the time of this writing.
Source: The figures reflect the authors' own calculations based on data collected from the private operators (most of the information is available from the operators' annual reports, and some additional information was collected through direct interviews and from regulators).

vided in table 2. Note that expenditures must equal revenue for each aggregate account.

The model identifies 21 domestic production sectors, 10 for goods and 11 for services. In addition to the usual services, the social accounting matrix identifies electricity generation, electricity distribution, gas, water, and communications as separate sectors. Three factors of production are accounted for: labor, physical capital, and financial capital. Labor and financial capital are assumed to be mobile across sectors, while physical capital is not. Domestic consumers are divided into five income classes, and there is only one foreign consumer and one foreign producer. We rely on the small assumption of an open economy, implying that Argentina is a price taker in international markets.

We had to make several critical assumptions concerning data.[2] First, some of the production data were not available for 1993, and we had to fill the holes with 1986 data, the last year for which detailed information was available. Second, the matrix of intermediate purchases is based on the 1984 data adjusted to the values of the national census of 1993. Third, the distribution of factor incomes across income groups is based on the distribution observed in the province of Buenos Aires in 1991. Finally, the composition of consumption is based on the 1986 household consumption survey updated with information available for 1991.

For both the input and output matrix and household consumption, we maintained consistency with national accounts data by relying on the RAS method (Bacharach 1970).[3] Data for the composition of spending by national and provincial governments are available for 1993. Municipal expenditures are assumed to be distributed in the same proportion as the average for the two other levels of government. (No information on expenditures is available for Argentina at the municipal level; however, most of those expenditures are in employment.) Infrastructure data are based on information on assets, inputs, and expenditures from annual balance sheets of companies and complementary data provided by the national regulatory entities and the sector secretariats (energy, water resources, communications). We used sensitivity analysis to confirm that the data are reasonable.

The behavioral assumptions are contained in the following equations.

Consumers

The representative consumer of income group h has a utility function:

$$U^h = U^h \{c^d(h), c^m(h), I^d(h), S(h), S_g(h), B(h), C_r[Q_{C(h)}, \pi]\}. \tag{1}$$

2. The data sources used to construct the accounts are detailed in an appendix (in Spanish) available from the authors. This appendix explains how the data were collected, how several partial studies conducted by the statistical office were used to update information on production and consumption, and the various techniques used to check the consistency of the information collected.

3. RAS is a code name that comes from the notation $r_{ij}\, a_{ij}\, s_j$, where r_i and s_j are adjustment coefficients for the a_{ij} (input-output coefficients).

Table 2. *Social Accounting Matrix and Economic Features of the* CGE *Model for 1993*

	Expenditures				
Revenue	*Domestic product sectors*	*Private*	*Government*	*Investment*	*External sector*
Domestic product sectors (21 sectors, including separated infrastructure services)	Domestic purchases: ($132.370 billion) • CES value added for private firms • Leontief value added for privatized firms • Market clearing prices for nontradables for given levels of rationing in factor markets • Combination with other goods and services in fixed proportions	Spending on domestic goods: ($175.082 billion) • Cobb-Douglas utility in goods • Fixed proportion with goods for retail trade • Separate quantity, price, and quality for each privatized service • Rationing possible	Spending on goods and services: ($6.085 billion) • Cobb-Douglas social welfare function in purchases of goods and services, bonds, retiree services, and investment • Purchases of goods and services in fixed proportions	Final demand for investment goods: ($42.16 billion)	Exports: ($16.237 billion) • Foreign consumer has a Cobb-Douglas utility in exports and imports • Foreign consumer can issue bonds to pay for net imports • Argentina is a price taker in exports and imports • Whatever Argentina cannot consume is sold abroad at given price
External sector	Imports: ($8.182 billion) fixed proportion with value added	Spending on imports: ($8.727 billion) imperfect substitution with domestic substitutes		Imports of capital goods: ($4.150 billion) fixed proportion with value added	

Government	Trade tax revenue: ($1.282 billion) Direct taxes paid by firms: ($22.461 billion) Indirect taxes: ($25.283 billion)	Trade tax revenue: ($1.133 billion) Direct taxes paid by households: ($4.519 billion)		
Families (five income classes)	Labor income net of taxes: unemployment in the benchmark year ($60.786 billion) Capital income net of taxes: can be domestic or foreign ($122.266 billion)		Salaries and public sector transfers: ($43.645 billion)	
Investment		Private savings: ($37.196 billion)	Public savings: ($4.948 billion)	Foreign savings: (4.822 billion)

Note: The figures in parentheses are values in current prices. GDP in 1993 was $256.329 billion.
Source: Authors' calculations based on data published or provided by INDEC (the National Statistics Office) and by the private operators and regulators of utilities.

Equation 1 is a Cobb-Douglas function for all goods except retail trade, assumed to be purchased in fixed proportions with the rest of the goods and services. The preferences of domestic agents are assumed to follow an Armington specification that implies no perfect substitutability between domestic and imported goods.[4] $S(h)$ stands for the supply of labor to the private sector, and $S_g(h)$ stands for the supply of labor to the public sector; this separation is useful for some simulations if it is assumed that it is not easy to instantaneously transform a public employee into a private worker.

Expenditures are distributed as follows:

- Domestic consumption goods c^d and investments I^d at price p.
- Imported goods c^m at prices p_m.
- Bond services B at prices p_b.
- Goods and services of privatized firms represented by an index C_r, combining the quantity Q_C with quality π at price r_C per unit of Q_C. A change in quality is not necessarily associated with a change in the price of the service provided by the privatized firm. C_r can follow a multiplicative form, such as $C_r = Q_C\ v(\pi/\pi^N)$, where π^N is the normal level of quality and v is a nondecreasing function of π/π^N. An increase in service failures raises costs for consumers of services because they need to buy a larger number of physical units to reach the desired flow of services. This "naive" modeling approach permits modeling the costs of power losses or interruptions as a share of unit costs.

In some simulations prices are differentiated by income groups r_C. Equation 2 gives the budget constraint for income group h:

$$(2)\quad (1 + t_i)[pI^d(h) + pc^d(h)] + (1 + t_m)\, p_m c^m(h) + (1 + t_{ir})\, r_C\, C_r(h) + p_b B(h)$$
$$= [wS(h) + w_g S_g(h) + \theta(h)(r_p K_{po} + r_p K_{pxo} + N^p + N^{px}) + \theta_r(h)(r_r K_{ro} + N^r)]$$
$$(1 - t_d) + p_b B^o(h) + p_R\, R^o.$$

The family pays indirect taxes at rates t_i and t_{ir}, depending on the type of good and service, direct taxes t_d, and taxes on imports t_m. Its income sources are labor income in the private sector S at salary w, labor income in the public sector S_g at salary w_g, capital K_{po} and K_{pxo} in private firms remunerated at rate r_p, revenue from profits on domestic sales N^p and sales abroad N^{px}, and revenue from participation in the privatized firm N^r in proportion to shares owned, indicated as θ_r. θ_r also represents the participation of the income group in each sector-specific capital $r_p K_{po}$, $r_p K_{pxo}$, and $r_r K_{ro}$. In the scenario in which capital is specific, the profit rates enter fully r_p or r_r. B^o represents holdings of private sector bonds. The initial holdings are negative if the consumption group is a net debtor in the benchmark simulation; in that case an increase in p_b probably results in an increase in the supply of labor and a reduction in the expenditures of the quintile. Families also receive public sector transfers represented as the purchase by the govern-

4. By assumption, the capital installed in the tradable sectors cannot be reallocated.

ment of a service with an inelastic supply R^o at price p_R. Income from private sector bonds, $P_b B^o(h)$, is not taxed.

Private Firms

Private firms are those for which there was no change in ownership.[5] They produce goods and services intended for intermediate and final consumption, as well as for export and investment. This differentiation is needed to be able to properly account for differences in the tax treatment of the various destinations (for instance, exporters do not pay the value added tax and benefit from discounts on their gross income tax). However, there is no technological differentiation across these sectors. In other words, the production function is the same for a specific product (say food) used at different stages of the production process (intermediate, final, or export).

Exporters of goods are price takers abroad, and exports of services are price inelastic (that is, their supply is constant). Nontradable prices are determined as solution variables and adjust with factor income until markets are in equilibrium.

The profit function for a private firm is

$$(3) \qquad N^p = \{p - ap_b - \alpha_{pE}\,[zr_E + (1 - z)\,r_C] - f(1 + t_i) - f_m(1 + t_m)p_m\}Q^p - wL_p(1 + t_{v1}) - r_pK_p(1 + t_{v2})$$

and for exporters, it can be adjusted as

$$(4) \qquad N^{px} = \{p_x - ap_b - \alpha_p[zr_E + (1 - z)r_C] - f(1 + t_i) - f_m\,(1 + t_m)p_m\}\,X^p - (wL_{px} + r_pK_{px})$$

where the parameter a is the credit requirement per unit of output, and α_p is the quantity of services provided by the privatized company to obtain a unit of output. The amount $1 - z$ is the share of privatized services required per unit of output purchased through distribution companies at price r_C, where z is the share purchased on the wholesale market at price r_E. Purchases of electricity in the wholesale market correspond to generation; purchases on the retail market correspond to distribution.[6] L_p is employment in the private sector that produces goods and services for the domestic market, while L_{px} is employment in the export sector. L_r is employment in the privatized sector.

Interindustrial transactions in these simplified expressions are represented by a coefficient f for national goods and f_m for imported intermediate inputs. These requirements are proportional to total production Q^p and to exports X^p, respectively. Privatized goods and services are also proportional to output, which is different from the assumption made for consumers in situations where rationing could take place. However, firms, like consumers, can be subject to adjustment in the quality of services and hence can face different costs for the same ser-

5. However, YPF, the former public oil company, was considered a private firm.

6. Although the model assumes no substitutability between the two types of inputs, some evidence in other countries suggests that this may be a strong assumption (see Seitz 1994).

vice.[7] An improvement in the quality of service is represented by a reduction in parameter α, that is, $\alpha'(\cdot) < 0$. If **A** is the $n \times n$ input-output matrix, this improvement in quality is measured indirectly through its effect on the increase in productivity of the input requirements. Remuneration r_p includes total payments to capital and hence amortization. Saving and investment decisions are made by households. The tax t_{v1} corresponds to the value added tax and to the labor taxes collected at the firm level, while t_{v2} corresponds to similar taxes on capital. To simplify, taxes on labor and capital that are levied on exports are not included here.

The product combines intermediate inputs and value added in fixed proportions. The value added itself is obtained by combining labor and capital inputs in a constant elasticity of substitution production function:

$$VA_p = F(L_p, K_p) = (b_1 L_p^k + b_2 K_p^k)^{1/k} \tag{5}$$

where k is the elasticity of substitution of labor and capital, and b_1 and b_2 are distribution parameters used to calibrate the model.

The value added function for exports is similar:

$$VA_{px} = F(L_{px}, K_{px}) = (b_3 L_{px}^k + b_4 K_{px}^k)^{1/k}. \tag{6}$$

More generally, the product of sector j, Q_{pj}, is obtained from a fixed coefficient function (Leontief) between intermediate consumption and value added:

$$Q_{pj} = \min \{Q_{1j}/a_{1j}, \ldots, Q_{nj}/a_{nj}, VA_{pj}/av_j\} \tag{7}$$

where Q_{ij} is the quantity of good i consumed in producing j.

Privatized Utilities

The privatized firms sell mostly to the domestic market. With the exception of some differentiation due to regulation, service obligations, or taxes, each utility sector is assumed to sell a single product. Their profit function includes any subsidy TG that could be transferred by the public sector. It is written as

$$\begin{aligned} N^r = {} & r_C Q_C + r_E Q_E + r_G Q_G - \{a^r p_b + \alpha_r[z r_E + (1 - z) r_C] \\ & + f(1 + t_i) + f_m(1 + t_m)\, p_m\}\, (Q_C + Q_E + Q_G) - w L_r(1 + t_{v1}) - r_r K_r(1 + t_{v2}) + TG \end{aligned} \tag{8}$$

where Q_C is the quantity of products sold to households at unit price r_C, Q_E is the quantity of goods and services sold to the firms at price r_E, and the index G is used for the public sector wherever a distinction is relevant. This also allows a differentiation of tariffs into retail, wholesale, or commercial and residential as necessary. The quality variables are modeled as an improvement in the overall efficiency of the sector, and TG is modeled as a subsidy to capital set equal to zero or prescribed to shrink to zero as spelled out in the privatization documents. TG is used as an adjustment variable (a fine-tuning variable) to ensure that the rate of return in the regulated sector continues to be consistent with the rate of

7. This assumes that there is no possibility of using homemade substitutes for infrastructure services.

return in the rest of the economy. Although this is an income transfer, it does not generate significant distortions. First, the transfer goes to sector-specific capital, and hence there is no reallocation across sectors. Second, although transfers go to the highest income group, their effect is offset by the reduction in other public expenditures to the same income group. Third, the amounts involved are quite small compared with the total public resources to be allocated.

All outputs are limited by capacity and transmission constraints incorporated through the value added function. The product of the privatized sector is also based on a fixed-proportion production function:

$$Q_{ri} = \min \{Q_{1i}/a_{1ri}, \ldots, Q_{ni}/a_{nri}, VA_{ri}/av_{ri}\} \tag{9}$$

where a_{jri} is the input requirement of j by privatized firm ri.

The value added functions in the privatized sector are assumed to be Cobb-Douglas.

$$VA_{ri} = A\, L_{ri}^{a}\, K_{ri}^{1-a} \tag{10}$$

where A is a constant. The installed capital of the firm is taken as given:

$$K_{ri} = K_{ri}^{o}. \tag{11}$$

Price regulation is modeled as $RPI - X$, where X is set to 0 at the beginning of the contract. This implies that the r_C is

$$r_C / r_C^{o} = (PQ^{o} / P^{o}Q^{o} - X)\, \beta$$

where P is the price vector of private and privatized domestic goods that make up the Laspeyres index of retail prices in the base year with weights given by Q^{o}, and β is a correction coefficient for the tariffs (with $\beta = 1$ in the benchmark scenario).

