PRO-POOR ECONOMIC GROWTH RESEARCH STUDIES
CONTRACT NO. PCE-I-02-00-00015-00

Educational Sector Study: Pro-Poor Economic Growth Effects of Policies and Activities

Deliverable 1: Research and Results Dissemination Plan and Time Line
Deliverable 2: Website
Deliverable 3: Pro-Poor Economic Growth: A Review of Recent Literature
Deliverable 4: Poverty-Problem Countries Typologies
Deliverable 5: Selection Criteria for Pro-Poor Economic Growth Policies
Deliverable 6: Preliminary Policy Recommendations
Deliverable 7: Poverty Reduction Strategy Papers: A Preliminary Analysis of the Process and Outputs
Deliverable 8: Meeting: Presentation of Phase One Results
Deliverable 9: Workshop: Phase One Results, Preliminary Findings, and Planned Research
Deliverable 10: Poverty, Economic Growth, and Development Policies and Activities: A Case Study of:

- *Brazil*
- *Egypt*
- *Indonesia*
- *Peru*
- *Uganda*
- *Ukraine*
- *Zambia*
- *India or Sri Lanka*

Deliverable 11: Pro-Poor (Sector) Policies, Reforms and Activities

- *Agriculture*
- *Education*
- *Finance*

Deliverable 12: Pro-Poor Economic Growth Policies, Reforms, and Activities
Deliverable 13: Workshop: Research Findings and Implications for USAID Programming
Deliverable 14: Pro-Poor Economic Growth and Poverty Reducing Policies, Reforms, and Activities (Guidance Manual)
Deliverable 15: Workshop: Final Findings and Presentation of Guidance Manual
Deliverable 16: Dissemination of Findings of Guidance Manual
Deliverable 17: Final Project Report
Deliverable 18: Issues Papers on:

- *Conflict and Post-Conflict Recovery*
- *The Environment*
- *Gender*
- *Health*
- *HIV/AIDS*
- *Privatization*

Educational Sector Study: Pro-Poor Economic Growth Effects of Policies and Activities

By

Jere R. Behrman*

April 2003

*Jere R. Behrman is the William R. Kenan, Jr. Professor of Economics and Director of the Population Studies Center at the University of Pennsylvania, McNeil 160, 3718 Locust Walk, Philadelphia, PA 19104-6297, USA, jbehrman@econ.sas.upenn.edu. He thanks Albert Berry for useful comments on a draft of this paper. This document represents part of Deliverable #11 under this task order. Behrman alone—and not DAI, BIDE or USAID—is responsible for all the interpretations given in this paper.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Education is widely thought to be central for promoting efficient growth and for increasing the assets of the poor to bring them out of poverty over time, if not immediately. Education is most usefully defined to include a broad range of learning activities. Formal schooling is the most-emphasized and the most-studied of those activities. But it is important to recognize that education encompasses activities other than just formal schooling—for example, pre-school learning at home and in formal programs, informal and formal general and specialized training programs, and learning from experience all may be important parts of education more broadly defined.

This paper reviews what we know and what we do not know about the pro-poor economic growth effects of policies and activities in the educational sector, drawing on the substantial literature in this area. For this purpose, education is broadly defined as noted, but given the concentration in the literature on formal schooling, this review also concentrates on formal schooling. The paper provides, based on the imperfect available knowledge, recommendations about what to do and not to do in this sector, the prioritization of strategies, and how such recommendations depend on different country conditions and thus the topology of countries as outlined in the project paper by Berry (2002).

Although the focus of the paper is on recommendations regarding what to do and what not to do in the educational sector, such recommendations cannot be made in a vacuum. To understand what we know and what we do not know, some framework for analysis is necessary. Such a framework should address what determines education and the effects of education, what are the problems in making inferences about the causal determinants of education and the effects of education, and what are the key motives for policy. Chapter One lays the foundation for the recommendation in this paper by summarizing these issues (with elaboration on estimation issues in Appendix A). Chapter Two describes basic aspects of patterns in the most-emphasized form of human capital—schooling—with regard to associations with regions, time, per capita income, and adult women=s schooling to provide some perspective (Appendix B provides elaboration). Chapter Three then provides recommendations for the educational sector (with details on one important program that is a useful model for some policies and, more important, for evaluation in Appendix C). The key policy recommendations for educational policy and pro-poor economic growth presented in Chapter Three are:

1. **Recognize that education is more than formal schooling.** If “education” means learning that may increase productivity in activities and improve welfare, the term includes not only schooling but also learning at home, in the community, and on the job prior to, during, and subsequent to schooling and learning through formal training programs. More attention needs to be paid to evaluating education outside of formal schooling as well as to improving the evaluation of formal schooling.

2. **Distinguish between private and social rates of return to different types of education and be conscious of policy hierarchies.** The basic policy motives are to increase efficiency (in the sense of more welfare for the same resources and technology) and to attain better distribution (which often, and in this project in particular, focuses on reducing poverty). It is important to consider policy alternatives with regard to both of these basic objectives. There may be some policies that are Awin-win@ by improving efficiency and benefiting the poor. The reduction of the impact of capital market restrictions on human capital investments is a likely example. But other policies involve tradeoffs between efficiency and distribution that should be considered in assessing policy choices. And just because some policies, such as those that increase schooling, have productivity gains does not mean that such policies are desirable in terms of efficiency—what is central for that assessment is the difference between the social and the private rates of return, which may be small or negative even if the private rates of return are high (as, for example, for most forms of tertiary schooling).

3. **Recognize that the rates of return to investing in the education of the poor appear substantial in many cases—comparable or exceeding those to many other investments, with those for investing in females often at least as high as those for investing in males.** The relatively small number of studies that evaluate the benefit-to-cost ratios for education in developing countries suggest that these ratios are greater than one, thus indicating that the returns to increasing education are relatively high in comparison with many other investments. Some careful studies also suggest that such returns are likely to be at least as high for investments in the education of females as of males because of gender specialization in tasks such that female education has greater impact on important outcomes related to health and nutrition and perhaps the education of the next generation. These studies, however, are limited to a few forms of education (primarily formal schooling) in a relatively few contexts. Although their results suggest the possibility that there may be relatively high rates of returns to expanding education in other developing country contexts, it is important that careful evaluation of educational policies be undertaken in the particular contexts in which a policy is being considered and, if the policy is implemented, that the evaluation be updated with new information that becomes available from careful monitoring of the policy.

4. **Recognize that rates of return to education depend importantly on market, policy, and cultural contexts.** The rates of return to education are likely to be higher in contexts in which there is an ongoing stream of market and technological changes as a result of more extensive integration into world markets. The rates of return to female education, for example, are likely to be higher in societies in which females have more flexible choices regarding time use and occupations.

5. **Focus on the parts of the lifecycle in which the returns are likely to be highest and include health and nutrition support among the policy tools.** The rates of return appear to be relatively high for early education and for complementary human capital investments in health and nutrition from conception onward. In many developing countries, therefore, there are likely to be win-win pro-poor policy improvements possible through shifting resources away from high per-student subsidies for tertiary

schooling (where most of the students are from middle- and upper-income families) toward basic education (including pre-school programs) and infant and early childhood health and nutrition.

6. **Improve school access for basic education.** Currently, in most parts of most developing countries, primary schools are accessible for most children, although significant pockets remain where primary schools are not available, which is likely to be costly in terms of pro-poor growth policies. Nevertheless, the margin at which questions of school access are increasingly important is for secondary schooling. Also, schooling policies should consider options that are not limited to certain classes of ownership such as public schools. Policies are likely to be much more effective in terms of pro-poor growth objectives if they are neutral with regard to ownership of school suppliers, thus encouraging a variety of new providers to increase competition.

7. **Consideration should be given to improving educational quality, not just to increasing the quantity of education.** School quality includes multiple dimensions—class size, teacher-to-student ratios, teacher education, curriculum development, supplies of textbooks and other materials, and decentralization so the provision of education is sensitive to local conditions. In many developing country contexts, the returns to improving some dimensions of school quality appear to considerable. But there are not magic formulas that fit everywhere. Consideration of policies that might improve school quality need to be evaluated with careful attention to local conditions, with monitoring and updating of the evaluations of those policies that are implemented, and policies directed toward all potential providers of educational services (not just a subset such as public schools), if resources are to be used best for pro-poor growth.

8. **Before using mandates to attempt to increase education, obtain more information about their effects.** Policy makers, in education as in most other areas, often think that mandates, perhaps in some legal form. are relatively effective ways of inducing desirable changes. If more schooling is desired, for example, why not mandate a higher level of compulsory schooling? But the associations between such mandates and the level of schooling over time or across societies for the most part do not indicate that the mandates caused more schooling but that, apparently, the mandates reflected market and cultural pressures that increased schooling. Moreover, such mandates often impinge particularly on the welfare of poorer members of society. The message here is not that such mandates never should be used but that their effectiveness should be considered carefully in comparison with alternatives, such an increasing incentives to attain the same objectives.

9. **Consider using incentives in the form of scholarships for students from poor families to increase their education.** A particular form of incentives that seems to have been effective in increasing schooling in several cases recently is to provide scholarships for students from poor families. Such programs might be emulated more broadly. But, as with other policies, they need to be evaluated and monitored carefully over time and not just blindly transplanted. In some contexts, for example, they may have significant negative congestion effects through inducing increased enrollments in schools that are already overcrowded. And that poor households would increase schooling of their

children much less if the same transfers were to be provided without being conditional on school attendance raises the question of whether such resources are increasing the welfare of the poor as much as they might were they unconstrained. From the point of view of decisions makers in poor households, the answer seems clearly that schooling children more is not as high priority as are other uses of such resources. Possibly, such constraints are justified on distributional grounds because household decision makers do not weigh enough (from a social point of view) the interest of their children. But redistribution from poor parents to their children is not obviously a pro-poor policy (although it may be pro-growth) because most (not all) children in the developing world can expect to grow up in higher per capita income economies than did their parents.

10. **Improve information about what educational providers are providing.** Imperfect information about what educational providers are doing is the basis for some central tendencies in the educational sector, such as favoring public over other providers and mandating changes that are thought to be in the interests of the poor, although they may not understand this. However, in such situations there are likely to be alternatives that are higher in the efficiency and pro-poor policy hierarchies. Examples include the alternatives of improving the information on which parents and students make their educational decisions and developing mechanisms so they can use such information effectively (for example, by making choices among educational providers).

11. **Improve information and evaluation of the rates of return to alternative strategies for increasing education.** Good systematic evaluations of educational policies are rare and probably are biased toward more successful programs. These evaluations, moreover, do not necessarily provide sufficient information for simply transplanting apparently successful programs in one context to other contexts. The resource costs of good evaluations, moreover, are small compared with the resource costs of the policies being evaluated. Therefore, there are likely to be high returns to adopting the policy that the systematic evaluation of educational policies is the norm, rather than the exception.

CHAPTER ONE
FRAMEWORKS FOR ANALYSIS

This chapter begins with a discussion of why frameworks for analysis are necessary, then summarizes a standard analytical framework for the determinants of investments in education and the impact of such investments, next notes the problems in estimating the impact of investments in education, then considers the basic motives for policy interventions, and finally discusses alternative methodologies for economic evaluation of investments in education.[1] The general framework that is summarized can be used for all investments in education—investments that range from improving human capital, for example, by improving pre-school home environments, pre-school programs, schooling, home environment while in school, training, and learning on the job and from other experiences. The wide range of possible investments in education indicates that there are numerous possibilities in which the use of current resources to increase learning affect future productivities in a wide range of market and nonmarket activities. Policy changes may improve such investments in education in terms of the basic policy motives of efficiency and distribution (with redistribution towards the poor a leading special case of the latter) that are discussed below.

WHY FRAMEWORKS FOR ANALYSIS ARE NECESSARY

Good analysis of the impact of investments in education has tripartite foundations: data, modeling and estimation. These three dimensions are critically interrelated. Data are essential for empirical analysis, limit the extent to which analyses can be undertaken, and shape most of the estimation problems. If there were available data from well-designed and well-implemented experiments, associations between observed investments in education and observed outcomes would reveal the underlying causality directly. But for numerous reasons, including costs and ethical concerns, such experimental data are rarely available. Therefore, while there may be high returns for some aspects of policy analysis to increase experimental data, most analysis has been and will continue to be based on behavioral data. Such behavioral data can "speak for themselves" regarding associations between investments in education and various outcomes. But they generally cannot "speak for themselves" with regard to what observed determinants—policies or otherwise—cause differences in investments in education or to what extent observed investments in education cause different outcomes. The problem is that most data are the result of a number of behavioral decisions taken by households, individuals, bureaucrats, policy-makers and others in light of a number of factors unobserved by analysts. Good analysis of what causes investments in education or of what effects such investments have is difficult, and requires a much more systematic approach than simply looking at associations among observed variables.

[1] The material in this section is presented in a very summary form; for a more extensive discussion within a more general context, see Knowles and Behrman (2003a,b).

Analytical frameworks permit exploring systematically investments in education, point to what data are needed for such explorations, facilitate the interpretation of empirical findings, and help to identify some of the probable estimation issues that should be addressed given the data used. The analytical frameworks provided by models are essential if the empirical estimates are based on behavioral data generated in the presence of unobservables such as innate ability and family connections. The problem, for example, is that individuals with greater ability and motivation and better innate health may be more productive directly and may also benefit from higher levels of investments in their education. Therefore it may be difficult to sort out the effect of investments in education per se as opposed to the fact that such investments are correlated with unobserved abilities, motivation and innate health.

For such reasons, the empirical effects of investments in education can be analyzed satisfactorily with nonexperimental data only within frameworks that incorporate well the essence of relevant behaviors. To be interpretable, estimates based on behavioral data require some model of the underlying behaviors, though far too often in the literature the models used are not explicit. Those who are not clear about their framework of analysis may think they are revealing underlying truths unconstrained by such frameworks, but they are instead usually making implicit assumptions that may upon examination not be plausible.

ANALYTICAL FRAMEWORKS FOR THE DETERMINANTS OF INVESTMENTS IN EDUCATION

Households and the individuals in them are the proximate sources of demands for many investments in education, given their predetermined assets (i.e., physical, financial, and human, including endowments[2]), production functions related to human resources, public and private services related to investments in education (i.e., schools, training programs), and current and expected prices for inputs used in investments in education and for outcomes of the investments. Policies, of course, may enter directly or indirectly into this process through a number of channels ranging from the accessibility and quality of public and private educational services to the functioning of capital markets for financing investments in education to the functioning of markets in which these investments are expected to have returns. Becker's (1967) Woytinsky Lecture provides a simple framework for investments in human resources that captures many of the critical aspects of such investments and which has been widely appealed to in rationalizing empirical studies of the determinants of investments in education and their impacts.

In Becker=s framework, human resource investment demands reflect the equating of expected marginal private benefits and expected marginal private costs (both in present discounted

[2] "Endowments" means characteristics that are given independent of behavioral decisions. From the point of view of investments in the education of a particular individual, genetically determined innate ability and innate health robustness are examples—but so are other aspects of family and community background at birth. Investments in that individual—including those in education but also, for example, in health—are likely to respond to those endowments. Investments in education, for example, often are presumed to be complementary with innate ability and family connections and thus responsive to these aspects of endowments.

terms) for investments in a given individual. The marginal private benefit curve depends importantly, *inter alia*, on expected private gains in productivity in all of the ways in which the investment may have impacts. The marginal private benefit curve is downward-sloping because of diminishing returns to investments in education (given genetic and other endowments) and because, to the extent that investments in education take time (such as schooling, training, as well as search in labor and other markets), greater investments imply greater lags in obtaining the returns and a shorter post-investment period in which to reap those returns. The marginal private cost may increase with investments in education because of higher opportunity costs of more time devoted to such investments (especially for schooling and training) and because of increasing marginal private costs of borrowing on financial markets. The equilibrium human resource investment for an individual is where the marginal private benefits and the marginal private costs are equalized. This equilibrium human resource investment is associated with an equilibrium rate of return that equates the present discounted value of expected marginal private benefits with the present discounted value of expected marginal private costs. This simple stylized representation of human resource determinants is based on a dynamic perspective, with both benefits and costs not only in the present but also those that are expected in the future and with current period options conditional on past decisions. Thus it is consistent with placing investments in education in a lifespan perspective, as has been emphasized from a number of different perspectives.

If the marginal private benefit curve is higher, all else equal, the equilibrium human resource investment and the equilibrium marginal private benefit both are greater. The marginal private benefit curve may be higher for one of two otherwise identical individuals except for the difference noted below that in many cases may be due directly or indirectly to policies because one individual (or whomever is investing in that individual, such as children=s parents): (1) has greater endowments that are effective in increasing schooling (or other forms of education) and in post-schooling labor markets and household production; (2) has lower discount rates so that the future benefits of investments in education have greater value at the time of the investment decision; (3) has investments in education options of higher quality (e.g., access to higher quality public schools or public health services) so that the marginal private benefits for a given level of private investments are higher, and the equilibrium investments greater;[3] (4) has better health and a longer expected life due to complementary investments, so that the post-investment period in which that individual reaps the returns to the investment is greater and therefore the expected returns greater; (5) has greater marginal private benefits to a given level of such investments because of labor market discrimination that favors that individual due to gender, race, language, family, village, or ethnic group; (6) has returns to human resources investments that are obtained more by the investor or the relevant decision maker (e.g., if traditional gender roles dictate that children of one sex, but not the other, provide old-age support for their parents, parental incentives

[3] If the investor must pay for greater human resource service related quality, investment does not necessarily increase with a higher quality option. What happens to the equilibrium investment tends to depend upon where the marginal private cost curve for the higher quality option is in relation to the location of the marginal private benefit curve.

may be greater to invest in youth who are likely to provide such support);[4] (7) has greater marginal private benefits to a given level of investment because of being in a more dynamic economy in which the returns to such investments are greater; (8) has greater marginal private benefits to a given level of such investments because of greater externalities from the human resource investments of others in the same society; or (9) lives in a more stable economy so that the discount rate for future returns is lower and thus the marginal private benefit of future returns greater.

If the marginal private cost is lower, all else equal, the equilibrium human resource investment is greater, with the marginal private benefit lower at the higher investment level. The marginal private cost might be lower for numerous reasons, again, most or all of which may be affected directly or indirectly by policies. Compare two otherwise identical individuals except that one individual: (1) has lower private cost access to educational services related to such investments because of closer proximity to such services or lesser user charges; (2) has less opportunity costs for time used for such investments (e.g., due to gender specialization in household and farm tasks performed by youth, tasks that appear to be more concentrated among poor than among better-off youth); (3) faces lower utility costs of such investments because of cultural norms that favor some activities associated with such investments more for some individuals than for others (e.g., in some societies, it is not thought desirable that girls past puberty mingle with males outside of the family in transit to school or in school so that the preference costs of schooling are lower for boys than for girls);[5] or (4) is from a household with greater access to credit because of greater wealth or status or better connections.

The maximization through equating expected private marginal benefits and costs leads to dynamic decision rules or demand relations for educational investments in individual i that depend on all relevant prices P, on all relevant resources R and on all the parameters of the relevant production functions, on preferences and on stochastic factors that, say, may enter into production processes (e.g., weather fluctuations) or reflect preference differences:

(1) $H_i = H(P, R \mid \text{production parameters, preference parameters, stochastic factors})$.

The prices include all prices that enter into the investing household=s decision-making process, including the prices paid by the household for goods and services related to investments in education, other uses of household resources and for transferring resources over time (i.e., the interest rate) and for insuring formally or informally against uncertainty. At the time that any human resource investment decision is made, these prices include all past and current prices (perhaps embodied in current stocks of human capital), as well as expected future prices (including expected future returns to investments). The resources include all resources of the individual, household (identified by ownership if there is intrahousehold bargaining), educational and health/nutrition institutions, and community that

[4] Though this tendency may be offset if, for example, human capital substitutes sufficiently for financial transfers in marriage markets (e.g., Rao 1993 and Behrman, et al. 1999 present some evidence regarding this possibility for rural India).

[5] For this case the marginal utilities of marginal private benefits and costs are equated.

affect any household decisions. These resources include human resources that reflect past investments, financial resources, physical resources, genetic endowments, school characteristics, and general learning environments. Policies may have impacts directly and indirectly through a number of the prices and resources.

This simple framework systematizes six critical, common sense, points for investigating dimensions of the determinants and the effects of investments in education—and how these relate to policy choices.

First, the impacts of changes in policies may be hard to predict by policymakers and analysts. If households or other entities face a policy or a market change, they can adjust all of their behaviors in response, with cross-effects on other outcomes, not only on the outcome(s) to which the policy is directed.

Second, aggregation to obtain macro outcomes will average out random stochastic terms across individuals or households. But such aggregation does not average out systematic behavioral responses at the micro level. Therefore associations among macro variables can reveal, conditional on the overall context, what are those associations—but not causal effects of processes occurring at the micro level. For most phenomena of relevance regarding the causal effects of what determines education for what impact does education have on the poor and on efficient growth for this paper, only careful systematic micro evidence is likely to be informative—though, of course, aggregate data may be suggestive of associations and patterns across time and space and other characteristics (see Chapter Two).