The Public Sector

The government maximizes social welfare y, which is a function of current collective goods H produced with goods purchased from the private sector G, goods purchased from the privatized sector G_r, and government employment L_g; bonds B_g (which can be sold domestically or internationally); retiree services R; and public investment I_g:

$$y = y[H(G,G_r,L_g), B_g, R, I_g]. \tag{12}$$

The function $y(.)$ is Cobb-Douglas and $H(.)$ is Leontief in G, L_g, and G_r, which includes all the privatized services in fixed proportions. Pensions, bond services, investments, and current operative expenses are a constant proportion of total government income in this model.

The government faces a budget constraint given by:

$$\begin{aligned} (13)\quad & t_i[f(pQ + p_xX) + pI^d + pc^d] + t_{v1}\, w(L_p + L_r) + t_{v2}(r_pK_p + r_rK_r) \\ & + t_m p_m f_m(Q + X) + t_m p_m c^m + t_d(wL + w_gS_g + rK^{o} + N^r + N^p - pI^d) + p_bB_g^{o} \\ & + \alpha_g\,(r_rK_{ro} + N^r) = p(G + I_g) + r_G\, G_r + w_gL_g + p_bB_g + p_RR + TG \end{aligned}$$

where $L = L_p + L_r + L_g$.

In this equation α_g is the participation of the public sector in the ownership of capital of the privatized utilities. This is an important parameter because, through α_g, the government is able to share monopoly rents.

The Rest of the World

The foreign consumer has a Cobb-Douglas utility function

$$u^F = u^F(M^c, X^c, B_x) \tag{14}$$

subject to the following constraints:

$$p_m M - z^* V^d = 0 \tag{15}$$

for imports M, produced with a single factor V^d at price z^*, and

$$p_x X^s - z^* V^x = 0 \tag{16}$$

for exports X, where V^x is the quantity of the foreign factor needed to produce X^s, a perfect substitute for Argentina's exports.

This foreign consumer faces the following budget constraint:

$$p_x X^c + p_m M^c + p_b B_x = p_b B_x^o + z^*(V^d + V^x) + (r_r^* K_{ro} + N^r) \tag{17}$$

that is, the foreign consumer's revenue comes from payments to V, from its share of capital in the privatized sector—and from bonds—and his expenditures are X^c in export markets and M^c in import markets.

Equation 18 sets export prices at the international level:

$$p_x X^a - pX = 0. \tag{18}$$

Considering that A_m and A_x are the foreign technological parameters, equations 19 and 20 determine a linear transformation curve abroad and fix the relative prices faced by Argentina:

$$M = V^d/A_m \tag{19}$$

$$X^s = V^x/A_x. \tag{20}$$

The Labor Market

Constraint 21 describes the imbalance in the labor market, and in the model it is replaced by equation 22, determining the salary in the private sector of the economy. The labor market for the public sector clears as shown by equation 23, accounting for the fact that S_g is an observation:

$$L_p + L_{px} + L_r \leq S \tag{21}$$

$$w = bw^* \tag{22}$$

$$L_g = S_g. \tag{23}$$

Parameter b is calibrated for the equilibrium salary in the economy so that the initial unemployment rate is equal to the observed unemployment rate. This value of b is then kept constant throughout the counterfactual exercises.

Investment Goods Industries

Investment goods industries are divided into two main categories: those providing capital goods for private firms and those constructing specific capital for each of the privatized utilities (electricity, gas, water, and telecommunications). This division allows us to recognize the differential impact of investment schedules established by regulatory contracts—for example, as network expansion commitments—on the rate of unemployment and the trade balance. Special effort was devoted to determine the input composition of each industry, but the model has not yet been fully exploited to estimate the social gains from investments in water and sanitation after privatization.

The Market for "Bonds"

The financial market in the model is simple compared with the sophistication of Argentina's financial sector, but it is sufficient to deal with the issues of interest here. There are fixed requirements of credit per unit of output in each production sector, including recently privatized utilities. Domestic consumers are separated into net debtors (the four poorest income brackets) and net creditors (the highest income bracket). The rest of the world is considered a net creditor. In the bond market debtors are issuers and creditors are subscribers. Recall that, according to equation 1, bonds are an argument in the household's utility function. These were financial transactions that had to be taken into account (this is particularly important for the consistency of the model).

The equilibrium condition for the bond market is therefore represented by:

$$(24)\quad B(h) + B_g + B_x + a(Q^p + X^p + I^p) + a^r\,(Q^C + Q^E + Q^G) = B^o(h) + B_g^o + B_x^o.$$

The information on sectoral and personal net financial positions was obtained from monetary authorities and estimated using purchases of durable goods and total capital holdings.

The domestic bond market equilibrates not only the internal credit disequilibria of families, but also the credit position of the government and of Argentina vis-à-vis the rest of the world. Internally, the first four quintiles sell "bonds" to the richest. A net increase in the demand for bonds thus reduces the purchasing power of the four poorest income groups. An increase in the price of bonds is compensated by a decline in the purchase of other goods and by an increase in the labor supply, which can contribute to an increase in unemployment. Firms demand bonds as a fixed proportion of their value added. For them an increase in the price of bonds implies a cut in the marginal product of labor, which in turns leads to a reduction in the demand for labor, adding to the unemployment problem.

Because the simulations of the model include both a positive level of unemployment and a commercial deficit, in addition to disequilibrium in the labor market, the rest of the world is financing consumption and domestic investment. For the bond market this means an increase in the demand for bonds issued by domestic agents and purchased by foreigners. If foreigners did not accept Argentine bonds, it would be impossible to have an equilibrium between total savings and total investments. With an increase in the international interest rate, as in the case of the Tequila Effect, foreign investors stop buying domestic bonds. Between October 1993 and October 1995 the LIBOR increased from 3.4 to 5.8 percent and the PRIME from 6 to 7.8 percent, while the domestic interest rate increased from 9 percent in October 1993 to 14 percent in November 1994, and to more than 33 percent in March 1995. Simultaneously, unemployment rose from 9.3 to 12.2 percent. The share of problem portfolio in total portfolio increased to more than 10 percent in the third quarter of 1994 and to more than 30 percent in the second quarter of 1995. This fact is used in the calibration of the model.

Two simulations are performed. The first assumes that tariffs on utilities are endogenous (within the limits imposed by regulation) so that productivity and quality gains are diffused throughout the economy. This would be the outcome expected under perfect regulation. The second simulation assumes fixed prices for utilities, which means that the gains from privatization are appropriated by the capital owners of the sector as a quasi-rent. This would be the outcome under ineffective regulation and is a lower bound for the gains from the private operation of utilities. The difference between the results of the first and second simulations provides an estimate of the potential quasi-rent for which the new owners are likely to fight, as well as an indication of the economic gains from effective regulation. An alternative interpretation is that the Walrasian solution illustrates what a full pass-through implies for the economy, while the fixed-price solution models a cost-plus regulation in which the "plus" factor is determined by the efficiency gains achieved by private operators or a price cap regulation in which the cap is equal to the price under public operation of the utility and productivity gains (the "x" factor in $RPI - x$) are set at 0 forever. With Walrasian prices these sectors cannot be financially sustainable without an explicit adjustment to their rate of return through some type of subsidy (TG in this case).

III. The Private Operation and Regulation of Utilities

The total gains from privatization are the sum of the effects of four changes:

- *Efficiency.* Reductions in inputs per unit of output modeled as decreases in a_{jri} in equation 9; the efficiency gains increase the capacity of the economy to generate a surplus (see Diewert 1985).
- *Productivity.* Increases in labor productivity modeled as a reduction in the relevant L_{ri} in equation 10. Productivity gains are computed as efficiency gains in work so that less employment is needed to obtain a given level of service.

- *Quality.* Improvements in quality measured as reductions in a_{jri} for all i, that is, reductions in the coefficients of the privatized inputs needed to produce one unit of output in other sectors.
- *Tariffs.* Regulated prices of privatized sectors modeled as observed changes in the price of utilities.

The measurement of these changes for each sector is based on the observations summarized in table 1. Unfortunately, no quality indicator could be estimated for the water sector.

The main purpose of the simulation is to track how these gains percolate through the economy along the following channels:

- Directly, through lower prices of the privatized services to final consumers.
- Indirectly, through lower input costs to industries using these services.
- Indirectly, through lower input prices for the privatized utilities themselves.
- Directly or indirectly through remuneration in factor markets.

Privatization increases labor productivity in utilities and reduces costs in sectors using utilities. But it also reduces input requirements of the utilities themselves, which buy 23 percent of value added in the manufacturing sector and 19 percent of value added in the service sector. Moreover, the interaction between utilities is significant as well. For example, the water sector is the largest client of the electricity sector.

But the effects of privatization depend on how private utilities are regulated. The benefit of effective regulation can be estimated by comparing the results from simulations assuming flexible prices—effective regulation—and simulations assuming fixed prices—ineffective regulation. Under effective regulation it is assumed that all domestic prices, including utility prices, adjust to clear the markets, except salaries, so there is unemployment in the model. The prices of tradable goods are fixed in foreign currency because Argentina is assumed to be a price taker in international markets. The capital market is somewhat peculiar because capital is sector specific and the rates of return are endogenous to each industry. Finally, the trade balance is offset in the bond market, and if the domestic economy requires financing, the prices of bonds increase. All of this implies that regulation is effective and that private providers of utility services are unable to take advantage of their monopolistic position to extract rents. So, this kind of simulation provides an upper bound for the gains from privatization in Argentina.

However, if the regulator is ineffective, rents could be significant. This can be simulated by keeping the prices of utility services fixed, assuming that any reduction in cost from reforms is captured by the private operator. The same rules as before determine the prices of tradables and nontradables, as well as employment in the labor market. Because the prices of the privatized utilities are mostly set in foreign currency, quantity variables are added to provide the required number of endogenous variables. This simulation provides not only estimates of maximum monopoly rents for private utilities but also a lower bound for the gains

Table 3. *Average Macroeconomic Effects of Private Management*
(percent)

	Electricity generation		*Electricity distribution*		*Gas*		*Water*		*Telecommunications*		*Total*	
Effect on	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*
GDP	0.05	0.10	0.17	0.21	0.36	0.31	0.02	0.00	0.07	0.19	0.70	0.79
Industrial production	–0.01	0.09	0.21	0.29	–0.07	0.20	–0.01	0.00	0.04	0.10	0.16	0.66
Unemployment (percentage change in unemployment rate)	0.00	–2.47	–1.08	1.17	–1.93	–6.76	–3.22	–2.36	6.75	3.21	2.35	–4.50
GDP per employment	0.09	–0.13	0.09	0.39	0.19	–0.42	–0.29	–0.22	0.88	0.60	1.01	0.32
Price of tradable per price of nontradable	–0.12	0.18	0.77	0.78	–0.33	0.64	–0.05	–0.02	0.22	0.88	0.49	2.48
Exports per import	0.09	0.67	–0.25	0.67	–2.95	0.42	–0.31	0.02	0.75	0.77	–2.47	2.52
Industrial exports	0.41	1.41	0.36	2.15	–6.84	–2.11	0.50	0.07	1.40	1.59	–4.91	2.72

Note: Values are measured in average percentage changes over base year 1993, except for unemployment, which is measured in absolute terms. "Good regulation" means that the regulators are effective and that prices are essentially flexible; "bad regulation" means that regulators are ineffective and that privatized companies keep all the rent from privatization.

Source: Authors' calculations.

from privatization. There is a major difference between the distributive effects in the two simulations because the distribution of ownership of capital is the key determinant of who receives rent.

IV. The Macroeconomic Effects

Table 3 summarizes the main macroeconomic results. Privatization of the gas sector has the greatest effect on gross domestic product (GDP). The smallest impact is realized from reform of the water sector, but this is probably because most of the gains would come from increased investments in this sector, which are not considered because of data problems. As for unemployment, reforms in gas and water lead to some decline even when the regulator performs poorly, while reforms in telecommunications increase unemployment. The impact of electricity reforms on unemployment depends on the effectiveness of the regulator but does not affect unemployment much in any case. Actual unemployment increased from 9.3 percent in 1993 to more than 18 percent in 1995. But besides privatization, Argentina was hit by the Tequila Effect at the end of 1994 and early 1995. This international shock can be captured through the net debt position of the industries and of the various income groups. These simulations are not reported here but are available on request from the authors. They show how an interest rate shock could lead to increases in the supply of labor and in costs, wiping out the cost reduction brought about by the reforms, which in turn could lead to reductions in the demand for labor. The two effects would lead to significant increases in unemployment, consistent with those observed between 1993 and 1995.

The predicted effects on labor productivity are surprising. Two factors must be considered: when employment rises in a sector, marginal productivity declines. And when output shifts to more labor-intensive sectors, average labor productivity declines. The less effective are regulators, the larger are the gains in labor productivity. In fact, gains in labor productivity under an ineffective regulator are three times larger than under an effective regulator. This is due largely to the gas sector, where dispersing the efficiency gains leads to a significant drop in labor productivity in the economy by shifting production to more labor-intensive sectors and reducing overall unemployment. The combination of these two effects explains why labor productivity ends up lower with a good regulator than with a bad regulator.

The effects on trade are clearer and closer to expectations. The utility reform has little impact on imports because there is little change in the sources of capital in these sectors. The effect on exports depends on the effectiveness of regulation. If effective, exports increase; if not, they decrease. Similarly, when rents are retained by private operators, the relative price of tradables increases only by one-fifth of what it increases when regulators are effective.

The most important result presented in table 3 is that the macroeconomic benefits from privatizing utilities in Argentina are significant and that gains are

Table 4. *Decomposition of Sector-Specific Distributional Effects*
(percent)

	Electricity distribution		*Gas*		*Water*		*Telecommunications*		*Total*	
Item	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*	*Bad regulation*	*Good regulation*
Gini	0.01	0.00	–0.05	–0.22	–0.06	–0.06	–0.06	0.07	–0.06	–0.24
EV for quintile 1 (poorest)	0.29	0.41	0.54	1.00	0.13	0.09	0.08	0.21	1.19	1.99
EV for quintile 2	0.21	0.29	0.47	0.74	0.10	0.07	0.11	0.26	1.03	1.57
EV for quintile 3	0.18	0.21	0.51	0.65	0.10	0.07	0.11	0.26	1.05	1.38
EV for quintile 4	0.16	0.17	0.39	0.56	0.09	0.06	0.04	0.24	0.78	1.20
EV for quintile 5 (richest)	0.25	0.32	0.43	0.45	0.00	–0.01	0.19	0.35	1.02	1.30
Average labor income	0.40	0.40	–0.19	0.33	–0.03	–0.01	0.12	0.49	0.24	1.29
Average capital income	0.44	0.56	0.51	0.71	0.01	0.00	0.54	0.17	1.60	1.68

Note: Values are measured in percentage changes over base year 1993, except for unemployment, which is measured in absolute terms. Gini and average factor income are expressed as percentage changes over the base year. The equivalent variation (EV) is in terms of total income of the quintile. "Good regulation" means that the regulators are effective and that prices are essentially flexible; "bad regulation" means that regulators are ineffective and that privatized companies keep all the rent from privatization.

Source: Authors' calculations.

larger when prices are flexible, that is, under effective regulation. This does not mean that there were no problems in distributing gains among different income classes, the government, and foreign owners.

V. Distributional Effects

There are many ways of looking at the distributional implications of the reforms. One way is to compare factor incomes. The most standard way is to compute the change in the Gini coefficient. More revealing, however, is to compute the impact on families' income levels in terms of some welfare indicator. In this article the impact is computed in terms of an equivalent variation adapted to measure the effect of changes in prices as well as in quality.

Consider $v(p, M, \gamma)$, the indirect utility function of the representative agent, which depends on the price vector p, the agent's revenue M, and a quality or a quantity variable γ, which can also represent rationing of a service. If, as a result of a change in policy, the price vector with initial value po becomes lower, say $p1$, the equivalent variation EV is computed as:

$$v(po, M + EV, \gamma) = v(p1, M, \gamma).$$

The equivalent variation is the variation in income that keeps the consumer at the same level of utility he or she would achieve from a price reduction at the initial income level. In other words, it is the amount the consumer would have to receive to make him or her accept the change in price. A similar approach can be used to assess the impact of an improvement in quality. Also, the equivalent variation can be computed for the equivalent monetary compensation of an improvement in quality or for an increase in access to a public service.

The welfare changes due to privatization for each income class depend on the relative importance of the cost of services provided by privatized sectors in different household budgets and the distribution of factor ownership across income classes. They can be measured as percentage changes in the Gini coefficient or in an equivalent variation.

Table 4 shows the distributional implications of privatization reforms for each sector individually and for all reforms together. It shows that privatization improves the overall distribution of income, as indicated by the negative sign on the Gini coefficient. The overall improvement in the Gini coefficient, however, is six times larger when regulation is effective. The largest gains are also for the poorest, as indicated by the highest equivalent variation for that group. But the distribution of gains in equality is different when regulation is not effective. This is because under ineffective regulation average gains in labor income, the major source of wealth among the poorest, is only about one-fifth of what it would be under effective regulation. Also, although privatization reforms increase both average labor and capital income, average gains in capital income, particularly under poor regulation, are greater than gains in average labor income. This might

lead those with large capital incomes to push hard for privatization but not for effective regulation.[8]

The poorest stand to gain the most from improvements in gas and electricity—major inputs in their consumption basket. They also stand to gain relatively more from improvements in water, although their main source of gain—access—is not included here. The middle-income classes stand to gain the most from improvements in telecommunications, but only if the regulator is effective. Otherwise, they end up paying a huge rent to the private operators of the services.

VI. Conclusions

It may be useful to provide some dollar estimates of the effects of the reforms. Table 5 presents the general equilibrium calculation of the levels and distribution of gains across income classes from the efficiency and quality improvements due to the privatization process and the gains that could be achieved through effective regulation.

The key results are:

- The spillover effects from the private operation of utilities represent about $2.3 billion or 0.9 percent of Argentina's GDP, and their distribution benefits all income groups. On average these gains represent the equivalent of 41 percent of what households spend on utility services, even when ineffective regulation allows new owners to keep as much as possible of these gains as rents.
- The gains from effective regulation add up to almost $1 billion or 0.35 percent of GDP. This represents 16 percent of the average utility bill. The size of the effect also indicates why private operators with some degree of monoply power in any country have a strong incentive to contest any decision by regulators that forces them to share rents with the rest of the economy.
- The direct gains are significantly higher for the higher income classes (59 percent compared with 29 percent for the poorest). This is because when regulation is not effective, the gains from privatization are turned into a quasi-rent captured by the richest, who are the largest domestic owners of capital in infrastructure services. Part of these gains is also captured by foreign consumers and by the government, because they own a large share of the "privatized" assets.
- The indirect gains achieved through effective regulation, in contrast, tend to favor the poorest income classes somewhat more, even though all share in the gains from efficient regulation. This suggests that how serious governments are about the fair distribution of gains from privatization reform is revealed by how serious they are about regulation.

8. The public sector is, in fact, a partner of the privatized firms and could also have an incentive not to press for effective regulation, because it shares in the rent.

Table 5. *Gains from Private Operation of Public Utilities*

Income quintile	*Savings from operational gains*[a] *(millions of 1993 U.S. dollars)*	*Expenditure on utilities*[a] *(percent)*	*Savings from effective regulation*[b] *(millions of 1993 U.S. dollars)*	*Expenditure on utilities*[b] *(percent)*
1 (poorest)	197	29	138	20
2	259	31	142	17
3	373	37	121	12
4	403	32	214	17
5 (richest)	1,047	59	302	17
Total	2,279	41	915	16

Note: These figures represent annual gains.

a. Figures are the equivalent variation computed in terms of the dollar revenue of each income class. They are calculated by applying the total gains in the fixed-price simulation to the income in the base year. In net present value and over a period of 10 years, the gains represent a total varying between $8.2 billion and $14.4 billion with discount rates varying between 12 and 18 percent and amortization rates between 0 and 10 percent. The gains from efficient regulation under similar assumptions vary between $3.3 billion and $5.8 billion.

b. Figures are computed by applying the differences in gains between the fixed-price and the flexible-price simulations.

Source: Authors' calculations.

In sum, these general equilibrium estimates suggest extremely high economic rates of return for both privatization and regulation projects, whether distributional weights are considered or not. Another key result is that the significant increase in unemployment observed in Argentina between 1993 and 1995 is unlikely to be due to the privatization of utilities. On the contrary, privatization probably increased employment and generated significant gains for the economy and all income classes.

References

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

Bacharach, Michael. 1970. *Biproportional Matrices and Input-Output Change.* Cambridge, U.K.: Cambridge University Press.

Bergman, Lars. 1990. "The Development of Computable General Equilibrium Modeling." In Lars Bergman, Dale Jorgenson, and Erno Zalai, eds., *General Equilibrium Modeling and Economic Policy Analysis.* Cambridge, Mass.: Basil Blackwell.

Burns, Philip, and T. G. Weyman-Jones. 1994. "Cost Drivers and Cost Efficiency in Electricity Distribution: A Stochastic Frontier Approach." Technical Paper 2. Centre for the Study of Regulated Industries, London. Processed.

Button, Kenneth, and T. G. Weyman-Jones. 1994. "Impact of Privatization Policy in Europe." *Contemporary Economic Policy* 12(October):22–33.

Chisari, Omar, Antonio Estache, and Carlos Romero. 1997. "Winners and Losers from the Privatization and Regulation of Utilities: Lessons from a General Equilibrium Model of Argentina." Policy Research Working Paper 1824. World Bank, Policy Research Department, Washington, D.C.

Chisari, Omar O., and C. A. Romero. 1996. "Distribución del ingreso, asignación de rescursos, y shocks macroeconómicos—Un modelo de equilibrio general computado para la Argentina en 1993." Serie Financiamiento del Desarrollo 36. United Nations, Comisión Económica para América Latina y el Caribe, Santiago de Chile. Processed.

Crampes, Claude, and Antonio Estache. 1997. "Regulating Water Concessions—Lessons from the Buenos Aires Concession." In International Forum for Utility Regulation and World Bank Group, "The Private Sector in Infrastructure: Strategy, Regulation, and Risk." World Bank, Finance, Private Sector, and Infrastructure Network, Washington, D.C. Processed.

Diewert, W. E. 1985. "The Measurement of Waste and Welfare in Applied General Equilibrium Models." In John Piggott and John Whalley, eds., *New Developments in Applied General Equilibrium Analysis*. Cambridge, U.K.: Cambridge University Press.

Estache, Antonio, and Martín A. Rodríguez-Pardina. 1997. "Regulatory Lessons from Argentina's Power Concessions." In International Forum for Utility Regulation and World Bank Group, "The Private Sector in Infrastructure: Strategy, Regulation, and Risk." World Bank, Finance, Private Sector, and Infrastructure Network, Washington, D.C. Processed.

Galal, Ahmed, and Mary Shirley, eds. 1994. *Does Privatization Deliver? Highlights from a World Bank Conference*. EDI Development Studies.Washington, D.C.: World Bank.

Seitz, Helmut. 1994. "Public Capital and the Demand for Private Inputs." *Journal of Public Economics* 54(2, June):287–307.

Shaikh, Abdul H. 1996. "Argentina's Privatization Program: A Review of Five Cases." World Bank, Private Sector Development Department, Washington, D.C. Processed.

Shoven, J. B., and John Whalley. 1992. *Applying General Equilibrium*. Cambridge Surveys of Economic Literature. Cambridge, U.K.: Cambridge University Press.

Spiller, Pablo T., and Luis Viana Mantorell. 1996. "How Should It Be Done? Electricity Regulation in Argentina, Brazil, Uruguay, and Chile." In Richard J. Gilbert and Edward Khan, eds., *International Comparisons of Electricity Regulation*. Cambridge, U.K.: Cambridge University Press.

Winston, Clifford. 1993. "Economic Deregulation: Days of Reckoning for Microeconomists." *Journal of Economic Literature* 31(3, September):1263–89.

THE WORLD BANK ECONOMIC REVIEW, VOL. 13, NO. 2: 379-408

Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence

Aslı Demirgüç-Kunt and Harry Huizinga

Using bank-level data for 80 countries in the years 1988–95, this article shows that differences in interest margins and bank profitability reflect a variety of determinants: bank characteristics, macroeconomic conditions, explicit and implicit bank taxation, deposit insurance regulation, overall financial structure, and underlying legal and institutional indicators. A larger ratio of bank assets to gross domestic product and a lower market concentration ratio lead to lower margins and profits, controlling for differences in bank activity, leverage, and the macroeconomic environment. Foreign banks have higher margins and profits than domestic banks in developing countries, while the opposite holds in industrial countries. Also, there is evidence that the corporate tax burden is fully passed onto bank customers, while higher reserve requirements are not, especially in developing countries.

As financial intermediaries, banks play a crucial role in the operation of most economies. Recent research, as surveyed by Levine (1997), shows that the efficacy of financial intermediation can affect economic growth. Crucially, financial intermediation affects the net return to savings and the gross return to investment. The spread between these two returns mirrors bank interest margins, in addition to transaction costs and taxes borne directly by savers and investors. Thus bank interest spreads could be interpreted as an indicator of the efficiency of the banking system. In this article we investigate how bank interest spreads are affected by taxation, the structure of the financial system, and financial regulations, such as deposit insurance.

A comprehensive review of the determinants of interest spreads is offered by Hanson and Rocha (1986), who summarize the role that implicit and explicit taxes play in raising spreads and discuss some of the determinants of bank costs and profits, such as inflation, scale economies, and market structure. Using aggregate interest data for 29 countries in the years 1975–83, the authors find a positive correlation between interest margins and inflation.

Recently, several studies have examined the impact of international differences in bank regulation using cross-country data. Analyzing interest rates in 13 Organisation for Economic Co-operation and Development (OECD) countries in

Aslı Demirgüç-Kunt is with the Development Research Group at the World Bank, and Harry Huizinga is with the Department of Economics, Tilburg University, and the Centre for Economic Policy Research in London. The authors gratefully acknowledge comments by three anonymous referees. They also thank Anqing Shi for excellent research assistance and Paulina Sintim-Aboagye for help with the manuscript.

the years 1985–90, Bartholdy, Boyle, and Stover (1997) find that the existence of explicit deposit insurance lowers the deposit interest rate by 25 basis points. Barth, Nolle, and Rice (1997) use 1993 data from 19 industrial countries to further examine the impact of banking powers on bank return on equity, controlling for several bank and market characteristics. They find that variations in banking powers, bank concentration, and the existence of explicit deposit insurance do not significantly affect the return on bank equity.

In this article we extend the existing literature in several ways. First, we use bank-level data for 80 industrial and developing countries in 1988–95 to provide summary statistics on the size and decomposition of bank interest margins and profitability. Second, we use regression analysis to examine the underlying determinants of interest spreads and bank profitability. This empirical work enables us to infer the extent of taxation and regulation on bank customers and on banks themselves.

Apart from covering many banks in many countries, this study is unique in its coverage of the determinants of interest margins and profitability. These determinants include a comprehensive set of bank characteristics (such as size, leverage, type of business, foreign or domestic ownership), macroeconomic indicators, taxation and regulatory variables, financial structure variables, and legal and institutional indexes. Among these, the ownership variable, the tax variables, some of the financial structure variables, and the legal and institutional indicators have not been included in any previous study in this area. To check whether some of these determinants affect banking differently in developing and industrial countries, we interact these variables with the country's gross domestic product (GDP) per capita.

I. Bank Interest Spreads and Profitability

We can measure the efficiency of bank intermediation using both ex ante and ex post spreads. The ex ante spread is the difference between the contractual rates charged on loans and rates paid on deposits. The ex post spread is the difference between banks' actual interest revenues and their actual interest expenses. The ex post spread differs from the ex ante spread by the amount of loan defaults. The ex post spread is a more useful measure because it controls for the fact that banks with high-yield, risky credits are likely to face more defaults. An additional problem with using the ex ante spread is that data are generally available at the aggregate industry level and are put together from a variety of sources. Thus they are not completely consistent. For these reasons we focus on ex post interest spreads in this article. There is, however, a problem with ex post spreads, in that the interest income and loan loss reserving associated with a particular loan tend to materialize in different time periods.