Third, the marginal benefits and marginal costs of investments in a particular individual differ depending upon the point of view from which they are evaluated: (i) There may be externalities or capital/insurance market imperfections so that the social returns differ from the private returns and (ii) there may be a difference between who makes the investment decision (e.g. parents) and in whom the investment is made (e.g. youth). The effectiveness of policies are likely to depend crucially on what are perceived to be the private effects from the point of view of private decision-makers, and these may differ from the social effects of interest to policymakers.

Fourth, investments in education are *determined* by a number of individual, family, community, (actual or potential) employer, market and policy characteristics, only a subset of which are observed in available data sets. To identify the impact of the observed characteristics on investments in education, it is important to control for the correlated unobserved characteristics.[6]

[6] For example, if schools with higher quality tend to be in areas in which expected rates of returns from investments in education tend to be greater but only indicators of school quality and not expected rates of return are observed in the data and if there is not control for the unobserved expected rates of return in the analysis, the impact of school quality on such investments is likely to be overestimated because in the estimates school quality proxies in part for unobserved expected rates of return to these investments. If expected rates of return differ across communities or clusters in the sample, they could be controlled in the estimates with community dummy variables (or fixed effects) (e.g., Alderman *et al.* 1996a for an example for Pakistan).

Fifth, to identify the *impact* of investments in education, it also is important to control for individual, family, community, market and policy characteristics that determine the investments in education and also have direct effects on outcomes of interest.

Sixth, empirically estimated determinants of, and effects of, investments in education are for a given macro economic, market, policy, schooling and regulatory environment in which there may be feedback both at the local and at a broader level. For this reason attention to the characteristics that determine how a particular country fits into the typologies discussed in Berry (2002) is likely to be important for understanding what are likely to be the more promising changes in educational policies related to pro-poor efficient growth.

EMPIRICAL ISSUES—MEASUREMENT

To assess the rates of return to investments in education, it is necessary to (a) measure what we mean by investments in education and by various outcomes that might be affected by investments, (b) estimate the impact of investments in education on the latter measures, and (c) measure the costs of these investments. These are not trivial tasks. This section considers some of the measurement difficulties. The next section turns to estimation problems.

Investments in education: Key variables for this paper are indicators of investments in education. In the case of schooling, for an important example, most empirical studies represent human resource investments empirically by years of schooling or highest grade of schooling (or level of schooling) completed. Though "years" and "grades" of schooling are often used as synonyms, they need not be the same if they are used literally and if there is grade repetition, as is widespread in many parts of the developing countries (e.g., in much of Latin America). And since repetition rates often differ by gender, not accounting appropriately for grade repetition may lead to misleading inferences about gender differences in the rate of return to schooling as well as the total rate of return to schooling. The point simply is that one of the major costs of schooling is the opportunity cost of time in school, which is greater if there is more grade repetition for a given schooling grade attainment. Putting aside the question of the time spent in school, there are other limitations of grades (years) of schooling as a measure of human resource investments. Probably most important is the implicit assumption that school quality is constant. But empirical measures indicate that school quality varies substantially, so it would be desirable in assessment of the impact of investments in education to represent not only the time (grades, years) that they spend in school, but also the quality of that schooling. If both the quantity and the quality of schooling should be included, but only the quantity is included as is usually the case, the likely result is to overstate the impact of time in school and to miss that there is likely to be an important quality-quantity tradeoff.

Besides schooling, as noted, there are many other investments in education—use of parental time and other resources to invest in education at home, pre-school programs, training programs and learning through experiences on jobs and in other activities. Information on these other forms of educational investments and on their effects is much more scarce and limited than information on schooling—which has led, as noted, to over focus on the role of

schooling (particularly time in schooling) in discussions of changing educational investments.

Outcome Variables: Unfortunately there are many problems in measuring these outcomes.[7] For some outcomes that may be affected importantly by investments in education, data usually used in the social science literature do not include direct measures—alleged external or spillover effects on others are important examples. This may mean that important outcomes are missed when assessing the impact of investments in education.

For some other outcomes there may be at best imperfect indicators—for instance, representing health by health-related inputs (e.g., nutrients), reported disease conditions, curative health care, and preventative health measures (e.g., vaccinations). Some of these measurement problems may be systematic, moreover, resulting in biases in the estimated impact of investments in education. If, for example, those who have less schooling report less sickness for the same health conditions as those who have more schooling (perhaps because the degree of sickness viewed as Anormal@ varies inversely with schooling), the impact of schooling on actual sickness will be underestimated.

For some other measures of outcomes conventions are used in the literature that may not be sensible. For example, economic gains from reducing mortality—which may be accomplished through educational investments— are represented at times by the present discounted value of foregone earnings of an individual (e.g., CGCED 2002). For young individuals, such gains can be considerable. For measuring the purely economic benefits of survivors, however, this seems an overstatement of the economic costs of mortality since such individuals also would have consumed—perhaps most or all of their earnings—over their lifetime.[8] For another example, the economic gains from improving the earnings capacities of individuals are at times measured by the reduction of their demands on governmental social welfare systems for transfers (e.g., CGCED 2002). But governmental transfers, though perhaps appropriately viewed as costs from a budgetary perspective, are not resource costs in the sense desirable to evaluate policies. There are likely to be some resource savings if social welfare programs are reduced because some resources are consumed in running and financing such programs. But such possible resource savings should not be confused with the amount of transfers involved (and the latter probably greatly exaggerate the true resource costs).

Even for the economic outcomes on which there long has been focus in evaluating the impact of the schooling component of education, there often are serious measurement problems. People with less schooling, for example, may remember income less well, may have more problems in assessing the value of their income because more is in kind or self-produced, and may be subject to greater variations in income because of sporadic employment with

[7] The discussion here is limited to micro data because the problems with the aggregate data are so severe (see Behrman and Rosenzweig 1994, de la Fuente and Doménech 2002, Knowles and Behrman 2003a,b, and Srinivasan 1994).

[8] This measure of the economic value of life also implies that there is no value to the life of an individual who is not productive because, for example, of age or disability.

seasonal fluctuations. If such factors lead to a tendency to underreport their incomes, then—all else equal—the returns to schooling will be overestimated due to these systematic errors.

Costs: There often are problems in measuring the costs of investments in education. The direct budgetary expenditures for particular investments, for example, may be intermingled with many other expenditures in budgets, and therefore be difficult to identify. They also may be spread among various budgets at various levels of aggregation—for example, if schooling staff salaries are paid by the Ministry of Education but other recurrent expenditures are paid by local governments as in the Philippines. There further frequently is a problem that the budgetary expenditures do not reflect the true costs because of distortions in market prices, perhaps created by policies. For example, governments may mandate wage rates for some workers such as teachers that may differ from the true scarcity value of such individuals, and then introduce other distortions such as in benefits or job security in order to attempt to attract enough qualified individuals to these positions. A major problem with evaluating costs, finally, pertains to the nongovernmental costs. Policy-makers often ignore these because they are focused on governmental budgets. But the costs of a program to the private sector may be considerable. For example, many programs require considerable amounts of time of private individuals—time that has opportunity costs in the form of other uses. For another example, as noted above, raising funds for governmental programs in itself may have large distortionary costs.

Measurement of Policies: It might seem strange to include policies among the key variables for which there are measurement problems. But there are serious measurement issues regarding what we know. For example, there is considerable emphasis on the potential for improved outcomes through policies that improve women=s education. But the empirical foundation for these claims in substantial part relates associations between years of schooling for females and outcomes that are viewed as desirable. Years of schooling is not a policy variable, but a behavioral outcome that reflects the schooling market interaction between household/individual demands for time in schooling and various aspects of schooling supply, both of which may be channels through which policies may work in complex ways and the workings of which may not be represented well by interpreting the world as if year of school is a policy variable.

EMPIRICAL ISSUES—ESTIMATION OF EFFECTS OF INVESTMENTS IN EDUCATION

This paper is concerned with assessing what we know about policies that effect education in developing countries, their rates of private and social returns and their impact on the poor. But obtaining good empirical estimates of the effects of investments in education is not easy, in part because of the measurement problems discussed above and in part for other reasons to which I now turn. I first consider the possibility of empirical estimation through experiments and then through econometric methods using nonexperimental (behavioral, observational) data.

Experimental Evaluation

To assess the impact of a particular change, such as increasing secondary school availability, the ideal would be a double-blind experiment with random assignment to treatment and control groups for the policies of interest. Experiments have been conducted to evaluate a relatively small number of policies related to investments in education in developing countries. One example is the Mexican Programa de Educación, Salud y Alimentación, PROGRESA (the Education, Health, and Nutrition Program), which was introduced as part of an effort to break the intergenerational transmission of poverty and which included an evaluation sample in which communities were randomly assigned to immediate versus delayed treatment. This design permitted evaluation of the impact of this program on a number of outcomes by comparing changes in treatment households with changes in control households. Another example is for urban Colombia, where students were selected randomly from poor applicants to receive scholarships to attend private schools. A third example is the random assignment of deworming programs, textbooks and flip charts among schools in Kenya. Studies based on these experimental data are summarized in Chapter Three.

But the possibilities for using such experiments for policy evaluation are limited. First, most such experiments cannot be double-blind, with neither the subjects nor the evaluator knowing who received treatment (though some medical experiments can be if, for example, the placebo appears to human senses to be identical with the treatment). That those who are treated know that they are treated may create better performance (i.e. the "Hawthorne effect"). That those who are not treated know that they are not treated may create incentives to obtain treatment through migration, political pressure, market purchases or other means. Second, the argument often is made that new programs can not easily be introduced at the same time throughout a country, so it may be effective in terms of evaluation, as was attempted with PROGRESA, to introduce them in a random set of treatment communities and only later in a random set of control communities. But if members of the control group perceive that they will eventually be affected by the program (perhaps with less than certainty) and if they can transfer resources over time, they should immediately adjust their behavior to reflect their changed command over resources due to the expected eventual future direct program impact. If so, the comparison between the program and the treatment groups probably underestimates the program impact.[9] Third, many experiments cannot be conducted because they are unethical or too costly. Imposing randomly some resources for education, for example, may be viewed as unethical. Even if some such possibilities are not viewed as unethical, they may be very costly in terms of resources or in terms of political costs. Consider the difficulties, for example, with the possibility of randomly assigning schooling among individuals in order to obtain better estimates of the effects of human

[9] Of course this is not a problem if members of the control group do not know that they eventually will be affected by the program, but such ignorance may be difficult to maintain because of interactions between members of the treatment and control groups and general information about the new program. In fact in some cases the administrators of the program may tell the control group directly that they will be included eventually in hopes of obtaining their agreement to serve in the control group and to enhance a sense of fair treatment. Atanasio, Meghir and Santiago (2002) analyze the extent of what they call "announcement effects" (but which might be more accurately labeled "anticipation effects" because there may be anticipations regarding eventual program participation among members of the control group even if there are no formal announcements) within a structural model and in some cases find the effects to be considerable.

resource investments in education as would be desirable for the present paper. Fourth, even for the policies for which good experiments can be conducted at a reasonable (resource, ethical and political) cost, they would reveal only the gross changes induced by the experimental treatment conditional on a particular situation, not what would happen in somewhat different circumstances. That is, experiments basically are Ablack boxes@ that reveal the total impact of some change, but do not reveal anything about the underlying structural relations that could be used to infer what would be the effects of other changes.[10] Fifth, there may be insufficient time to observe effects of interest to policy makers. This is a particular problem for many investments in education, since many of the desired outcomes are expected to occur much later in life or even in subsequent generations through children (e.g., see the evaluations in Knowles and Behrman 2003a).

For such reasons, though it would be desirable to increase experimental evaluation of policies and to assure that the experiments that are undertaken are of high quality (e.g., with good baseline data and random assignment of treatment versus control groups), there are severe limits on what policies can be evaluated by experimental means. Nevertheless, the experimental design is an important benchmark against which other means of evaluation should be compared and judged to aid in understanding what are the probable biases that may arise from non-experimental evaluation.

Econometric Estimates of Impacts of Investments in Education

Econometric or statistical methods are used to attempt to circumvent some of the limitations of the data, including that most data that are available for evaluation of the impact of investments in education are behavioral data and not experimental data. Appendix A provides some discussion of these issues. The basic point is that there are a number of reasons—including measurement errors in variables, right-side variables that reflect current or past choices, important variables that are not observed in data sets (e.g., innate ability, preferences), and selected samples (e.g., cognitive achievement test results often only are available from those who apply for or go to secondary or tertiary schools)—why the stochastic terms in relations being estimated often are not likely to be independent of the right-side variables. As a result, estimates are likely to be biased and therefore misleading unless special data and estimation procedures, grounded firmly in the analytical framework used for the analysis, are used to control for these problems. The systematic use of such methods is a central criterion for the selection of studies using behavioral (observational) data on which emphasis is placed in Chapter Three.

[10] For example, if scholarships of a particular magnitude are assigned randomly for schooling, the experimental evaluation only tells what is the impact of scholarships of that magnitude—not what would happen if smaller or larger scholarships were given or if in other ways the program were changed. Structural models, on the other hand, can be used to explore such questions—see, for example, Todd and Wolpin (2003), who develop, test and use such a model to assess the impact of changing the scholarship payments in the case of PROGRESA.

Implications for Analysis of Endogenous Policies

Policies of all sorts—including those related to investments in education—are not predetermined, but are made by individuals or groups of individuals with various objectives in mind, including accommodating to pressures from and needs of constituents. This means that it may not be possible to evaluate the impact of governmental policies on various outcomes without controlling for the fact that governmental policies themselves are determined, implemented and monitored as part of a larger set of behavioral decisions.[11] The failure to control for the determinants of governmental policies may cause substantial mis-estimates of their effectiveness. These mis-estimates, moreover, may either overestimate policy effectiveness or underestimate policy effectiveness, depending on what is the nature of the governmental decision. Therefore the usual estimates of policy effectiveness that do not control for the determinants of policies generally can not be assumed to give either lower or upper bounds on the true effects.

Rosenzweig and Wolpin (1986) formally develop these points. But the basic intuition is clear from considering the simple example of evaluating the impact on child education of special schooling-enrichment programs (e.g., related to the availability of textbooks, the training of teachers, the length of the school day) from cross-sectional data from a number of communities. If the resources devoted to such special programs tend to be concentrated in communities that have greater political power, wealth and better-educated individuals net of the effects of the special programs and of characteristics that are observed in the data, the association between child education and resources devoted to these special programs without control for resource allocation among these special programs in different communities overstates the effectiveness of the programs. Those communities that receive more special program resources would have had better child education for other reasons that are correlated positively with the allocation of special program resources (and *vice versa*). On the other hand, if the resources devoted to such special programs tend to be concentrated in communities that have poorer educational environments, greater poverty and less-healthy individuals net of the effects of the special programs and of characteristics that are observed in the data, the association between child education and resources devoted to these special programs without control for resource allocation among the special programs in different communities understates the effectiveness of the programs on child education. As a result, even if the special programs are effective in improving child education, the simple cross-sectional association between special program resources in various communities and child education even may be negative (and, if positive, is an underestimate of the effectiveness of special program resources).

Casual observations and a number of studies suggest that (a) governments do allocate resources, including those related to education, differentially among various types of communities and (b) the failure to control for such allocations of resources across communities may lead to substantial misunderstanding of program effects (e.g., Behrman and Birdsall 1988a, Gershberg and Schuermann 1999, Pitt, Rosenzweig and Gibbons 1993). Therefore it is essential to control for the determinants of program placement and intensity across recipients in order to evaluate with confidence the impact of programs.

[11] If programs are allocated randomly, as is discussed in Chapter One, this problem is avoided.

PRIVATE VERSUS SOCIAL RETURNS, EFFICIENCY, AND DISTRIBUTION

Often analyses of the impact of investments in education are undertaken without consideration of the general rationale for policies. It is just presumed that policies that, say, through increased investments in education increase some outcome such as subsequent productivity or health must be good. But such analyses are of little help in convincing skeptics that scarce resources should be allocated for these purposes, given many competing alternative uses. Moreover they may not provide much in the way of guidelines for choosing among policy alternatives. Therefore it is useful to ask why policy interventions with respect to investments in education might be desirable.

At a general enough level of abstraction, policy should be chosen in order to maximize social welfare. That begs, of course, the critical political economy question of how the social welfare function is determined. Even if that difficult question is put aside, the practical guidance offered by the injunction to maximize social welfare may seem quite limited. For that reason it often is useful to think separately of the two standard economic justifications for governmental policy interventions: 1) to increase efficiency/productivity and 2) to redistribute resources (which often is particularized to lessening poverty).[12]

Policy justifications based on efficiency and on distribution are both firmly rooted in micro dimensions of behaviors. *Both of these standard economic motivations for policy are concerned ultimately with the welfare of individuals as judged by those individuals*.[13]

Efficiency/Productivity

Resources are used efficiently in the economic sense of the term if they are used to obtain the maximum product possible given the quantities of the resources and the available production technologies at a point of time, and over time, and if the composition of that product increases the welfare of members of society as much as is possible given the resource and technological constraints, preferences and the distribution of resource ownership.[14] An investment (or expenditure) is efficient if the marginal social benefit of the last unit of that

[12] These two justifications include some other common concerns about policies, such as questions of access and quality of services and sustainability of overall economic development and of particular programs as is discussed in Behrman and Knowles (1998a).

[13] This last statement is emphasized by placing it in italics because economic efficiency is viewed by some as a concern about allocation of "things" and technical and financial concerns, but *not* as a concern about people. However this reflects a fundamental misunderstanding. Economic efficiency ultimately is a concern with the welfare of people as judged not by policymakers or international experts, but by the individual decision-makers involved. There, in addition, is a separate important concern about the distribution of decision-making powers, including the distribution between females and males. But it is important to recognize that the efficiency motive for policy, far from being purely a mechanical or technical concern of "dismal scientists" devoid of concerns about people, is based fundamentally in people's perceptions of their own welfare.

[14] It is important to note that efficiency is not just a concern about the static use of resources at a point in time, but also is a concern about the use of resources over time and thus productivity and productivity growth over time.

investment just equals its marginal social cost.[15] If the marginal social benefit of a particular investment is greater (less) than the marginal social cost, society is not investing enough (is investing too much) and would benefit from increasing (decreasing) the level of investment until the marginal social benefits and costs are equalized. Note that increasing efficiency is not the same as increasing economic growth. Economic growth may be increased, for example, by mandates that restrict current consumption below efficient levels in order to increase current investment and future growth. What is of policy concern therefore should not be to maximize economic growth, but to assure as efficient growth as possible.

Although applying the above rule maximizes social gains, private maximizing behavior leads to investments (including those related to education) at the level at which the marginal private benefit of the investment tends to equal its marginal private cost under the assumption that, given the information available to them and the constraints that they face, individuals act in what they perceive to be their best interests. Consider the possibility that private incentives for investments differ from social incentives for such investments, first with respect to the marginal benefits and then with respect to the marginal costs.

Why might marginal social benefits exceed marginal private benefits for investments in education? The most frequent answers to this question include: (1) There are "externalities" or "spillovers" in the form of effects on others that are transferred "external to markets." Investments in education are thought to have not only private benefits to the person being educated, but, by adding to society=s stock of knowledge, social benefits beyond the private benefits (positive "externalities" in the form of effects that are transferred external to markets). (2) Private information on which educational investments are made may misrepresent the private rates of return to these investments because it is incomplete or incorrect.[16] For example, individuals in general may have less basis for predicting future market developments that will affect future returns to education than "experts." The "public good" nature of information (i.e., that the marginal cost of providing information to another consumer is virtually zero) leads to under-production and dissemination of information from a social point of view by private markets because private providers cannot cover their costs if they price information at the social marginal cost as required for efficiency. (3) The combination of uncertainty, risk aversion and imperfect insurance markets may result in

[15] Four points should be noted. First, economic efficiency is not the same as engineering efficiency because of the incorporation of marginal benefits and marginal costs rather than focusing exclusively on technological efficiency. Second, economic efficiency is not the same as efficiency as sometimes used in the schooling literature (i.e., progression rates) because generally maximizing progression rates does not equate the marginal benefits and marginal costs. Third, these marginal conditions for efficiency may not hold if there are, for example, large discontinuities in production processes. In such cases choices may have to be made among a number of different alternatives, using an explicit welfare function to compare among the alternatives. Fourth, because of uncertainty in the real world this discussion could be recast in terms of expected values, with concern about possible risk aversion (or something other than risk neutrality). But, for simplicity, I do not do so.