As a measure of what we call bank "efficiency," we consider the accounting value of a bank's net interest income divided by total assets (*TA*), or the net interest margin (*NIM*). Bank "profitability" is a bank's before-tax profits (*BTP*) divided by total assets. Profitability could also be measured by the return on

equity as opposed to the return on assets. It is well known that, ceteris paribus, a bank with a higher equity ratio will have a higher return on assets and a lower return on equity than a bank with a lower equity ratio. The problem in some developing countries is that banks operate with extremely low equity capital, often supported by implicit state guarantees, which inflates their return on equity. Using unadjusted returns on equity may then be more distortionary than using returns on assets. Ideally, we should use risk-adjusted returns on equity, but since these are not available, we analyze returns on assets after controlling for the banks' equity ratio. We do this by entering the equity ratio as an independent variable in the profit regression. Thus, by straightforward accounting,

$$\frac{BTA}{TA} \equiv \frac{ATP}{TA} + \frac{TX}{TA} \tag{1}$$

where ATP is after-tax profits. From the bank's income statement, before-tax profits divided by total assets further satisfies the following accounting identity:

$$\frac{BTA}{TA} \equiv NIM + \frac{NII}{TA} - \frac{OV}{TA} - \frac{LLP}{TA} \tag{2}$$

where NII is noninterest income, OV is overhead, and LLP is loan loss provisioning. NII/TA reflects the fact that many banks also engage in nonlending activities, such as investment banking and brokerage services, OV/TA accounts for the bank's entire overhead associated with all of its activities, and LLP/TA measures actual provisioning for bad debts.

Although the net interest margin can be interpreted as a rough index of bank efficiency, this does not mean that its reduction always signals improved efficiency. A reduction in the net interest margin can, for example, reflect a reduction in bank taxation or, alternatively, a higher loan default rate. In the first instance the reduction in the net interest margin may reflect the improved functioning of the banking system, while in the second case the opposite may be true. Also, variation in an accounting ratio, such as the net interest margin, may reflect differences in net interest income (the numerator) or differences in, say, nonlending assets (a component of the denominator).

In the data set the accounting data are organized so as to be comparable internationally. However, differences in accounting conventions regarding the valuation of assets, loan loss provisioning, hidden reserves, and so on may remain. Vittas (1991) reviews the pitfalls in interpreting bank operating ratios. Accounting data also tend to reflect economic realities with a long lag so that they are not able to flag pending banking crises, such as those that have recently occurred in Southeast Asia.

This article focuses on accounting measures of income and profitability as investors equalize (risk-adjusted) financial returns on bank stocks in the absence of prohibitive barriers. Gorton and Rosen (1995) and Schranz (1993) also focus

on accounting measures of profitability when examining managerial entrenchment and bank takeovers.

The above accounting identity (equation 2) suggests a useful decomposition of the realized interest spread—the net interest margin—into its constituent parts: noninterest income, overhead, taxes, loan loss provisions, and after-tax bank profits. Hanson and Rocha (1986) take this approach, with some modifications. As a first step to analyzing the data, in section III we provide an accounting breakdown of the net interest margin for individual countries and for selected aggregates. Although it may be misleading to compare accounting ratios without controlling for differences in the macroeconomic environment in which banks operate and the differences in their business, product mix, and leverage, these breakdowns are still a useful initial indicator of differences across countries.

Next, controlling for bank characteristics and the macroeconomic environment, we provide an economic analysis of the determinants of the interest and profitability variables—the net interest margin and before-tax profits divided by total assets. This empirical work offers insights into how bank customers and banks themselves are affected by these variables. The net interest margin regressions tell us how the spread determinants affect the combined welfare of depositors and lenders. The relationship between the interest spread and a bank's corporate taxes, for instance, reveals the extent to which a bank is able to shift its tax bill forward to its depositors and lenders.

Generally, taxes and other variables can affect interest rates as well as the volume of loans and deposits. In the short term the major effects may come through pricing changes, in which case the net interest margin and before-tax profits as a share of total assets immediately reveal easily interpreted welfare consequences for banks and their customers. With market imperfections in the form of credit rationing or imperfect competition in credit markets, changes in quantities generally have first-order welfare implications independent of changes in prices. We do not, however, evaluate changes in quantities in this article. Lastly, the before-tax profit regressions show how spread determinants affect bank shareholders.

The regression analysis starts from the following equation:

$$I_{ijt} = \alpha_o + \alpha_i B_{ijt} + \beta_j X_{jt} + \gamma_t T_t + \delta_j C_j + \varepsilon_{ijt} \tag{3}$$

where I_{ijt} is the dependent variable (either the *NIM* or *BTP/TA*) for bank i in country j at time t, B_{ijt} are characteristics of bank i in country j at time t, X_{jt} are characteristics of country j at time t, T_t and C_j are time and country dummy variables, and ε_{ijt} is a white-noise error term. We estimate several specifications of equation 2 including different bank and country variables.

II. Data

In this study we use income statement and balance sheet data of commercial banks from the BankScope database provided by IBCA. IBCA's coverage is compre-

hensive in most countries, accounting for 90 percent of all bank assets. We begin with all commercial banks worldwide, with the exception of France, Germany, and the United States, for which we include only several hundred commercial banks listed as "large." To ensure reasonable coverage for individual countries, we include only countries with at least three banks for a given year. We end up with a data set that includes 80 countries during the years 1988–95, with about 7,900 individual commercial bank observations. This data set includes all OECD countries, as well as many developing countries and transition economies (table 1).

Several countries, such as Luxembourg, the Netherlands, and Egypt, have a net interest margin close to 1 percent (column 2 of table 1). This is the low end of the distribution. Egypt's low net interest margin can be explained by a predominance of low-interest directed credits by the large state banking sector. Generally, developing countries, and especially Latin American countries, such as Argentina, Brazil, Costa Rica, Ecuador, and Jamaica, have relatively large spreads. This is also true for certain Eastern European countries, such as Lithuania and Romania.

Columns 3–6 in table 1 break down the net interest income into its four components: overhead minus noninterest income, taxes, loan loss provisioning, and net profits. Taxes as a share of net interest income (column 4) reflect the explicit taxes that banks pay (mostly corporate income taxes). Banks also face implicit taxation because of reserve and liquidity requirements and other restrictions on lending that come through directed or subsidized credit policies. These indirect forms of taxation directly lower the net interest income rather than the tax variable. Nonetheless, the tax variable indicates that there is considerable international variation in the explicit taxation of commercial banks. Several countries in Eastern Europe impose high explicit taxes on banking. (For example, taxes as a percentage of net interest income are only 17.5 in Lithuania and 13.7 in Hungary compared with 26.2 in Romania, 83.3 in Russia, and 23.2 in the Czech Republic.) The lowest share of taxes in net interest income is 0 for Qatar, where there is no significant taxation of banking. For some countries, such as Norway, Sweden, and Costa Rica, low tax shares reflect the tax deductibility of bad debts, which are plentiful.

Loan loss provisioning as a share of net interest income is a direct measure of differences in credit quality across countries (column 5). It also reflects differences in provisioning regulations. This variable is high for some Eastern European countries. It is also high for some industrial countries, such as France and the Nordic countries. The fourth component of net interest income is net profits (column 6). As a residual, net profits as a share of net interest income reflect the extent to which the net interest margin translates into net-of-tax profitability.

The remaining columns in table 1 tabulate the various accounting ratios (relative to total assets) in the accounting identity (equation 2). Noninterest income as a share of total assets reveals the importance of fee-based services for banks in different countries (column 7). Banks in Eastern Europe—for example, those in Estonia, Hungary, and Russia—seem to rely heavily on fee-based operations.

Table 1. *Bank Interest Spreads and Profitability: Economy Averages, 1988–95*
(percent)

Economy	Net interest margin[a]	Net interest income[b]: Overhead minus noninterest income	Taxes	Loan loss provisions	Net profits	Noninterest income	Overhead	Taxes	Loan loss provisions	Net profits
		(as a percentage of net interest income)				(as a percentage of total assets)				
Argentina	7.3	35.6	5.5	28.5	30.3	6.3	9.4	0.4	1.8	2.0
Australia	3.0	32.8	8.5	28.1	33.2	1.3	2.8	0.3	0.7	0.6
Austria	1.7	54.9	6.8	25.7	24.8	0.5	1.5	0.1	0.5	0.3
Bahrain	2.2	30.8	2.7	32.9	43.3	0.8	1.4	0.0	0.6	1.1
Belgium	2.0	48.7	10.2	20.4	24.0	0.9	2.1	0.2	0.4	0.4
Bolivia	3.1	73.7	1.9	12.6	12.6	2.0	5.2	0.6	0.6	-1.1
Botswana	6.0	43.7	11.8	12.9	31.5	2.8	5.4	0.7	0.7	1.9
Brazil	8.9	60.5	11.6	13.8	17.5	4.5	10.2	1.1	1.3	1.4
Canada	2.9	47.3	12.7	21.5	19.0	1.2	2.5	0.4	0.6	0.6
Chile	4.3	71.7	2.4	13.8	14.2	-0.1	3.0	0.1	0.6	0.5
China	2.1	30.0	15.9	—	54.2	1.0	1.6	0.3	—	1.2
Colombia	6.0	53.8	10.3	12.9	27.7	5.8	8.3	0.7	0.9	2.2
Costa Rica	13.6	40.7	4.7	57.6	9.9	3.5	8.1	0.8	5.7	3.5
Cyprus	1.0	59.1	11.1	15.8	22.0	3.1	3.2	0.3	0.3	0.5
Czech Rep.	3.3	13.5	23.2	53.4	13.6	1.5	2.1	0.6	2.0	0.3
Denmark	4.8	52.9	5.4	33.3	8.6	1.0	3.7	0.3	1.6	0.3
Dominican Rep.	6.6	52.8	8.6	9.1	30.9	3.1	6.3	0.6	0.5	2.3
Ecuador	7.7	52.8	4.7	12.8	34.9	3.8	8.1	0.4	1.0	2.5
Egypt	1.4	-32.7	11.2	62.6	63.5	2.1	1.4	0.3	0.7	1.2
El Salvador	3.2	34.3	1.9	14.0	49.8	1.6	2.9	0.1	0.4	1.5
Estonia	4.7	-35.9	24.1	—	111.7	8.7	7.0	1.1	—	5.3
Finland	1.8	50.1	9.4	55.6	-10.7	1.2	2.1	0.2	0.8	-0.1
France	2.4	48.3	7.0	50.8	-1.7	1.4	2.6	0.2	1.0	0.1
Germany	2.0	51.6	12.3	29.5	12.6	1.1	2.1	0.3	0.6	0.3
Greece	3.0	33.8	12.7	25.7	29.7	2.2	3.4	0.4	0.6	1.0
Guatemala	5.6	80.5	3.6	—	16.0	1.4	5.7	0.2	—	1.1
Haiti	2.8	53.8	7.3	12.4	26.5	2.8	4.2	0.2	0.4	0.8
Honduras	4.3	72.3	9.8	—	17.9	0.9	4.0	0.4	—	0.8
Hong Kong, China	2.5	17.1	10.2	6.0	67.8	1.3	1.4	0.3	0.2	2.0

Hungary	4.7	17.5	13.7	68.8	29.9	5.8	7.0	0.6	2.7	1.4
India	4.0	18.2	12.4	19.3	50.2	1.6	2.0	0.6	0.7	2.3
Indonesia	3.6	47.5	10.9	17.8	26.2	1.2	2.9	0.4	0.7	0.9
Israel	2.8	41.9	17.1	23.6	17.3	1.8	3.2	0.4	0.7	0.4
Italy	3.4	56.5	14.3	17.4	11.9	1.4	3.3	0.5	0.5	0.4
Jamaica	10.5	33.9	21.2	2.2	43.1	2.8	6.3	2.2	0.3	4.5
Japan	1.6	61.9	16.2	10.0	12.1	0.2	1.3	0.2	0.1	0.2
Jordan	2.1	48.0	10.4	24.1	24.5	1.4	2.4	0.2	0.5	0.5
Korea, Rep. of	1.8	36.4	12.5	34.0	29.9	1.5	2.2	0.2	0.5	0.5
Lebanon	2.7	45.6	9.2	13.7	35.3	0.9	2.1	0.3	0.5	0.9
Lithuania	10.6	29.8	17.5	81.7	–22.2	5.0	—	—	—	—
Luxembourg	0.8	–11.5	28.2	52.7	46.0	0.9	1.0	0.2	0.3	0.3
Malaysia	2.7	40.0	15.9	17.3	29.2	0.8	1.9	0.4	0.4	0.8
Malta	2.4	37.6	18.0	6.2	39.1	1.1	2.0	0.4	0.1	0.9
Mexico	4.6	40.8	6.1	42.2	15.4	2.1	4.5	0.3	1.1	0.9
Morocco	3.4	66.8	13.6	0.1	21.9	1.3	3.5	0.5	0.0	0.8
Nepal	3.6	10.5	25.3	16.1	48.1	2.1	2.4	1.0	0.5	1.8
Netherlands	1.4	43.1	9.7	21.4	26.1	1.0	1.7	0.1	0.3	0.4
Nicaragua	4.4	85.2	8.0	18.5	–10.4	3.3	6.3	0.3	0.9	0.2
Nigeria	5.3	–29.3	13.1	88.3	27.8	5.8	7.0	0.7	1.6	1.8
Norway	3.2	51.6	4.6	44.3	3.2	1.2	2.8	0.1	1.4	0.2
Oman	4.1	43.1	5.5	15.2	36.2	1.4	3.3	0.2	0.6	1.4
Pakistan	2.8	38.8	28.6	—	32.6	1.8	2.9	0.9	—	0.8
Panama	2.1	29.9	4.3	20.3	46.4	1.4	2.0	0.1	0.4	1.0
Papua New Guinea	3.2	–2.6	20.2	40.8	45.8	4.2	5.0	0.4	0.9	1.1
Paraguay	5.9	63.5	5.5	11.4	23.3	2.5	6.2	0.4	0.7	1.5
Peru	6.5	43.8	14.3	47.0	12.1	5.7	9.6	0.7	1.7	0.8
Philippines	4.1	29.8	6.6	10.3	55.0	3.0	4.3	0.3	0.4	2.2
Poland	6.1	16.8	27.9	23.3	34.9	2.4	3.6	1.6	1.3	2.1
Portugal	3.3	45.9	8.0	25.5	23.7	1.0	2.5	0.3	0.9	0.7

(continued on following page.)