[16] Realized marginal benefits (or costs), of course, may differ from expected marginal benefits (costs) because of unanticipated events. The point here, however, is not that realized marginal benefits may differ from expected marginal benefits, but the private expected marginal benefits may differ from social expected marginal benefits due to market failures regarding information. That is, private individuals are likely to be making their investment decisions based on expectations that are based on less information than is currently available.

private incentives to underinvest in educations (and other assets) because from a social point of view the risks are pooled. (4) The social discount rate may be lower than private discount rates because individuals value future outcomes more collectively than they do individually. (5) Prices for outcomes affected by such investments may understate social gains, in part because of distortions due to policies (e.g., policy restrictions on salaries in the health and education sectors). (6) Society may value an investment for individuals, or for certain types of individuals, more than do individual decision makers. For example, society may value heavily that every member of society has some basic level of education, but some parents (given the constraints under which they are operating, including current survival, which may require their children to be working) may not place sufficient priority on basic education for all of their children to assure that all of their children obtain basic education. Note that if society imposes its values/views in such a case (e.g., through an effective minimal schooling law), education of such children is increased to the basic level but the welfare of the (presumably relatively poor) family as judged by the parents is worsened. Therefore, though such a policy may reduce the welfare of the poor, or at least of the poor as evaluated by currently poor parents (though not necessarily by their children) unless it is completed by other policies such as income transfers. This is part of the reason that some programs (e.g., PROGRESA in Mexico, see below) have provided income transfers to poor households conditional on children attending school so that at least part of the income that the children would have earned were they working instead of in school is not lost to the household.

Why might marginal social costs be less than marginal private costs for investments in education? (1) There may be capital market imperfections for some types of investments because these forms of capital are not accepted as collateral such that the marginal private costs for such investments exceed their true marginal social costs, which probably is more relevant for individuals from poorer families who cannot relatively easily self-finance such investments. (2) The sectors that provide some types of services (e.g., schools, training programs) may produce inefficiently because institutional arrangements do not induce efficient production of an efficient basket of commodities. School teachers and staff, for example, might be oriented towards rewards established by the Ministry of Education or union negotiations, not towards satisfying the demands of clients. (3) The sectors that provide services related to investments in education may produce inefficiently because regulations preclude efficient production of an efficient basket of commodities. For example, regulations that limit hours during which schools are open, or limit textbook choices, or that impose quality standards based on different conditions in other economies or that limit provision of services to public providers, all may result in much greater costs of attaining specific investments than would be possible with less regulations.[17]

What are the implications of differences in the marginal private and social benefits or cost? Simply, if there are such differences, private incentives for investments in education differ from social incentives. In such circumstances, policies may increase efficiency by reducing

[17] This is not to say that all regulations are bad. In some contexts regulations may be the most efficient means of attaining a goal, particularly if there are certain types of information problems (e.g., those related to the quality of educational services that can not be easily discerned by consumers). But often regulations, no matter how good might be their intent, are not very effective policy tools (see Chapter One for further consideration of policy choices).

the differences between the private and the social incentives or through other means that cause outcomes to be closer to the socially-desirable outcomes.

Distribution

Distribution is a major policy motive distinct from efficiency. Distributional concerns, at least officially in pronouncements of governments and of international agencies, often focus on the command over resources of the poorer members of society—as in the concern of this project with pro-poor growth.[18] Society might well want to assure, for example, that everyone has basic schooling even at some efficiency cost currently or in terms of efficient growth. Though distributional concerns are often characterized by focus on the distribution of income or other resources among households, there may also be important distributional considerations within households.[19] Household decision-makers are not likely to consider equally the preferences of all household members in allocating household resources. For example, if women have preferences for using more resources to invest in children than do their husbands, these preferences may not be weighed equally as those of their husbands in decisions made by their husbands. Moreover, even if some households as aggregates have sufficient resources to cover what society considers basic needs, certain types of individuals in households may not be allocated what society considers to be sufficient resources for their individual satisfaction of basic needs. A particularly germane example may be child labor. Such labor may contribute importantly to the resources and the welfare of the household decision makers, but may detract from improving the human resources of the child by, for example, exposing the child to health and other risks and limiting the education of the child. Therefore there may be an important intergenerational distributional tradeoff.

Policy Choices to Increase Efficiency and to Improve Distribution

If all other markets in the economy are operating efficiently and there are differences between marginal private and social incentives in a given market related to investments in education, policies that induce investments at the socially efficient levels increase

[18] Many policies, whatever their official justification, however, distribute resources to middle and upper class households. For some examples for human resource-related policies in Viet Nam, see Behrman and Knowles (1998b, 1999) and World Bank (1995).

[19] A counter consideration about intrahousehold distribution is that one educated household member may provide benefits in the form of information or the interpretation of information for other household members. This point is developed formally by Basu and Foster (1998) and some empirical studies suggest that what accounts for major decisions (e.g., adoption of new technologies) is the level of schooling of the most highly-schooled household member (e.g., Foster and Rosenzweig 1995, 1996). On the other hand, households are not fixed units—individuals and subunits break off, and more schooling appears to increase the probability of breaking off (e.g., Foster and Rosenzweig 2002), perhaps more so for younger individuals (who tend to be more schooled given the secular increases in schooling that generally have predominated—see Chapter Two). Moreover, even in intact households, more education may shift bargaining power towards those with more education (generally males, though increasing less so; see Chapter Two). Therefore, even if there are what Basu and Foster call "intrahousehold externalities" to education, it is not clear that it would be desirable to focus on the highest educational level of current household members rather than on individuals in designing policies.

efficiency.[20] That still does not indicate what policies would be best to induce investments in education at the socially desirable levels. There is a large set of possibilities, including governmental fiats, governmental provision of services at subsidized prices, price incentives in markets related to investments in education, price incentives in other markets, and changing institutional arrangements in various markets. To choose among alternatives based on efficiency alone, there are two important considerations.

First, policies have costs. These costs include the direct costs of implementing and monitoring policies and the distortionary costs introduced by policies that may encourage socially inefficient behavior (including rent-seeking by both public and private entities). Often policymakers focus only on the direct costs and ignore the distortionary costs because only the direct costs have obvious and visible direct ramifications for governmental budgets. In fact the costs may be sufficiently high that it is not desirable to try to offset some market failures by policies.[21] But, if it is desirable to do so, there is a case generally for making policy changes that are directed as specifically as possible to the distortion of concern because that tends to lessen the distortionary costs. An *efficiency policy hierarchy* can be defined in which alternative policies to attain the same improvement in efficiency are ranked according to their social marginal costs, including direct and distortionary costs. This hierarchy indicates the preferential ordering of policies to deal with particular divergences between private and social incentives.[22]

Second, there are tremendous information problems regarding exactly what effects policies have, particularly in a rapidly changing world. This is an argument in favor of policies that are as transparent as possible, which generally means higher in the efficiency policy hierarchy with regard at least to distortionary costs because more direct policies are likely to be more transparent.[23] Information problems also provide an argument for price policies (taxes or subsidies) because if there are shifts in the underlying demand and supply relations they are likely to be more visible in a more timely fashion to policymakers if they have impact on the governmental budget than if they only change the distortions faced by private

---

[20] If all other markets in the economy are *not* operating efficiently, then policies that narrow the differences between private and social incentives in a particular market related to investments in education do *not* necessarily increase efficiency and productivity. But, in the absence of specific information to the contrary, such as the existence of two counterbalancing distortions, a reasonable operating presumption is that lessening any one distortion between social and private incentives is likely to increase efficiency.

[21] If the policies involve public expenditures as do most policies, it is important to consider the cost of raising the necessary tax revenue to finance the policy. For example, it has been estimated that the distortionary cost (often called the "deadweight loss") of raising a dollar of tax revenue in the United States ranges from $0.17 to $1.00, depending on the type of tax used (e.g., Feldstein 1995). Also see Devarajan, Squire and Suthiwart-Narueput (1997) and Harberger (1997).

[22] For example, it sometimes is argued that female schooling should be subsidized because more-schooled women have fewer children, which relieves budgetary pressures on subsidized schooling and health services. But in this case increasing female schooling through such subsidizes would not seem obviously to be high in the efficiency policy hierarchy. Higher in the efficiency policy hierarchy might be the elimination of any public subsidies for education and health that are not warranted by the marginal social benefits exceeding the marginal private benefits.

[23] This also is an argument for considering an experimental approach to evaluating policy alternatives when possible—e.g., rather than introducing a reform country-wide, introduce variants of reforms for schools (and other education-related services) in randomly selected sites with careful monitoring of the results for both the experimental groups and the control groups (Chapters One and Three).

entities as tends to happen with quantitative policies.[24] Finally, information problems in the presence of heterogeneities across communities point to the possible desirability of decentralization and empowerment of users of educational services in order to increase the efficacy of the provision of those services, though such considerations must be balanced off against possible economies of scale, higher quality of staff and possibly lower levels of corruption at more centralized levels, as well as intercommunity distributional concerns.

Thus, for efficiency/productivity reasons, particularly given that in the real world information is imperfect and changes are frequent, there is an argument generally for choosing policies as high as possible in the efficiency policy hierarchy defined by the extent of marginal direct and distortionary costs—and thereby using interventions that are focused as directly on the problem as possible. Note that this means that, for example, if there is a good efficiency reason for public support for investments in education, that does *not* mean that the best way to provide that support is through governmental provision of the relevant services. Higher in the efficiency policy hierarchy than direct governmental provision of such services, for example, may be subsidies or taxes that create incentives for the efficient provision of these services whether the actual providers are public, private or some mixture. On the other hand, policies that discriminate against one type of provider—for example, by making the availability of such subsidies dependent on whether the educational service provider is public —are generally likely to be lower in the efficiency policy hierarchy than policies that do not have such conditions.

Now consider distribution. Generally speaking the subsidization of specific goods and services (and even less, the direct provision by governments of goods and services at subsidized prices) is <u>not</u> a very efficient way of lessening distributional problems. Because subsidies are designed to lower prices to consumers, they induce inefficient consumption behavior. Instead, it generally is more efficient (and thus less costly in terms of alternative resource uses) to redistribute income to consumers, allowing them to allocate the income in ways that lead to efficient patterns of consumption.[25] Nevertheless, there are some cases in which subsidization of selected goods and services may be defensible to attain distributional objectives. For example, in cases where it is difficult (and therefore costly) to target the poor households or poor types of individuals within households, subsidizing certain goods and services that are mainly consumed by the poor may be the most efficient policy alternative.

Rather than being concerned with the general command over resources of its poorer members, as noted above, society may deem it desirable that everyone enjoy basic human resource related (and other) services, including basic education (as well as basic nutrition, housing and health care). Such an objective might be obtained through many means. But presumably it is desirable to assure that everyone have these basic options at as little cost in terms of productivity as possible so, rather than ignoring efficiency considerations, it is desirable to choose policies as high as possible in the efficiency policy hierarchy and still

[24] Nevertheless there are likely to be some cases, such as providing information regarding the quality of goods and services related to investments in education, for which quantitative regulations may be higher in the efficiency policy hierarchy than price policies because of the nature of the information requirements.

[25] However, even redistributing income may lead to inefficiency because it can affect the work effort of those on both the tax-paying and tax-receiving sides.

assure that the basic service objectives are met. Thus, to obtain a given distributional objective it is possible to define a *distributional policy hierarchy* in which policy alternatives that obtain that objective are ordered from lowest to highest marginal costs, including both direct and indirect cots. Efficiency goals thus should play an important role in interaction with the pursuit of distributional goals, not as independent considerations.

Methodology for Economic Evaluation of Investments in Education

If one has reliable estimates of the effectiveness of a set of alternative investments in education, how can one best evaluate them against the criteria of efficiency and distribution? There are several methodologies that are used in the literature.

Perhaps the simplest is cost effectiveness analysis (CEA). This consists of ranking a set of related investments according to their cost per unit of effectiveness, where the measure of "effectiveness" should be clearly defined and as narrow as practical, given the range of investments to be analyzed. CEA has been used widely to evaluate alternative investments *within* a given sector. However, CEA has several shortcomings in the context of evaluating alternative investments in education. First, it requires a single effectiveness measure. This is impractical in the case of many educational investments because they involve such a wide range of outcomes. Second, CEA does not provide any basis for comparing investments in education to alternative investments. Third, while CEA addresses aspects of the efficiency motive for policies, it does not do so comprehensively (i.e., it does not address whether the objective used to measure "effectiveness" represents itself an efficient use of resources).[26]

Cost-benefit analysis (CBA) is an alternative methodology for evaluating investments that is designed to handle alternative investments that may have several different outcomes. Because it values benefits in monetary terms, CBA obtains results (i.e., benefit-cost ratios or internal rates of return) that readily permit comparisons with alternative investments (e.g., investments in governance, investments in infrastructure). If one has reliable estimates of effects and of costs, the problem is valuing them in monetary terms. This is often technically challenging and can be politically sensitive as well (e.g., assigning a monetary value to a human life). Cost-benefit analysis should compare the present discounted value of the total extra (or marginal) benefits to the present discounted value of the total extra (or marginal) costs of a project or policy change. Several features of these benefit-cost ratios should be noted. First, the comparison of interest is at the margin—the additional benefits relative to the additional costs. But often empirical estimates may be of average, not marginal, benefits and costs. Second, the comparison should be in terms of present discounted values of additional benefits and additional costs because there is an advantage of obtaining the same benefit sooner rather than later (and of incurring the same cost later rather than sooner). If a benefit is received sooner rather than later, it can be reinvested to obtain further benefits (and similarly if costs can be postponed, the resources can be used for other purposes with positive returns in the mean time). With any reasonable discount rate, benefits received immediately are much more valuable than benefits received in the distant future. For example, with a 5

[26] On the other hand CEA can be, and often is, explicitly used to assess alternative means of attaining specific distributional goals, such as improving primary schooling enrollment rates for poor populations.

percent discount rate, $100 received in 14 years is worth only $50 received now. Sometimes in the literature, however, the importance of discounting seems to be ignored. That is, at times there is advocacy of more basic schooling because it will affect events that are decades in the future, such as reduced completed fertility or better health in old age, without adjusting these alleged benefits for the time lags necessary to obtain them.

Although there are several approaches that are used to do this in the literature, the most promising for use in the present context is to build on micro estimates of direct productivity effects that can be measured in monetary terms on the one hand and, for the effects that cannot be easily translated into monetary terms, use the resource cost of the most cost-effective alternative to achieve the same effects on the other hand.[27] This strategy is used by Summers (1992, 1994) to analyze the benefits and cost of investing in female education. Using Pakistani data, Summers began with estimates of the effect of an additional year of women=s schooling on her lifetime earnings. Next he developed estimates of the effect of an additional year of women=s schooling on child mortality, fertility and maternal mortality. He placed a monetary value on these latter effects by using estimates of the cost of producing similar effects using alternative cost-effective interventions (e.g., the cost per child life saved through measles immunization).[28] He compared these estimates (with some discounting to reflect the lagged nature of the effects) of Asocial benefits@ to the cost of providing an extra year of schooling to women and concluded that investing in girls= education yielded relatively high returns. The main problems with this approach are that (1) it is partial equilibrium and therefore market feedbacks, particularly within relatively closed economies, may be missed and (2) information with which to assess private versus social rates of return, and therefore, the efficiency motive, is relatively rare. Nevertheless, this seems the best available methodology, and therefore it is a useful benchmark for this paper. Some illustrations of this approach relevant to investments in education in developing countries (as well as to other investments in youth) are given in Knowles and Behrman (2003a).

Both the benefits and the costs are likely to vary from country to country, for a variety of reasons (see Berry 2002). One of the main determinants of both benefits and costs is the

---

[27] One prominent alternative is to define benefits in terms of an investment=s impact on economic growth, typically measured in terms of growth in GNP per capita. This approach was used in the pioneering study by Coale and Hoover (1958) of the economic benefits of fertility reduction. More recently, this methodology has been used to describe the association of investments in schooling with economic growth (Barro and Sala-I-Martin 1995). There are three main problems with this approach. First, the associations that are found in cross-national analysis may not represent unbiased estimates of the causal effects of investments (due to omitted variable and endogeneity bias). Second, the available cross-national data are subject to considerable errors and permit only fairly limited representation of the multiple facets of education (e.g., Behrman and Rosenzweig 1994, Srinivasan 1994, de la Fuente and Doménech 2002). Third, this approach does not deal well with either the distributional nor the efficiency motives for policies. Because it is on an aggregate level, it is not surprising that it does not deal well with distribution. But it also needs to be made clear that maximizing growth is not the same as the efficiency motive for policy. Too high growth may be inefficient just as may be too low growth.

[28] Summers' methodology involves using estimates of cost effectiveness ratios as a basis for valuing benefits. The implicit assumption is that discounted social benefits are at least equal to discounted social costs in the case of investments in which the cost-effectiveness ratio is at a minimum (and they would be exactly equal if the investment is at its optimal level). If instead, for instance, the social costs for alternative ways of attaining the same objective are higher than the social benefits, the cost-effectiveness ratio does not provide an accurate estimate of social benefits and would have to be adjusted downwards.

average level of income in a given country (as perhaps measured by GDP per capita in PPP). This affects benefits in terms of increased earnings, which generally are an important component of the benefits of investments in education. The average level of income in a given country also affects labor costs, which are often an important component of the cost of investments in education. Often (but not always) the two will virtually offset one another, so that the benefit-cost ratio or internal rate of return is unaffected. However, in the case of investments where this is not the case, it is desirable to adjust benefits and costs for differences in average levels of income.

CHAPTER TWO
PATTERNS IN ONE CENTRAL FORM OF EDUCATION—SCHOOLING—AS ASSOCIATED WITH REGIONS, TIME, INCOME PER CAPITA, AND ADULT WOMEN=S SCHOOLING

Schooling, as noted, is by far the most emphasized form of education in the literature. In fact many observers appear to equate "schooling" with "education," which I argue is too narrow a focus and may result in missing some important policy options outside of formal schooling. Perhaps because of the emphasis on schooling, however, only for schooling is there the possibility of making even rough comparisons among many countries over a number of years. Such comparisons, nevertheless, are useful for providing some perspective about the important schooling component of education. Because there has been considerable emphasis, both in the policy community and in the scholarly literature, on female schooling and possible gender gaps in schooling, this summary focuses on female schooling and male-female gaps in schooling. These descriptions are based on regressions with aggregate data for the 1960s through the 1990s, using World Bank categories to classify countries by income groups.[29] Though these data have limitations, they permit useful comparisons on four levels of disaggregation that are considered one at a time, and then the implications of considering them together in multivariate relations are noted. In each case in Appendix—there is consideration of female primary and secondary enrollment rates and the implied expected schooling for a synthetic cohort based on these enrollment rates,[30] plus the male-female gaps in the primary and secondary enrollment rates and in the expected schooling. This summary focuses on female expected schooling and the male-female gap in expected schooling because these are summary measures of enrollment rates at various schooling levels.

REGION

Expected years of female schooling on the average for most of the developing world has been much lower than in the developed countries: Sub-Saharan Africa and South Asia have averaged about nine years less, the Middle East and North Africa over six years less, East Asia and the Pacific over five years less and Latin America and the Caribbean over four years less (Europe and Central Asia is not significantly different). The gender gap in expected schooling favoring males has been largest in South Asia (2.3 years), the Middle

[29] The income groups include low-income, lower-middle-income, high-middle income and high income. Over this period of time the World Bank classification of some countries into income groups changed because, for example, some countries (e.g., in East and Southeast Asia, Botswana) had relatively high growth rates. In this analysis the unit of observation is the country in a particular year, so it is included with which ever country it was assigned by the World Bank for that year (which might be different from that assigned for a different year). In these estimates all country-year observations that are available are included in all the estimates and are weighted equally.

[30] Expected schooling for a synthetic cohort is the schooling that would be attained on average if a group of individuals on average experience the enrollment rates for different schooling levels observed at a point of time. See Appendix—for more details.

East and North Africa (2.2 years) and sub-Saharan Africa (1.4 years), much smaller but still statistically significant in East Asia and the Pacific (0.8 years) and Europe and Central Asia (0.3 years)—and insignificant in Latin America and the Caribbean. These patterns suggest that appropriate policies may differ significantly by regions and, to a lesser extent, among countries within regions. In parts of Sub-Saharan Africa and South Asia, for example, there may be significant shortfalls below universal enrollments at the primary level, with those left out of the system being primarily from poor families, girls more than boys. In such a context relatively high gains from pro-poor and growth/efficiency schooling policies may be focused on primary schools with particular attention to girls. In contrast, the most important margin for schooling policies in East Asia and the Pacific and Latin America and the Caribbean often is likely to be at the secondary school level because primary school enrollments already are close to universal—though, of course, there still may be important questions about improving school quality at the primary level. Likewise, concern about gender gaps favoring boys is an important issue in South Asia and the Middle East and North Africa, but for recent cohorts in Latin America and the Caribbean gender gaps favor girls (Behrman, Duryea and Székely 1999), though not yet at the high levels experienced in the United States and a number of developed economies.