Table 1. *(continued)*
(percent)

Economy	Net interest margin[a]	Net interest income[b]: Overhead minus noninterest income	Taxes	Loan loss provisions	Net profits	Noninterest income	Overhead	Taxes	Loan loss provisions	Net profits
		(as a percentage of net interest income)				(as a percentage of total assets)				
Qatar	1.9	6.6	0.0	15.0	85.2	1.1	1.3	0.0	0.2	1.6
Romania	9.7	1.9	26.2	36.8	44.3	2.4	2.8	2.3	3.7	4.3
Russia	4.7	–5.0	33.3	47.2	37.1	10.9	7.0	1.9	2.6	4.7
Singapore	2.2	20.7	21.6	8.7	56.4	1.0	1.4	0.5	0.1	1.3
South Africa	3.9	45.1	11.8	16.1	29.0	1.9	3.6	0.5	0.7	1.1
Spain	3.6	60.3	10.2	17.7	12.7	1.2	3.2	0.4	0.6	0.7
Sri Lanka	3.7	31.8	11.1	9.7	52.5	2.0	3.0	0.5	0.4	2.1
Swaziland	5.4	52.1	16.3	2.8	30.9	2.7	5.5	0.9	0.2	1.7
Sweden	3.1	26.3	1.9	64.6	11.2	1.5	2.5	0.1	1.9	0.3
Taiwan (China)	2.0	34.6	10.1	10.8	45.5	1.0	1.6	0.2	0.2	1.0
Tunisia	2.3	31.4	9.9	56.1	48.0	2.2	3.1	0.2	1.1	0.8
Turkey	6.3	11.7	10.0	32.9	47.2	4.0	5.4	0.8	0.8	3.3
United Kingdom	2.3	18.4	20.6	29.8	40.9	2.3	3.0	0.4	0.7	0.8
United States	3.9	47.6	12.5	15.2	25.8	1.8	3.6	0.5	0.7	1.0
Venezuela	7.2	49.9	2.7	16.7	30.6	2.8	6.4	0.2	1.0	2.5
Yemen	4.0	48.8	14.1	2.6	34.6	–0.5	1.4	0.6	0.1	1.4
Zambia	–4.7	186.1	–6.6	–49.1	–30.4	9.5	0.4	0.3	2.4	1.7

— Not available.

Note: Ratios are calculated for each bank in each country and then averaged over the country's sample period.

a. The net interest margin is defined as net interest income divided by total assets.

b. Columns 3 though 6 show the shares of the four components of net interest income. These shares add to 100 percent except for cases where information on loan loss provisioning is missing.

Source: Authors' calculations based on data from the BankScope database of the IBCA.

This is also the case in some Latin American countries, such as Argentina, Brazil, Colombia, and Peru, and in a few African countries, such as Nigeria and Zambia.

Overhead as a share of total assets reveals variations in operating costs across banking systems (column 8). This variable reflects variations in employment and in wage levels. Despite high wages, overhead as a share of total assets appears to be lowest at around 1 percent for high-income countries, such as Japan and Luxembourg. It is notably high at 3.6 percent for the United States, perhaps reflecting the proliferation of banks and bank branches because of banking restrictions.

Jamaica, Lithuania, and Romania stand out with high tax-to-asset ratios of around 2 percent (column 9). Loan loss provisioning, proxied by loan loss provisioning as a share of total assets, is equally high in Eastern Europe and in some developing countries (column 10). Finally, net profits divided by total assets also tend to be relatively high in developing countries (column 11).

Table 2 presents statistics on interest spreads and profitability for selected aggregates. The first breakdown is by ownership; a bank is said to be foreign-owned if 50 percent or more of its shares are owned by foreign residents. There is only a small difference in the net interest margin for domestic banks (3.7 percent) and foreign banks (2.9 percent). This small difference, however, masks the fact that foreign banks tend to achieve higher interest margins in developing countries and lower interest margins in industrial countries.[1] This may reflect the fact that foreign banks are less subject to credit allocation rules and have technical advantages (in developing countries), but also have distinct informational disadvantages relative to domestic banks (everywhere).

Foreign banks pay somewhat lower taxes than domestic banks (column 4). This gap may reflect differences in the tax rules governing domestic and foreign banks, as well as the ability of foreign banks to shift profits internationally to minimize their global tax bill. Foreign banks also have relatively low provisioning, as indicated by loan loss provisioning as a share of total assets, which is consistent with the view that foreign banks generally do not engage in retail banking operations.

The next breakdown is by bank size. For countries with at least 20 banks, large banks are defined as the 10 largest banks according to the value of their assets. Large banks tend to have lower margins and profits and smaller overheads. They also pay relatively low direct taxes and have lower loan loss provisioning.

Table 2 also considers bank groupings by national income levels and location.[2] Breaking down the data into four income levels, we see that the net interest margin is highest for countries in the middle-income groups. Banks operating in middle-income countries also have the highest values for overhead, taxes, and loan loss provisioning as shares of total assets. Net profits as a share of total

1. See Claessens, Demirgüç-Kunt, and Huizinga (1997) for more detailed information on the average spreads of domestic and foreign banks for different groupings of countries by income. That article also considers how entry by foreign banks affects the interest spreads and operating costs of domestic banks.

2. For country groupings by income, see World Bank (1996).

Table 2. *Bank Interest Spreads and Profitability, Selected Aggregates, 1988–95*
(percent)

Bank groupings	Net interest margin[a]	Net interest income[b]: Overhead minus noninterest income	Taxes	Loan loss provisions	Net profits	Noninterest income	Overhead	Taxes	Loan loss provisions	Net profits
		(as a percentage of net interest income)				(as a percentage of total assets)				
All banks	3.5	43.1	11.5	24.8	20.6	1.6	3.2	0.3	0.8	0.8
Bank ownership										
Domestic	3.7	46.2	11.1	22.8	19.9	1.6	3.3	0.4	0.8	0.8
Foreign[c]	2.9	29.0	13.1	33.5	24.4	1.6	2.8	0.3	0.7	0.8
Bank size[d]										
Large	2.6	35.5	13.1	27.5	23.9	1.2	2.5	0.3	0.6	0.5
Small	3.4	48.0	11.9	22.0	18.2	1.5	3.1	0.4	0.7	0.7
Country income										
Low	2.8	37.9	11.3	20.0	30.8	3.2	3.1	0.5	0.8	1.5
Lower-middle	5.7	36.8	11.0	24.9	27.2	3.2	5.1	0.7	1.3	1.8
Upper-middle	4.1	32.7	11.2	27.3	28.8	2.1	3.8	0.4	1.0	0.9
High	2.6	30.0	10.3	31.8	27.9	1.2	2.3	0.2	0.7	0.5
Region										
Africa	3.3	59.2	9.6	14.2	16.9	4.5	4.4	0.6	1.1	1.6
Asia	3.0	20.1	14.7	17.3	47.9	1.8	2.4	0.4	0.5	1.5
Latin America	6.2	48.7	6.8	21.1	23.4	3.1	6.2	0.5	1.1	1.5
Middle East and North Africa	2.9	26.1	8.5	23.4	41.9	1.6	2.6	0.3	0.5	1.1
Transition economies[e]	6.4	13.2	21.8	51.9	13.1	4.4	4.5	1.4	3.0	1.9
Industrial economies	2.7	32.9	10.4	34.7	21.9	1.2	2.5	0.3	0.8	0.4

Note: The data by income group and by region are means of country averages. Income and region classifications follow World Bank definitions as published in World Bank (1996).

a. The net interest margin is defined as net interest income divide by total assets.

b. Columns 3 though 6 show the shares of the four components of net interest income. These shares add to 100 percent.

c. A foreign bank is defined as having at least 50 percent foreign ownership.

d. Large includes the largest 10 banks; the remaining banks are classified as small. The large versus small distinction is made only if there are more than 20 banks in a given year.

e. The transition economies are China, Czech Republic, Estonia, Hungary, Lithuania, Poland, Romania, Russia, and Slovenia.

Source: Authors' calculations based on data from the BankScope database of the IBCA.

assets tends to be highest for banks operating in lower-income countries. Banks operating in higher-income countries, instead, achieve the lowest net interest margins, and they face the lowest values of overhead, taxes, loan loss provisioning, and net profits as shares of total assets.

The breakdown by region reveals that the net interest margin is highest for banks operating in the transition economies at 6.4 percent and is also high in Latin America at 6.2 percent. It is lowest for banks operating in industrial countries at 2.7 percent. The transition economies further stand out with high values of overhead, taxes, loan loss provisioning, and net profits as shares of total assets. Banks operating in industrial countries have the lowest ratio of net profits to total assets at 0.4 percent, probably because of the high level of competition in banking services.

Table 3 provides information on some of the macroeconomic and institutional variables used in the regression analysis. The data are for 1995 or the most recent year available. The tax rate is computed on a bank-by-bank basis as taxes paid divided by before-tax profits. The figure reported in the table is the average for all banks in the country in 1995. Reserves divided by deposits are the banking system's aggregate central bank reserves divided by aggregate banking system deposits. Actual reserve holdings reflect required and excess reserves. Reserves are generally remunerated at less-than-market rates, and therefore actual reserves may be a reasonable proxy for required reserves, as averaged over the different deposit categories. For several developing countries—Botswana, Costa Rica, Greece, and Jordan—the reserve ratio is above 40 percent, indicating substantial financial repression. In contrast, this ratio is low in Belgium, France, and Luxembourg at 0.01.

The deposit insurance variable is a dummy variable that takes on a value of 1 if there is an explicit deposit insurance scheme (with defined insurance premia and insurance coverage) and a value of 0 otherwise. Even if there is an explicit deposit insurance scheme, however, the ex post insurance coverage may prove to be higher than the de jure coverage, if the deposit insurance agency chooses to guarantee all depositors. With a value of 0 there is no explicit deposit insurance, even if the authorities offer some type of implicit insurance.

Next, table 3 presents some indicators of the structure of financial markets. The concentration variable is defined as the ratio of the assets of the three largest banks to the assets of the total banking sector. As is well known, the concentration of the U.S. banking market is low, at 16 percent, compared with values of about 50 percent for France and Germany. Note, however, that the U.S. figure may understate the concentration ratio in individual banking markets, which are protected from outside competition by interstate banking and branching restrictions. The number of banks in the table reflects the number of banks in the data set with complete information. The ratio of bank assets to GDP is defined as the total assets of the deposit-money banks divided by GDP. This ratio reflects the banking sector's overall level of development. The ratio of stock market capitalization to GDP measures the extent of stock market development. Developing coun-

Table 3. *Economic and Institutional Indicators*

Economy	*GDP per capita (U.S. dollars)*	*Inflation*[a] *(percent)*	*Tax rate*[b] *(percent)*	*Reserves/ deposits*	*Deposit insurance*[c]	*Market concentration*[d]	*Number of banks*[e]	*Bank assets/ GDP*[f]	*Stock market capital/GDP*	*Law and order*[g]
Argentina	3,825	0.02	0.15	0.06	1	0.48	11	0.23	0.13	5
Australia	14,542	0.03	0.47	0.02	0	0.45	44	0.77	0.70	6
Austria	16,947	0.02	0.23	0.04	1	0.75	12	1.27	0.14	6
Bahrain	7,902	0.02	0.01	0.10	—	0.94	7	0.49	—	5
Belgium	16,197	0.03	0.27	0.01	1	0.46	49	1.53	0.39	6
Bolivia	665	0.11	0.10	—	0	0.57	11	0.43	0.01	3
Botswana	1,844	0.18	0.32	0.41	—	0.94	5	0.14	0.09	—
Brazil	2,113	22.95	0.38	0.13	0	0.43	56	0.32	0.21	3
Canada	16,091	0.02	0.35	0.01	1	0.56	72	0.72	0.65	6
Chile	2,481	0.12	0.07	0.08	1	0.40	23	0.45	1.10	5
China	468	0.12	0.21	—	0	0.99	5	0.80	0.06	—
Colombia	1,445	0.23	0.18	0.23	1	0.35	28	0.20	0.22	2
Costa Rica	1,936	0.23	0.09	0.59	0	0.76	22	0.15	0.07	—
Cyprus	7,500	0.03	0.34	0.16	—	0.75	9	0.91	0.30	5
Czech Rep.	3,165	0.15	2.01	0.18	1	0.76	15	0.87	0.33	—
Denmark	22,386	0.02	0.16	0.07	1	0.77	56	0.55	0.33	6
Dominican Rep.	829	0.13	0.19	0.25	1	0.62	13	0.17	—	4
Ecuador	1,243	0.23	0.12	0.10	0	0.89	6	0.24	0.15	4
Egypt	709	0.12	0.25	0.20	0	0.89	9	0.65	0.13	4
El Salvador	994	0.12	0.04	0.32	1	0.86	4	0.28	0.67	3
Estonia	2,820	0.29	0.20	—	0	—	7	—	—	—
Finland	18,275	0.03	0.48	0.15	1	0.70	12	0.70	0.35	6
France	18,128	0.01	0.26	0.01	1	0.48	98	0.99	0.34	6
Germany	16,572	0.02	0.56	0.04	1	0.50	82	1.19	0.24	6
Greece	5,140	0.10	0.21	0.46	1	0.70	16	0.40	0.15	6
Guatemala	898	0.10	0.10	0.32	0	0.29	24	0.16	—	3
Haiti	230	0.72	0.25	0.35	—	1.00	3	0.10	—	3
Honduras	900	0.25	0.33	0.14	0	1.00	3	0.22	0.09	3
Hong Kong, China	11,911	0.02	0.13	—	0	0.44	35	—	2.17	6
Hungary	2,330	0.20	0.13	—	1	0.40	22	0.40	0.04	—

India	423	0.06	0.04	0.16	1	0.90	5	0.35	0.38	4
Indonesia	718	0.13	0.30	—	0	0.38	21	0.41	0.33	5
Ireland	13,653	0.01	0.26	0.05	1	0.75	12	0.46	0.42	6
Israel	10,515	0.10	0.53	0.07	0	0.48	26	0.91	0.42	5
Italy	15,491	0.05	0.48	—	1	0.27	66	0.62	0.11	6
Jamaica	1,573	0.33	0.24	0.33	0	0.52	10	0.30	0.41	3
Japan	23,960	0.02	0.57	0.01	1	0.21	81	1.32	0.72	6
Jordan	1,263	0.02	0.31	0.49	0	0.93	7	0.70	0.70	5
Korea, Rep. of	5,663	0.06	0.26	0.10	0	0.17	43	0.55	0.40	5
Lebanon	1,800	—	0.24	0.16	1	0.61	6	0.79	—	—
Lithuania	1,233	0.36	0.37	0.14	0	0.76	8	0.17	0.03	—
Luxembourg	21,433	0.07	0.45	—	1	0.30	108	0.41	0.14	6
Malaysia	3,108	0.06	0.32	0.12	0	0.31	49	0.84	2.82	5
Malta	6,102	0.04	0.29	0.08	—	0.69	7	0.75	—	—
Mexico	1,749	0.45	0.20	0.23	1	0.59	20	0.35	0.32	3
Morocco	853	0.07	0.34	0.07	—	0.63	8	0.46	0.18	6
Nepal	203	0.07	0.35	—	—	1.00	3	0.22	0.06	—
Netherlands	17,187	0.02	0.21	0.01	1	0.84	25	1.14	0.90	6
New Zealand	12,008	0.08	0.28	0.03	0	0.52	8	0.87	0.53	6
Nicaragua	786	0.09	0.22	0.27	—	0.63	13	0.32	—	—
Nigeria	339	0.65	0.06	0.14	1	0.87	9	0.13	0.03	3
Norway	23,083	0.03	0.16	0.01	1	0.52	27	0.68	0.31	6
Oman	5,696	0.04	0.12	0.05	0	0.69	6	0.29	0.15	5
Pakistan	377	0.14	0.53	0.19	0	0.73	15	0.37	0.16	2
Panama	2,435	0.05	0.08	—	—	0.54	9	0.69	0.10	3
Papua New Guinea	1,104	0.03	0.23	0.03	—	0.78	5	0.31	—	3
Paraguay	1,049	0.13	0.17	0.33	0	0.35	23	0.20	—	4
Peru	1,046	0.13	0.43	0.32	1	0.65	22	0.13	0.20	3
Philippines	615	0.07	0.12	0.12	1	0.44	21	0.43	0.79	4
Poland	1,903	4.65	0.40	0.10	1	0.45	32	0.29	0.04	—
Portugal	5,199	0.11	0.14	0.03	1	0.32	38	0.88	0.19	6
Qatar	12,820	—	0.01	0.04	—	1.00	2	0.71	—	—
Romania	1,341	2.16	0.30	0.33	0	0.70	7	0.15	—	—
Russia	1,989	1.90	0.46	0.19	0	0.44	18	0.12	0.05	—

(Table continued on following page.)

Table 3. *(continued)*

Economy	*GDP per capita (U.S. dollars)*	*Inflation*[a] *(percent)*	*Tax rate*[b] *(percent)*	*Reserves/ deposits*	*Deposit insurance*[c]	*Market concentration*[d]	*Number of banks*[e]	*Bank assets/ GDP*[f]	*Stock market capital/GDP*	*Law and order*[g]
South Africa	2,176	0.09	0.23	0.04	0	0.71	15	0.67	2.09	4
Saudi Arabia	5,316	0.04	0.00	—	0	0.96	4	0.41	0.33	5
Singapore	13,436	0.04	0.29	0.08	0	0.48	19	0.96	1.74	6
Spain	9,137	0.05	0.26	0.08	1	0.50	49	1.01	0.35	4
Sri Lanka	640	0.11	0.23	—	0	0.63	7	—	—	—
Swaziland	787	0.47	0.16	0.21	—	—	4	0.23	0.30	—
Sweden	19,387	0.04	–4.91	—	0	0.41	18	0.62	0.78	6
Taiwan (China)	7,268	0.02	0.18	—	1	0.40	25	—	—	5
Thailand	1,807	0.04	0.31	0.04	0	0.49	14	0.98	0.86	5
Tunisia	1,464	0.05	0.20	0.04	—	0.55	8	0.55	0.22	—
Turkey	1,848	1.06	0.14	0.27	1	0.43	29	0.17	0.17	5
United Kingdom	13,478	0.02	0.36	—	1	0.39	71	1.12	1.27	6
United States	20,931	0.01	0.27	0.03	1	0.16	372	0.48	0.95	6
Venezuela	2,651	0.51	0.06	0.25	1	0.46	17	0.13	0.05	4
Yemen	280	—	0.51	—	—	1.00	3	—	—	—
Zambia	247	0.55	0.34	0.08	0	1.00	3	0.10	—	4

— Not available.

Note: For all variables 1995 figures were reported, if available. Otherwise figures are for the most recent year available.

a. Inflation is the annual inflation of the GDP deflator.

b. The tax rate is defined as total taxes paid by banks divided by before-tax profits.

c. Deposit insurance is a dummy variable that takes the value 1 if there is an explicit deposit insurance scheme and 0 otherwise.

d. Market concentration is defined as the ratio of the assets of the largest three banks to total banking assets.

e. This value is the number of banks in the data set with complete information.

f. Bank assets include the total assets of the deposit money banks.

g. The law and order indicator reflects the degree to which the citizens of a country are willing to accept the established institutions to make and implement laws and adjudicate disputes. It is scored 0–6 with higher scores indicating sound political institutions and a strong court system. Lower scores indicate a tradition of depending on physical force or illegal means to settle claims.

Source: GDP per capita and inflation are from World Bank National Accounts. The tax rate, market concentration, and number of banks are from IBCA's BankScope database. Reserves/deposits and bank assets/GDP are from the International Monetary Fund, *International Financial Statistics*. Deposit insurance is compiled from Kyei (1995) and Talley and Mas (1990). Stock market data are from International Finance Corporation's emerging market database. The law and order indicator is produced by the International County Risk Rating Agency.

tries tend to have lower bank-to-GDP and capitalization-to-GDP ratios, with some notable exceptions. Malaysia, South Africa, and Thailand, for instance, have relatively high ratios for both variables.

The final column in the table provides an index of law and order, which is one of the institutional variables used in the regression analysis. This variable is scaled from 0 to 6, with higher scores indicating sound political institutions and a strong court system. Lower scores reflect a tradition in which physical force or illegal means are used to settle claims. The table shows considerable variation among countries in the sample.

III. Empirical Results

Tables 4 and 5 report the results of regressions of the net interest margin and before-tax profits as a share of total assets, respectively. Measuring profitability using the return on equity (as opposed to using the return on assets and controlling for equity ratios as we do here) does not lead to significantly different results and thus is not reported. All regressions include country and year fixed effects. The tables report several specifications, the basic one including a set of bank and macroeconomic indicators as regressors. These are important control variables to which we subsequently add the taxation variables, the deposit insurance index, the financial structure variables, and the legal and institutional indicators. We drop some variables from these two regressions because we want to ensure that banks from a reasonable number of countries are included. The estimation technique is weighted least squares, with the weight being the inverse of the number of banks for the country in a given year. This weighting corrects for the fact that the number of banks varies considerably across countries.

Bank Characteristics and Macroeconomic Indicators

The first bank characteristic is the book value of equity divided by total assets lagged one period (E/TA_{t-1}). We lag total assets by one period to correct for the fact that profits, if not paid out in dividends, have a contemporaneous impact on bank equity. Buser, Chen, and Kane (1981) examine the theoretical relationship between bank profitability and bank capitalization. They find that banks generally have an interior optimal capitalization ratio in the presence of deposit insurance. Banks with a high franchise value, reflecting costly bank entry, have incentives to remain well-capitalized and to engage in prudent lending behavior (see Caprio and Summers 1993 and Stiglitz and Uy 1996). Berger (1995b) provides empirical evidence that U.S. banks show a positive relationship between bank profitability and capitalization, although the evidence is not conclusive. The author notes that well-capitalized firms face lower expected bankruptcy costs for themselves and their customers, thereby reducing their cost of funding.

The basic specification (column 1 in tables 4 and 5) confirms that there is a positive relationship between E/TA_{t-1} and net interest income and bank profitability. In the regressions this variable is also interacted with GDP per capita (measured

Table 4. *Determinants of Net Interest Margins*

	Regression results				
Independent variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Bank characteristics					
Equity/lagged total assets (E/TA_{t-1})	0.046***	0.047***	0.044***	0.064***	0.063***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Equity/lagged total assets interacted with GDP per capita	–0.001	0.000	–0.001	–0.002**	–0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Loans/total assets	0.017***	0.008**	0.012***	0.022***	0.019***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Loans/total assets interacted with GDP per capita	–0.000	0.001***	0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-interest earning assets/total assets	–0.016**	–0.020***	–0.021***	–0.011	–0.020***
	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
Non-interest earning assets/ total assets interacted with GDP per capita	–0.001*	–0.001	0.000	–0.002**	–0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Customer and short-term funding/total assets	–0.007	0.003	0.004	–0.000	–0.004
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Customer and short-term funding/total assets interacted with GDP per capita	0.000	–0.000	0.000	–0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Overhead/total assets	0.173***	0.025***	0.213***	0.141***	0.310***
	(0.022)	(0.019)	(0.019)	(0.018)	(0.019)
Overhead/total assets interacted with GDP per capita	0.002***	0.004*	0.004*	0.009***	0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Foreign ownership dummy	0.004***	0.003**	0.004***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Foreign ownership dummy interacted with GDP per capita	–0.001***	–0.001***	–0.001***	–0.000***	–0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Macroeconomic indicators					
GDP per capita	0.000	0.000	0.000	0.000	–0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Growth rate	0.004	0.005	0.006	–0.011	–0.020)**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007
Inflation rate	0.021***	0.026***	0.025***	0.020***	0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Real interest rate	0.044***	0.060***	0.058***	0.051***	0.025***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Real interest rate interacted with GDP per capita	0.001	–0.004	–0.003*	–0.005***	–0.000
	(0.002)	(0.002)**	(0.002)	(0.002)	(0.002)

Table 4. *(continued)*

	Regression results				
Independent variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Taxation					
Reserves		−0.076***	−0.076***	−0.024*	−0.104***
		(0.015)	(0.015)	(0.016)	(0.016)
Reserves interacted with GDP per capita		0.011***	0.011***	0.009***	0.004
		(0.003)	(0.003)	(0.003)	(0.004)
Tax rate		0.016***	0.015***	0.017***	0.017***
		(0.002)	(0.002)	(0.002)	(0.002)
Tax rate interacted with GDP per capita		−0.001***	−0.001***	−0.001***	−0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Deposit insurance					
Deposit insurance dummy			−0.009***		
			(0.003)		
Financial structure					
Bank assets/GDP				−0.024**	
				(0.010)	
Bank assets/GDP interacted with GDP per capita				0.001*	
				(0.001)	
Stock market capitalization/GDP				0.016***	
				(0.005)	
Stock market capitalization/GDP interacted with GDP per capita				−0.002***	
				(0.001)	
Stock market capitalization/bank assets				−0.013***	
				(0.003)	
Stock market capitalization/bank assets interacted with GDP per capita				0.001**	
				(0.001)	
Number of banks				−0.001	
				(0.015)	
Market concentration				0.004	
				(0.005)	
Total assets (U.S. dollars)				0.003***	
				(0.000)	
Legal and institutional indicators					
Contract enforcement dummy					−0.042***
					(0.007)
Contract enforcement dummy interacted with GDP per capita					0.003***
					(0.001)
Law and order index					−0.003***
					(0.001)
Law and order index interacted with GDP per capita					−0.000***
					(0.000)
Corruption					−0.009***
					(0.001)

(Table continued on following page.)

Table 4. *(continued)*

	Regression results				
Independent variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Corruption interacted with GDP per capita					0.001***
					(0.000)
Adjusted R^2	0.50	0.51	0.50	0.58	0.63
Number of observations	5,841	5,276	5,212	5,054	4,497

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

Note: The regressions are estimated using weighted least squares pooling bank-level data across 80 countries for the 1988–95 time period. The number of banks in each period is used to weight the observations. The regressions also include country and time dummy variables that are not reported. The dependent variable is the net interest margin defined as net interest income divided by total assets. Standard errors are given in parentheses.

Source: Authors' calculations.

Table 5. *Determinants of Bank Profitability*

	Regression results				
Independent variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Bank characteristics					
Equity/lagged total assets (E/TA_{t-1})	0.047***	0.051***	0.055***	0.058**	0.015***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.006)
Equity/lagged total assets interacted with GDP per capita	0.002	0.002***	0.003***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Loans/total assets	–0.013***	–0.024***	–0.023**	–0.015***	–0.018***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Loans/total assets interacted with GDP per capita	0.001***	0.003***	0.003***	0.003***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-interest earning assets/ total assets	–0.005	–0.010	–0.011	–0.014	–0.033***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.007)
Non-interest earning assets/ total assets interacted with GDP per capita	–0.007***	–0.007***	–0.007***	–0.008***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Customer and short-term funding/total assets	–0.029***	–0.017**	–0.014***	–0.031***	–0.051***
	(0.006)	(0.007)	(0.008)	(0.001)	(0.005)
Customer and short-term funding/total assets interacted with GDP per capita	0.002***	–0.000	–0.000	0.001	0.002***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Overhead/total assets	–0.023	–0.006	–0.004	–0.024	–0.114***
	(0.025)	(0.026)	(0.026)	(0.026)	(0.019)
Overhead/total assets interacted with GDP per capita	–0.030***	–0.049***	–0.049***	–0.048***	0.007***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)

Table 5. *(continued)*

	Regression results				
Independent variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Foreign ownership dummy	0.005***	0.006***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Foreign ownership dummy interacted with GDP per capita	−0.001***	−0.001***	−0.001***	−0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Macroeconomic indicators					
GDP per capita	0.008***	0.008***	0.008***	0.007***	0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Growth rate	0.002	−0.006	−0.007	−0.019	0.004
	(0.010)	(0.011)	(0.011)	(0.011)	(0.007)
Inflation rate	0.011	0.015*	0.014*	0.009	0.011*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.005)
Real interest rate	0.023***	0.029***	0.029***	0.023***	0.026***
	(0.009)	(0.010)	(0.010)	(0.009)	(0.006)
Real interest rate interacted with GDP per capita	−0.000	−0.001	−0.001	−0.000	−0.003**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Taxation					
Reserves		−0.126***	−0.129***	−0.106***	−0.091***
		(0.021)	(0.021)	(0.023)	(0.016)
Reserves interacted with GDP per capita		0.029***	0.031***	0.032***	0.005***
		(0.004)	(0.004)	(0.004)	(0.004)
Tax rate		0.022***	0.022***	0.021***	0.017***
		(0.003)	(0.003)	(0.003)	(0.002)
Tax rate interacted with GDP per capita		−0.000	−0.000**	−0.003**	0.000***
		(0.000)	(0.000)	(0.000)	(0.000)
Deposit insurance					
Deposit insurance dummy			−0.005		
			(0.004)		
Financial structure					
Bank assets/GDP				−0.028*	
				(0.014)	
Bank assets/GDP interacted with GDP per capita				0.002*	
				(0.001)	
Stock market capitalization/GDP				0.010	
				(0.007)	
Stock market capitalization/GDP interacted with GDP per capita				0.000	
				(0.001)	
Stock market capitalization/bank assets				−0.001	
				(0.001)	

(Table continued on following page.)