TIME

Female expected schooling for a synthetic cohort increased significantly over time at close to a linear rate of about a sixth of a grade per year. Female expected schooling in the 1970s averaged 1.4 years above that in the 1960s, that in the 1980s averaged 3.8 years above that in the 1960s, and that in the 1990s averaged 5.2 years above that in the 1960s. This implied an increase from an average female expected schooling of 5.8 years in the 1960s to 11.0 years in the 1990s. These are considerable gains—a 90 percent increase. These changes over time, of course, as suggested above imply that the margin for schooling policies related to enrollment of children from poor families in many countries has moved to secondary schooling rather than to primary schooling. The male-female expected schooling gap declined significantly over time, probably at somewhat increasing rates. The male-female expected schooling gap in the 1970s was not significantly different than that in the 1960s, but that in the 1980s averaged—0.4 years below that in the 1960s, and that in the 1990s averaged—0.9 years below that in the 1960s. This implied a decrease from an average male-female expected schooling gap of 1.3 grades in the 1960s to 0.4 grades in the 1990s. Thus there has been a substantial reduction in the gender gap favoring boys in school attainment and, as noted, in Latin America and the Caribbean and in an increasing number of countries elsewhere, a shift towards gender gaps favoring girls.

INCOME

Female expected schooling for a synthetic cohort significantly increases as income increases, though at a diminishing rate. The expected years of female schooling averaged—8.8 grades less for the low-income group than for the high-income group, -4.9 grades less for the lower-middle-income group and B3.1 grades less for the upper-middle-income group than for the

high-income group. These are considerable differences across income groups. These patterns suggest that schooling investments are fairly responsive to income, so that general economic growth tends to result in significant schooling increases, which may have the effect of increasing future growth. The male-female expected schooling gap decreases significantly as income increases, probably at somewhat increasing rates. The estimated male-female expected schooling gap on average has been 1.5 years greater for low-income countries than in high-income countries, 0.9 years greater for lower-middle-income countries than in high-income countries and 0.2 years greater for upper-middle-income countries than in high-income countries. This implies a decrease from an average male-female expected schooling gap of 1.7 grades in low-income countries to 1.1 grades in lower-middle-income countries to 0.4 in upper-middle-income countries to 0.2 grades for high-income countries. Thus, growth on the average tends to favor female schooling more than male schooling.

ADULT FEMALE SCHOOLING

There is considerable emphasis in the literature on the proposition that increased female schooling much more than increased male schooling causes more investment in children, including in their schooling.[31] The variations in female enrollment rates and expected schooling and in gender gaps in these variables are more consistent with variations in adult female schooling than with region, time or income with the single exception that variations in female secondary school enrollment rates are more consistent with regional variations. Thus there are quite strong intergenerational female schooling associations. The expected female schooling for a synthetic cohort increases significantly with adult female schooling (at a diminishing rate). The female expected schooling in medium-adult-female-schooling countries averaged 6.7 years above those in low-adult-female-schooling countries and those in high-adult-female-schooling countries averaged 8.2 years above those in low-adult-female-schooling countries. The male-female expected years of schooling gap decreases significantly as adult female schooling increases. The male-female expected years of schooling gaps in middle- and high-adult-female-schooling countries averaged respectively—1.3 and -1.4 grades below that in low-adult-female-schooling countries. This implies a decrease from an average male-female expected years of schooling gap of 1.3 grades in low-adult-female-schooling countries to basically zero in middle- adult-female-

[31] The empirical basis for this proposition is much weaker than is commonly believed. For example, several years ago I surveyed every study that I could find from any country in the world on the associations between parental schooling and child schooling (Behrman 1997). While it is true that more studies reported larger associations with mother's schooling than with father=s schooling than *vice versa*, the difference was very small—52 percent versus 48 percent. More importantly, these studies do not control for unobserved intergenerationally-associated endowments (e.g., genetics, other aspects of family background) and assortative mating on unobserved endowments—which may make a considerable difference. In Behrman and Rosenzweig (2002), for example, we present cross-sectional estimates for the United States in which both mother's and father's schooling have significantly positive coefficient estimates, but—when we control for endowments using special data—the coefficient on mother's (but not father=s) schooling becomes negative, consistent with more-schooled women (holding all else fixed) in that economy spending more time in the labor force and less time in raising children. This result does depend on context. In rural India where the labor market returns to female schooling are not very high, in contrast, in another study we find a strong impact of female schooling on their children=s schooling (Behrman, et al. 1999).

schooling countries and to -0.1 (i.e., favoring females over males) in high-adult-female-schooling countries.

MULTIVARIATE ASSOCIATIONS

When all four variables—region, time, income and adult female schooling—are included, the overall relations are more consistent with the variations in female schooling enrollments and expected years of schooling and related gaps than if only one of these four variables are considered at a time. In almost all cases, however, the associations with each of these four individual variables are lessened somewhat in comparison with those when only one is included at a time. But the associations with adult female schooling tend to be lessened least, again reinforcing the strength of the intergenerational female schooling associations. The associations with income also remain fairly strong. But those with region and time tend to become much weaker than when they are included alone in the estimates—suggesting that in substantial part they when they alone are included, they are representing primarily differences in adult female schooling and in income. The gender gaps in schooling enrollment rates associated with region, for example, are much smaller or are not significant or even are reversed to favor females (e.g., for secondary school enrollments in Latin American and the Caribbean) once there is control for time, income and adult female schooling. For such reasons, as recognized in Berry (2002), useful classification of countries by different typologies is difficult in part because the central factors for possible classifications are likely to be correlated, with the result that it is difficult to identify what are the most important factors.

CHAPTER THREE
KEY RECOMMENDATIONS CONCERNING THE EDUCATIONAL SECTOR BASED ON AVAILABLE LITERATURE

NECESSARY CAVEATS ABOUT THE IMPLICATIONS OF THE LITERATURE FOR EDUCATIONAL POLICIES TO PROMOTE PRO-POOR GROWTH

There is an enormous literature that presents estimates of associations between years of schooling (and much less frequently, other aspects of education) and a number of outcomes in developing countries, particularly wages but also fertility, health, nutrition, etc. Were it the case that such associations measured the causal effects of education as they apply to the poorer members of society, then it might appear that they would provide a substantial basis for assessing the role of education in pro-poor efficient-growth strategies. However, based on the considerations elaborated in Chapter One, most of these many estimates do NOT provide much information about the role of education in pro-poor growth strategies for a number of reasons. First, most of these studies do not use approaches that recognize the implications of education being determined by behavioral choices. For the most part, instead, they interpret associations as causal under the implicit assumption that education is determined randomly—and NOT associated with family and community background, innate ability, preferences and motivation, and characteristics of educational options, such as the quality of schooling. This means that they are likely to be misleading regarding the impacts of education on the average, and probably even more so for the poorer members of society who are likely to be more poorly situated with regard to many of these unobserved factors. Second. most of these studies focus on the impact of grades (or years or levels) of completed schooling, which is but one aspect of education. Third, most of these studies only consider associations of education (schooling) with various outcomes, but do not integrate cost considerations within a dynamic framework in order to be able to obtain benefit-cost ratios or rates of return to alternative strategies. And often if costs are incorporated, there is confusion between budgetary costs and true resource costs. Fourth, almost none of these studies address the distinction between private and social rates of returns that might warrant public subsidies of education on efficiency grounds. And most of those that claim to do so, in fact so not do so persuasively. Instead, for example, they suggest that Anoneconomic@ effects are Asocial@ effects even though such effects (e.g., on one=s own health or prospects in marriage markets) may be private, not social. Fifth, many studies of the impact of education (again, usually grades of schooling) limit themselves to a subset of the possible impacts that are suggested in the literature (see Knowles and Behrman 2003a,b).

For one or more (and in many cases, all) of these reasons, most of the existing literature on education in developing countries is not very informative regarding what educational policies are likely to be of relatively high priority in promoting pro-poor economic growth. The most solid bases for making policy recommendations, instead, is based on a combination of a small number of mostly recent studies that deal with some of these problems relatively well, in conjunction with *a priori* considerations such as are suggested by some of the aspects of

Chapter One. I now turn to key recommendations on these bases, with references to some relatively high quality studies that underlie the recommendations.[32]

KEY POLICY RECOMMENDATIONS FOR EDUCATIONAL POLICY AND PRO-POOR ECONOMIC GROWTH

1. Recognize that education is more than formal schooling: Education, as noted above, has a number of important components over the life cycle, including pre-school education, various levels of schooling, adult education, formal training programs and informal training/learning through work and other experience. It is important that educational policy not be reduced to schooling policy, though there often are tendencies to do so because of the formal and long-established nature of schooling and the strong lobbies for schooling within countries and internationally. A few careful studies, for example, on both developed and developing countries suggest that the returns may be relatively high for pre-school programs. The benefit-to-cost ratios for the PIDI pre-school program in poor communities in urban Bolivia, to illustrate, are estimated to be from 1.7 to 3.7 with control for unobserved characteristics that affect selection into the program using marginal matching methods (Behrman, Cheng, and Todd 2003). Estimates for the benefit-to-cost ratios of adult basic education and literacy (ABEL) in Bangladesh, Ghana and Senegal. albeit conditional on a number of assumptions changes in which alter the estimates a fair amount, are 19.9 (Knowles and Behrman 2003a; also see box in recommendation 11 below). Because of the paucity of systematic studies on aspects of education other than formal schooling the rest of this section focuses on formal schooling. But considerable resources have been devoted to other types of education and it is generally not known whether these are fairly high return use of resources or if these resources better could be used for formal schooling or other purposes. More attention should be given to evaluating other types of education so that more informed policy choices could be made.

2. Distinguish between private and social rates of return to different types of education and be conscious of policy hierarchies: Private rates of return to investments in education of the poor are important for considering poverty alleviation through improving the education of the poor. If the private rates of return to education for the poor are high, and society has a strong interest in alleviating longer-run poverty, then it may be good social policy to subsidize such education with public funds even if the social rates of return to such educational investments are not so high—depending, of course, on the returns to alternative investments as well as the relative weight on longer-run antipoverty alleviation versus higher income. On the other hand high private rates of return on the average to educational

[32] There are some topics that are of considerable concern in the educational literature that I do not address here because I am unaware of relevant systematic studies that permit inferences relevant for the broader developing world. A leading example pertains to curricular content, which is touched on below only indirectly with regard to schooling quality. *A priori* it would appear the curricular content is very important. Strong English language programs, for example, would seem to be an important factor in the capacity for the Philippines to export many types of laborers and, in combination with strong quantitative training, for India to have become a player in world software markets. There would seem to be high social returns to systematic studies of curricular choices that had broader implications.

investments in themselves may not provide a policy justification for public subsidies to education. They may be high because of large returns to higher levels of education that are rarely attained by the poor (e.g., Brazil over much of the second half of the 20th century), so there is not an anti-poverty justification for subsidies. And that the private rates of returns are high does NOT mean that there the social rates or returns are higher yet, as would be required to justify public subsidies on efficiency grounds.