Table 5. *(continued)*

	Regression results				
Indicator	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Stock market capitalization/bank assets interacted with GDP per capita				–0.001	
				(0.001)	
Number of banks				0.000	
				(0.000)	
Market concentration				0.010*	
				(0.007)	
Total assets (U.S. dollars)				0.000	
				(0.000)	
Legal and institutional indicators					
Contract enforcement dummy					–0.022***
					(0.007)
Contract enforcement dummy interacted with GDP per capita					0.001***
					(0.001)
Law and order index					–0.000
					(0.001)
Law and order index interacted with GDP per capita					–0.000*
					(0.000)
Corruption					–0.002*
					(0.001)
Corruption interacted with GDP per capita					–0.000
					(0.000)
Adjusted R^2	0.21	0.27	0.27	0.31	0.35
Number of observations	5,841	5,276	5,212	5,054	4,497

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.
Note: The regression is estimated using weighted least squares pooling bank-level data across 80 countries for the 1988–95 time period. The number of banks in each period is used to weight the observations. The regressions also include country and time dummy variables that are not reported. The dependent variable is before-tax profits divided by total assets. Standard errors are given in parentheses.
Source: Authors' calculations.

in constant thousands dollars for the year 1987). The positive coefficient on the interaction variables in the before-tax profits regression may reflect a higher bank franchise value in wealthier countries. The coefficients on E/TA_{t-1} and the interaction variable together indicate how the ratio of equity to assets affects bank variables in countries with different income levels. For a country with a per capita GDP of $10,000, for instance, the point estimate of the effect of E/TA_{t-1} on before-tax profits divided by total assets is 0.067 (or $0.047 + 10 \times 0.002$).

There is a negative and significant coefficient on non-interest-earning assets as a share of total assets in the net interest margin equation, but there is no significant relationship for the before-tax profits equation. Note that the sign on this variable interacted with per capita GDP is negative in both the net interest margin

and the before-tax profits specifications. Apparently, the presence of non-interest earning assets depresses net interest income and profitability more in wealthier countries than in poorer countries. By contrast, the sign on loans divided by total assets is positive in the net interest margin equation and negative in the before-tax profits equation. However, the coefficient of this variable interacted with GDP in the profits equation is positive, indicating that at higher income levels bank lending activities tend to be more profitable.

On the liability side, customer and short-term funding consists of demand deposits, savings deposits, and time deposits. On average, this type of customer funding may carry a low interest cost, but it is costly in terms of the required branching network. This liability category does not significantly affect the net interest variable, although there is evidence that it lowers bank profitability.

Differences in overhead may also capture differences in bank business and product mix, as well as variation in the range and quality of services. The ratio of overhead to total assets has an estimated coefficient of 0.173 in the net interest margin regression, which suggests that about a sixth of a bank's overhead cost is passed on to its depositors and lenders. The interaction of overhead with per capita GDP also enters with a positive coefficient, indicating that a larger share of overhead is passed onto financial customers in wealthier countries. This may reflect more competitive conditions in banking markets in industrial than in developing countries. In the before-tax profits regression the interaction of overhead with per capita GDP enters negatively, indicating that higher overheads eat into bank profits.

The foreign ownership variable equals 1 if at least 50 percent of the bank's stock is in foreign hands and equals 0 otherwise. In both tables 4 and 5 this variable has a positive coefficient, while its interaction with per capita GDP has a negative coefficient. These results suggest that foreign banks realize relatively high net interest margins and profitability in relatively poor countries. It may be that foreign banks are frequently exempt from unfavorable domestic banking regulations and apply superior banking techniques. Note, however, that the point estimate of the effect of foreign ownership for a wealthy country with a per capita GDP of $20,000 is negative in the net interest margin equation at –0.016 (that is, $0.004 - 20 \times 0.001$) and in the profitability equation at –0.015 (that is, $0.005 - 20 \times 0.001$). Foreign banks' technological and efficiency advantages in countries may be insignificant because, while there, they face informational disadvantages. This could explain why foreign banks in industrial countries are relatively unprofitable on average.

Turning to the macroeconomic indicators, we see, first, that per capita GDP has no significant impact on the realized net interest margin, although it enters the profitability equation with a positive coefficient. Per capita GDP is a general index of economic development, and it thus reflects differences in banking technology, the mix of banking opportunities, and any aspects of banking regulations omitted from the regression. Growth, defined as the growth rate of real per capita GDP, is insignificant in both regressions. The percentage change in the GDP defla-

tor, or inflation, is estimated to increase the net interest margin and bank profitability. However, the significance of the coefficients in the profitability regressions is low, possibly because banks obtain higher earnings from float or because there are delays in crediting customer accounts in an inflationary environment. With inflation, bank costs also tend to rise. A larger number of transactions may lead to higher labor costs and, as shown by Hanson and Rocha (1986: 40), result in a higher ratio of bank branches per capita. On net, however, the regression results suggest that the impact of inflation on profitability, although not very significant, is positive throughout.

We constructed the real interest rate using the short-term government debt yield and, if that measure was not available, other short-term market rates. The real interest rate enters the net interest margin and before-tax profits regressions positively, although this variable interacted with per capita GDP has a significantly negative coefficient in the ne t interest margin equation. Thus there is some evidence that increases in the real interest rate do not raise spreads as much in industrial countries, perhaps because their deposit rates are not tied down by deposit rate ceilings.

Taxation Variables

Banks are subject to direct taxation through the corporate income tax and other taxes, and they are subject to indirect taxation through reserve requirements. Reserve requirements are an implicit tax on banks if, as is usual, official reserves are remunerated at less-than-market rates. The corporate income tax and the reserve tax differ in important respects. First, the corporate income tax, in principle at least, can be targeted at pure profit. To the extent that it is a profit tax, the corporate income tax is relatively nondistorting. In practice, however, it may not be a pure profit tax if complete expensing of costs is not allowed.

The reserve tax, by its very nature, is proportional to the volume of deposit taking and is therefore a distorting tax. From a welfare perspective the corporate income tax thus appears to be superior to the reserve tax. A second important difference is that the severity of the reserve tax depends on the opportunity cost of holding reserves. This may depend on financial market conditions as much as on any tax code. Related to this second condition, reserve requirements are also an instrument of monetary policy.

As far as we know, there has been no other empirical research on the effect of the corporate income tax on the banking sector. Several studies have considered the impact of reserve requirements on bank profitability. Some, in particular, have examined how Federal Reserve membership affected the profitability of U.S. commercial banks in the 1970s (see Rose and Rose 1979 and Gilbert and Rasche 1980). Federal Reserve membership subjected banks to generally higher reserve requirements. Most of the studies in this area support the notion that nonmember banks were more profitable than member banks (with similar characteristics) because nonmember banks held relatively little cash. Competition among member and nonmember banks in the same market appears to have pre-

vented member banks from passing their higher reserve costs onto their customers. In related work Kolari, Mahajan, and Saunders (1988) study the impact of announced changes in reserve requirements on bank stock prices using an event study methodology. Huizinga (1996) and Eijffinger, Huizinga, and Lemmen (1998) examine how nonresident withholding taxes affect interest rates, while Fabozzi and Thurston (1986) examine how differences in reserve requirements are priced into money-market instruments.

Because detailed information on the reserve regulation of all countries in our sample is not available, we use a proxy to capture bank reserves. We construct this variable in the regressions as the product of the banking system's ratio of aggregate reserves to deposits (as in table 3) and the individual bank's ratio of short-term funding to total assets. Customer and short-term funding, consisting of demand deposits, savings deposits, and time deposits, here proxy for reservable deposits. The reserves variable is thus an approximation of actual bank reserves that reflects differences in systemwide reserve requirement rules.

In tables 4 and 5 the reserves variable enters the regressions negatively. The coefficients in the net interest margin equations show two effects: less-than-market remuneration and the impact on banks' lending and deposit rates. The first effect is expected to be negative because underremunerated reserves lower a bank's net interest income and profitability. The second effect could be either zero, in which case the bank bears the full cost of higher reserves, or positive, in which case the cost of reserves is passed onto bank customers in the form of higher interest margins. In table 5 we see that the reserves variable negatively affects bank profitability. This suggests that the second, or pass-through, effect is either nonexistent or too small to offset the first, or direct, effect. Abstracting from any pass-through, the coefficient on the reserves variable in either regression can also be interpreted as a bank's opportunity cost of holding reserves. The reserves variable interacted with per capita GDP enters both regressions positively, which may reflect the fact that the opportunity cost of holding reserves is higher in wealthier countries.

We capture the explicit taxes that banks pay with the variable tax rate, which is measured by a bank's tax bill divided by its pretax profits. This variable has a significantly positive impact on interest margins and profitability. The tax rate interacted with per capita GDP is negative and significant in both regressions. These results suggest that both the net interest margin and profitability increase with tax rates, but less so in richer countries. Thus the corporate income tax is passed through to bank customers to some degree.

To calculate the extent of this pass-through, we use the estimated coefficients on the tax rate variable and its interaction with per capita GDP. Let the pass-through be defined as the increase in before-tax profits following a one-unit increase in the corporate tax bill, or $\partial BTP/\partial TX$. Next, note that $(\partial BTP/\partial\tau)/TA = \beta$, where τ is the tax rate, TA is total assets, and β is estimated at $0.022 - (0.0004) \times$ (per capita GDP). Further, $\partial TX/\partial\tau = (\partial BTP/\partial\tau) + BTP$, as $TX = \tau \times BTP$. It now follows that $\partial BTP/\partial TX = \beta/[\beta\tau + (BTP/TA)]$. We can evaluate this expression using mean val-

ues of τ, *BTP/TA*, and per capita GDP separately for countries in four different income groups (low, lower-middle, upper-middle, and high), where per capita GDP is the international average for 1995. The calculations suggest that the pass-through coefficient, $\partial BTP/\partial TX$, equals 1.01, 0.72, 1.00, and 1.21 for countries in the four income groups, respectively. For low-income countries in 1995 *BTP/TA* and τ have mean values of 0.016 and 0.225 for all banks, while the average GDP per capita is $426. The calculations reflect the fact that in high-income countries the mean value of *BTP/TA* is lower, while the value of τ changes little.

Essentially, these results suggest that the corporate income tax completely passes through to bank cutomers. Thus there is no support for the notion that the corporate income tax is a nondistorting tax on bank profits. Generally, it is a source-based tax on domestically employed capital resources. A complete pass-through of this tax is consistent with the assumption that international investors demand a net-of-tax return on capital invested in a particular country independent of the country's source-based taxes.

Deposit Insurance

Several studies have examined the impact of deposit insurance using international data. Demirgüç-Kunt and Detragiache (1997) find that the existence of explicit deposit insurance is positively associated with the probability of banking crises. Barth, Nolle, and Rice (1997), however, find that deposit insurance has no significant impact on banks' return on equity for a sample of 142 banks in 1993. Bartholdy, Boyle, and Stover (1997) estimate that deposit insurance lowers the deposit rate by 25 basis points, using aggregate deposit interest rate data for 13 OECD countries during 1985–90. These authors discuss why deposit insurance has a theoretically ambiguous effect on interest margins. On the one hand, the deposit rate for insured deposits should decrease given the insurance protection. On the other hand, mispriced deposit insurance provides banks with an incentive to engage in more risky lending strategies to increase the contingent pay-out from the deposit insurance agency.

Brewer and Mondschean (1994) offer empirical support for the notion that deposit insurance creates incentives for banks to acquire risky assets by examining the junk bond holdings of U.S. banks, while Demirgüç-Kunt and Huizinga (1993) argue that deposit insurance was an important determinant of bank stock prices during the international debt crisis of the 1980s. This moral hazard problem and the associated risks can lead bank creditors to demand a higher interest rate. Also, for a given level of risk, deposit insurance may lead banks to lend money more cheaply than they otherwise would, depressing net interest margins and profitability. Even banks that do not engage in risky lending strategies themselves may experience a downward effect on interest margins because of bank competition.

The deposit insurance variable equals 1 if the country has an explicit deposit insurance regime. For some countries this variable changes with time, reflecting changes in the regime during the sample period. The results suggest that deposit

insurance lowers net interest margins. Deposit insurance may also influence margins and profits through its effect on financial structure—it encourages new entry and enables small banks to operate. However, when we include financial structure variables in the regression, the results do not change. The impact on bank profits is negative, but it is not significant, possibly because of the offsetting impact of mispriced subsidies in actual deposit insurance schemes. These results suggest that explicit deposit insurance regimes do not produce higher bank profitability and margins, perhaps because of design and implementation problems.

Financial Structure Variables

In the regressions reported in column 4 of tables 4 and 5 we include two sets of financial structure variables. The first set comprises the market concentration ratio, the number of banks, and total bank assets as indicators of market structure and scale effects. Various authors, such as Gilbert (1984), Berger (1995a), and Goldberg and Rai (1996), have pointed out that such variables may proxy for market power as well as for differences in bank efficiency. We do not attempt here to distinguish between the corresponding market power and efficient structure hypotheses.

The second set consists of financial structure variables that measure the importance of bank and stock market finance relative to GDP and to each other. Reasons why these variables matter may also hinge on market power arguments. A high ratio of bank credit to GDP, for instance, may reflect a high demand for banking services fueling competition among banks. Or these variables may reflect the complementarity of or substitutability between bank and stock market finance. The Miller-Modigliani theorem states that debt and equity finance are purely substitutes in the absence of taxes and bankruptcy costs. In practice, however, debt and equity finance may also be complementary, as modeled in Boyd and Smith (1996). Demirgüç-Kunt and Maksimovic (1996) provide empirical evidence that an ability to attract equity capital may also enhance firms' borrowing capacity, especially in developing countries' financial markets. In this setting easier equity finance may increase rather than decrease the demand for debt finance, reflecting that these sources of finance are complements.

Turning to the first set of market concentration and scale variables, we see that the bank concentration ratio has a significant and positive impact on bank profitability, while bank size, as proxied by total assets, has a significant and positive impact on interest margins. The number of banks has no significant impact on either interest margins or profits.

The second set of financial structure variables affects bank margins more significantly than bank profits. This may indicate that these variables have a greater impact on banks' loan and deposit customers than on other clients. The ratio of bank assets to GDP has a significantly negative impact on margins and profits, perhaps reflecting more intense interbank competition in well-developed financial systems. This effect is smaller in richer countries that already have relatively developed banking sectors. The ratio of stock market capitalization to GDP enters

the net interest margin equation positively, which suggests that a larger stock market per se enables banks to obtain higher interest margins, supporting the complementarity hypothesis between debt and equity financing discussed above. As stock markets develop, better availability of information increases the potential pool of borrowers, making it easier for banks to identify and monitor them. This raises the volume of business for banks, making higher margins possible. In the regression the ratio of stock market capitalization to banking assets enters the interest margin equation negatively. Thus it may be that a larger stock market relative to the banking sector lowers bank margins, reflecting substitution possibilities between debt and equity. For both stock market development indicators the interaction with per capita GDP enters the interest margin equation with the opposite sign, suggesting that the impact of any stock market development on interest margins is muted in wealthier countries.

Legal and Institutional Indicators

The final regressions reported in tables 4 and 5 include a variety of legal and institutional variables (column 5). The contract enforcement dummy, ranging from 1 to 4, measures the degree to which contractual agreements are honored and not subject to language and mentality differences. A higher value means greater contract enforcement. In both the net interest margin and before-tax profits regressions, the contract enforcement variable has a negative and significant sign. Poor contract enforcement may prompt banks to require higher interest margins and investors to require higher profitability to compensate for the additional risk. In both regressions the contract enforcement variable interacted with per capita GDP enters positively, suggesting a muted effect of this variable in wealthier countries.

The law and order index, ranging from 0 to 6, captures how well the legal system works in adjudicating disputes. From table 4 we see that a higher value of this index is significantly associated with lower interest margins. The reason may again be that an effective legal system reduces the required risk premia on bank lending. The interaction between the law and order index and per capita GDP enters the equation negatively, however.

Finally, the corruption index, ranging from 0 to 6, measures the degree of government corruption. A higher score indicates that government officials are less likely to take bribes. Table 4 indicates that a cleaner government is associated with lower realized interest spreads, and this relationship is weaker in wealthier countries. Again, banks may require a lower risk premium on their investments in countries that are relatively free of corruption. Overall, the regressions indicate that the underlying legal and institutional variables are important in explaining cross-country variation in interest spreads and bank profitability. For two of the three variables the interaction with per capita GDP has a coefficient with the opposite sign, suggesting that the effects of institutional differences are muted in wealthier countries.

IV. Conclusions

Banking systems around the world differ widely in their size and operation. Across countries commercial banks have to deal with different macroeconomic environments, explicit and implicit tax policies, deposit insurance regimes, financial market conditions, and legal and institutional realities. Using a comprehensive cross-country data set with bank-level data, this article analyzed how bank characteristics and the overall banking environment affect how banks function as reflected in interest margins and bank profitability.

We can confirm some findings of earlier research: for instance, a positive relationship between capitalization and profitability and a negative relationship between reserves and profitability. But other important determinants of bank margins and profitability, such as ownership, corporate taxation, financial structure, and the legal and institutional setting, have not been treated extensively in the literature.

Differences in the mix of bank activity also have an impact on spreads and profitability. Our results show that banks with relatively high non-interest earning assets are less profitable. Banks that rely largely on deposits for their funding are also less profitable, because deposits apparently entail high branching and other expenses. Similarly, variation in overhead and other operating costs is reflected in variation in bank interest margins, because banks pass on their operating costs to their depositors and lenders.

The international ownership of banks also has a significant impact on bank spreads and profitability. Foreign banks, specifically, realize higher interest margins and higher profitability than domestic banks in developing countries. This finding may reflect the fact that in developing countries a foreign bank's technological edge is relatively strong, apparently strong enough to overcome any informational disadvantage in lending or raising funds locally. Foreign banks, however, are shown to be less profitable in industrial countries, where they may not have a technological edge.

Macroeconomic factors also explain variation in interest margins. We found that inflation is associated with higher realized interest margins and higher profitability. Inflation entails higher costs—more transactions and generally more extensive branch networks—and also higher income from bank float. The positive relationship between inflation and bank profitability implies that bank income increases more with inflation than do bank costs. Further, high real interest rates are associated with higher interest margins and profitability, especially in developing countries. This may reflect the fact that in developing countries demand deposits frequently pay zero or below-market interest rates.

Regarding financial structure, banks in countries with more competitive banking sectors—where banking assets constitute a larger portion of GDP—have smaller margins and are less profitable. The bank concentration ratio positively affects bank profitability, and larger banks tend to have higher margins. A larger ratio of stock market capitalization to GDP increases bank margins, suggesting possible

complementarity between debt and equity financing. A larger ratio, however, is negatively related to margins, suggesting that relatively well-developed stock markets can substitute for bank finance.

Similarly, several institutional factors, such as indexes of credit rights, law and order, and corruption, and differences in financial structure, have more pronounced effects on interest margins and bank profitability in developing than in industrial countries. These results may reflect the relatively closed nature of banking markets in developing countries. Coupled with earlier empirical evidence that a weak institutional environment makes banking crises more likely (Demirgüç-Kunt and Detragiache 1997), these results suggest that returns to improving underlying institutions are indeed high. Reserves also have a more pronounced impact on margins and profitability in developing than in industrial countries. This result may simply reflect the relatively high opportunity cost of holding reserves in poorer and more inflationary countries.

The corporate income tax appears to be passed on fully to bank customers in both developing and industrial countries. This finding is consistent with the notion that bank stock investors require net-of-company-tax returns independent of the level of company taxation. It also implies that the corporate income tax on banks is likely to distort the underlying saving and investment decisions, with possibly negative implications for economic growth. These considerations must weigh heavily in considering the merits of the corporate income tax on banks as part of the overall tax system.

However, we also found that official reserves depress bank profits. Prima facie, this suggests that reserve requirements are a better instrument with which to tax bank profits than the corporate income tax. Note, however, that the implicit reserve tax in many countries is much more variable than the corporate income tax. The level of banking investment and activity is therefore unlikely to be adjusted to each change in the implicit reserve tax. Variability in the reserve tax can thus go a long way toward explaining the responsiveness of bank profits to this tax. These issues are pursued further in Demirgüç-Kunt and Huizinga (1997).

Policymakers have an interest in promoting banking sectors that are both stable and efficient. Stability clearly requires sufficient banking profitability, while economic efficiency requires banking spreads that are not too large. A prerequisite to formulating effective banking policies is thus to understand the determinants of bank profitability and interest margins.

Several other topics remain for further study. Countries worldwide differ considerably in the extent of foreign ownership of their banking systems. An interesting issue is how entry by foreign banks affects the operation of domestic banking firms. In principle, foreign entry can affect pricing by domestic firms and force them to reduce their operating costs and to remain competitive. Both of these effects determine whether the entry of foreign firms is welfare-improving overall. We address the impact of foreign entry in Claessens, Demirgüç-Kunt, and Huizinga (1997).

As a related issue, it would be interesting to consider what determines foreign bank entry. Foreign bank entry, and foreign direct investment generally, may be driven by the different (worldwide) taxation of domestic and foreign firms rather than simply by countries' comparative advantage in providing financial services.

We have found some evidence that government regulations, such as the design of deposit insurance schemes, have an impact on bank margins. It would be interesting to analyze this issue further by taking into account differences in design features. We intend to return to these issues in future work.

REFERENCES

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

Barth, James R., Daniel E. Nolle, and Tara N. Rice. 1997. "Commercial Banking Structure, Regulation, and Performance: An International Comparison." Comptroller of the Currency Economics Working Paper 97-6. Comptroller of the Currency, Washington, D.C. Processed.

Bartholdy, Jan, Glenn W. Boyle, and Roger D. Stover. 1997. "Deposit Insurance, Bank Regulation, and Interest Rates: Some International Evidence." University of Otago, New Zealand. Processed.

Berger, Allen N. 1995a. "The Profit-Structure Relationship in Banking: Tests of Market-Power and Efficient Structure Hypotheses." *Journal of Money, Credit, and Banking* 27(2):404–31.

———. 1995b. "The Relationship between Capital and Earnings in Banking." *Journal of Money, Credit, and Banking* 27(2):432–56.

Boyd, John, and Bruce Smith. 1996. "The Co-Evolution of the Real and Financial Sectors in the Growth Process." *The World Bank Economic Review* 10(2):371–96.

Brewer, Elijah III, and Thomas H. Mondschean. 1994. "An Empirical Test of the Incentive Effects of Deposit Insurance." *Journal of Money, Credit, and Banking* 26(1): 146–64.

Buser, Stephen A., Andrew H. Chen, and Edward J. Kane. 1981. "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital." *Journal of Finance* 35(1):51–60.

Caprio, Gerard, Jr., and Lawrence H. Summers. 1993. "Finance and Its Reform: Beyond Laissez-Faire." Policy Research Working Paper 1171. World Bank, Policy Research Department, Washington, D.C.

Claessens, Stijn, Aslı Demirgüç-Kunt, and Harry Huizinga. 1997. "How Does Foreign Entry Affect the Domestic Banking Market?" World Bank, Policy Research Department, Washington, D.C. Processed.

Demirgüç-Kunt, Aslı, and Enrica Detragiache. 1997. "The Determinants of Banking Crises: Evidence from Developed and Developing Countries." World Bank, Policy Research Department, Washington, D.C. Processed.

Demirgüç-Kunt, Aslı, and Harry Huizinga. 1993. "Official Credits to Developing Countries: Implicit Transfers to the Banks." *Journal of Money, Credit, and Banking* 25(3):430–44.

———. 1997. "The Taxation of Domestic and Foreign Banking." World Bank, Policy Research Department, Washington, D.C. Processed.

Demirgüç-Kunt, Aslı, and Vojislav Maksimovic. 1996. "Stock Market Development and Financing Choices of Firms." *The World Bank Economic Review* 10(2): 341–69.

Eijffinger, Sylvester, Harry Huizinga, and Jan Lemmen. 1998. "Short-Term and Long-Term Government Debt and Nonresident Interest Withholding Taxes." *Journal of Public Economics* 68(3):309–34.

Fabozzi, Frank J., and Thom B. Thurston. 1986. "State Taxes and Reserve Requirements as Major Determinants of Yield Spreads among Money Market Instruments." *Journal of Financial and Quantitative Analysis* 21(4):427–36.

Gilbert, R. Alton. 1984. "Bank Market Structure and Competition." *Journal of Money, Credit, and Banking* 16(4):617–60.

Gilbert, R. Alton, and Robert H. Rasche. 1980. "Federal Reserve Bank Membership: Effects on Bank Profits." *Journal of Money, Credit, and Banking* 12(3):448–61.

Goldberg, Lawrence G., and Anoop Rai. 1996. "The Structure-Performance Relationship for European Banking." *Journal of Banking and Finance* 20(4):745–71.

Gorton, Gary, and Richard Rosen. 1995. "Corporate Control, Portfolio Choice, and the Decline of Banking." *Journal of Banking* 50(5):1377–420.

Hanson, James A., and Roberto de Rezende Rocha. 1986. "High Interest Rates, Spreads, and the Cost of Intermediation: Two Studies." Industry and Finance Series 18. World Bank, Industry Department, Washington, D.C.

Huizinga, Harry. 1996. "The Incidence of Interest Withholding Taxes: Evidence from the LDC Loan Market." *Journal of Public Economics* 59(3):435–51.

International Monetary Fund. Various years. *International Financial Statistics.* Washington, D.C.

Kolari, James, Arvind Mahajan, and Edward M. Saunders. 1988. "The Effects of Changes in Reserve Requirements on Bank Stock Prices." *Journal of Banking and Finance* 12(2):183–98.

Kyei, Alexander. 1995. "Deposit Protection Arrangements: A Comparative Study." IMF Working Paper 95/134. International Monetary Fund, Washington, D.C.

Levine, Ross. 1997. "Financial Development and Economic Growth." *Journal of Economic Literature* 35(2):688–726.

Rose, John T., and Peter S. Rose. 1979. "The Burden of Federal Reserve System Membership: A Review of the Evidence." *Journal of Banking and Finance* 3(2):331–45.

Schranz, Mary S. 1993. "Takeovers Improve Firm Performance: Evidence from the Banking Industry." *Journal of Political Economy* 101(2):299–326.

Stiglitz, Joseph E., and Marilou Uy. 1996. "Financial Markets, Public Policy, and the East Asian Miracle." *The World Bank Research Observer* 11(2):249–76.

Talley, Sam, and Ignacio Mas. 1990. "Deposit Insurance in Developing Countries." World Bank, Latin America and the Caribbean Regional Office, Washington, D.C. Processed.

Vittas, Dimitri. 1991. "Measuring Commercial Bank Efficiency, Use, and Misuse of Bank Operating Ratios." Policy Research Working Paper 806. World Bank, Policy Research Department, Washington, D.C.

World Bank. 1996. *World Development Report 1996: From Plan to Market.* New York: Oxford University Press.

The World Bank

1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.
Telephone: (202) 477-1234
Facsimile: (202) 477-6391
Telex: MCI 64145 WORLDBANK or MCI 248423 WORLDBANK
World Wide Web: http://www.worldbank.org/
E-mail: books@worldbank.org

14362
9 780821 343623

ISBN 0-8213-4362-9