The difference between the private and the social rates of return is what is important for the efficiency motive for policy and for efficient economic growth. Too often the distinction is not made or the term "social rates of return" is used in a sense different from that defined in Chapter One. For example, Psacharopoulos (1994) is one of a series of well-known surveys that presents estimates of "private" and "social" rates of returns to school. But the "social" rates of return are just the private rates of return adjusted downward to reflect any public resources that are used—with the result that the "social" rates of return are below the private ones. If the estimates were credible representations of the terms used in Chapter One, this pattern would imply that schooling should be taxed more (i.e, subsidies reduced) rather than the subsidies increased (as often as been argued on the basis of these estimates). A basic problem is that such estimates do not consider that private marginal benefits may differ from social marginal benefits (e.g., due to spillovers) nor that private marginal costs may differ from social marginal costs for the same inputs (e.g., due to capital and insurance market imperfections). For another example, returns that are not directly economic are often refereed to as "social returns"—such as the impact of a woman=s schooling on the health, nutrition and education of her family members. But surely the returns to her own health and nutrition are private returns. And it would seem that the returns to family or household members also are private returns from the point of view of her family—just as the returns to a firm from having a more educated worker who increases the marginal productivity of other workers in the same firm is a private return to the firm. This is not to say that such "noneconomic" returns of education are not important, but that clarity is desired in characterizing to what extent they are private and to what extent they are social in the sense defined in Chapter One.

Good policy should be cognizant of both distributional motives and efficiency motives. Some policy changes have the "hackneyed" but still potentially important "win-win" characteristics of permitting better attainment of both distributional and efficiency goals. For example, distributional goals often focus on improving options for the poor and the poor often are subjected more than others to market and policy failures that have efficiency costs, such as in capital, labor, insurance and information markets. Thus there is likely to be an argument for public subsidies to education used by the poor on both anti-poverty and pro-efficiency/growth grounds. But it should be noted that this argument is likely to be relevant primarily for pre-schooling, basic schooling (primary and perhaps secondary), and training programs—but not for general subsidies for tertiary schooling where the proportion of students from poor families in most societies is very small. Indeed, in the absence of compelling evidence that there are significant market failures associated with particular types of tertiary schooling, "win-win" policies in most countries should shift subsidies from tertiary schooling to other types of education, while instituting or expanding "need-based" scholarship funding at the tertiary level for the small proportion of students from poor families who gain admittance.

Finally, whatever the weight on distributional objectives such as reducing poverty and on efficiency/productivity objectives, it is important to recognize that generally policy goals can be attained by many goals, with very different costs. Therefore it is not sufficient to show that education has a significant impact on some outcome to warrant public subsidies for education. It should be demonstrated, in addition, that education is relative high in the policy hierarchies that are discussed in Chapter One. For instance, many studies suggest that there are efficiency reasons for subsidizing female schooling because more female schooling reduces fertility and therefore reduces negative "spillover" demands on public educational and health systems. However, it is not at all obvious that more female schooling to reduce fertility two or three decades in the future is very high in the policy hierarchy for reducing undesired public educational and health subsidies. It would seem much higher in the policy hierarchy to simply correct the pricing for education and health services now to reduce any discrepancies between their private and social marginal costs.

3. Recognize that the rates of return to investing in the education of the poor appear fairly substantial in many cases—comparable or exceeding those to many other investments, with those for investing in females often at least as high as those for investing in males: There are hundreds if not thousands of studies that report positive *associations* between schooling and "good" outcomes such as increased productivity and higher wages, as well as (particularly for female schooling) better health and nutrition, and lower fertility. Most of these studies, however, do not provide compelling evidence that schooling has strong positive *causal* effects on these outcomes because they are undertaken as if schooling were distributed randomly and independently of ability, motivation and family background. Further, many—perhaps most—of these studies do not recognize that schooling effects may be nonlinear and therefore different for the poorer members of society than for the average member of society. A much smaller number of studies control for the determinants of schooling in estimating the impact of schooling for the poor in developing country contexts. They generally suggest fairly substantial returns to investments in schooling, particularly in contexts in which there are technological and market changes, such as during the Green Revolution in Asia (Foster and Rosenzweig 1995, 1996, Rosenzweig 1995). All in all these studies suggest rates of return to education in such contexts that probably are high relative to many alternative investments. For instance, the estimated benefit-cost ratio for a scholarship program in Colombia is 3.3, for the Progresa conditional transfer program in Mexico is 1.7, and for the PIDI pre-school program in Bolivia is 1.7 (Angrist, et al. 2002, Knowles and Behrman 2003a, Behrman, Cheng and Todd 2003, Behrman, Sengupta and Todd 2002). However, as emphasized elsewhere in this paper, there are very few such estimates available and the estimates are sensitive to the underlying assumptions and to the contexts in which such programs operate. Therefore the reader should take away from the available examples that there are possibilities that such programs in other contexts have high benefit-to-cost ratios, but that it is important to evaluate particular programs in the particular contexts in which they are under consideration for implementation and, if they are implemented, to update such evaluations with information gathered from monitoring these particular programs.

The more systematic estimates also in a number of cases suggest higher returns to schooling for females than for males. In some of these the higher returns for female than for male

schooling have been found in terms of economic productivity or wages even with control for unobserved factors related to family background (e.g., family connections, assortative mating on ability and preferences), such as for higher levels of schooling in Indonesia (Behrman and Deolalikar 1995). Such a finding, however, is not likely to provide support for investing more in females on anti-poverty grounds because those females who obtained higher levels of schooling generally are not from poorer families. In a number of cases higher returns to schooling for females than males have been found in terms of outcomes other than economic productivity narrowly defined. In well-known papers then Vice-President of the World Bank Summers (1992, 1994), for example, presents rough estimates that suggest that such returns are considerable in Pakistan. King and Mason (2000) give references to a number of such estimates in different countries. A common assumption is that such returns are additional to those in terms of productivity. The validity of this assumption, however, is not clear. If more schooling induces more time in market-oriented production and less in home production, for instance, the result may be reduced "noneconomic" returns as found by Behrman and Rosenzweig (2002) for the United States.[33]

4. Recognize that rates of return to education depend importantly on market, policy and cultural contexts: The rates of returns to education depend strongly on the overall economic and social environment, and are likely to be the highest in contexts in which there are expected to be needs for ongoing learning due to ongoing changes in markets and technologies. The greater are expected future changes, moreover, the more important is likely to be general education—learning how to learn—relative to specific or technical education (e.g. Rosenzweig 1995). This means that the rates of returns to education are likely to be higher in developing economies that are more integrated into international markets and thus more exposed to technological innovations and international market fluctuations. It also means that rates of return to education are likely to be greater in economies with more fluid labor markets and less restrictions on, for example, worker turnover. For instance, Behrman, Birdsall and Székely (2003) provide estimates of significant impacts of liberalizing foreign-sector and domestic-market policies on the rates of return to secondary and tertiary schooling relative to primary schooling by using micro data from 71 household surveys for 1977-98 for 18 Latin America countries that include about 95 percent of the total population of the region and that permit control for unobserved characteristics that may affect wage levels. Not only are the private returns to education likely to be higher in more dynamic economies in which there are more changes, but the social returns above the private returns are likely to be higher. This is the case because more education facilities not only those with the greater education learning more quickly about the best production choices in a changing context, but such individuals also provide models for others to imitate in adjusting to changes (e.g., Foster and Rosenzweig (1995) analyze an important example with regard to the local adoption of Green Revolution agricultural technology in India). Such considerations suggest that schooling is likely to be more effective in and more important for improving the options of the poor (who tend to be less well-schooled) in economies that are more dynamic and more integrated into world markets. In particularly, increasing secondary schooling for students from poor families may have relatively high pro-poor and growth benefits in such a

[33] See the note in Chapter Two.

context.[34] Perhaps fortunately, for many developing countries, primary enrollments are close to universal so the primary operational margin for increasing schooling enrollments is at the secondary level.

Moreover, there are major changes in age compositions of populations, which have implications for pressures for resources for schooling and other forms of education and for the rates of return to education. Looking forward, almost all developing regions will be aging, and, indeed, the average ages in Latin America and the Caribbean and in South Asia are projected to increase in coming decades more than in any other regions of the world (Behrman, Duryea and Székely 2003). Based on past experiences, that means that resources for schooling and possibly other forms of education are likely to increase on a per capita basis because younger cohorts are becoming relatively smaller (e.g., Truong, et al. 1996, Behrman, Duryea and Székely 2003), though with a greater lag in Africa than elsewhere. On the other hand, these changes in the age composition of populations also mean that workers entering the labor market in relatively large cohorts in the next decades may face lower returns to schooling due to their relatively large sizes, as found historically, for example, for Brazil (Behrman and Birdsall 1988b).

Further, it should be noted that such considerations mean that the important margin for focusing on pro-poor and efficient-growth educational policies are likely to differ among countries and, in particular, change with overall development. While it may have been in attaining universal primary schooling in many developing countries two decades ago, for example, now it increasingly is in expanding secondary schooling in many of those same countries in the first part of the 21st century, in part because of the substantial changes in education experienced in recent decades (Chapter Two) as well as rapid globalization and aggregate policy changes such as noted above.

There also are likely to be important gender dimensions of context that go beyond the obvious point that females can not reap labor market returns in certain sectors if cultural or legal restrictions preclude them working in certain sectors. Based on experiences in East and South Asia, it appears that opening up economies is likely to shift production towards sectors in which women may have a comparative advantage due to more knowledge and skill-based technologies, less rewards for strength and more rewards for manual dexterity, and greater rewards for disciplined performance. Another gender-related implication, of course, is that laws designed to "protect" women (e.g., by limiting their hours, by excluding them from certain occupations), if effective, may preclude them from activities in which they could gain higher returns.

---

[34] An important factor underlying the apparently high returns to schooling in the so-called "East Asian miracle" economies (e.g., see Page, et al. 1993) appears to have been market and macro conditions such as discussed i the text for which the rates of return to schooling were relatively high. Schooling alone, in the absence of such conditions, did not lead to high economic growth. For instance, the country that had the highest level of schooling in Asia conditional on per capita income in the 1960s was not one of the "miracle" economies, but the Philippines. Also relatively high in terms of schooling were other non-miracle countries such as Sri Lanka and, at the secondary school level, India. (See Behrman and Schneider 1994).

One obvious, but none-the-less, important implication of the dependence on context is that there is not likely to be a "one-size-fits-all" specific silver bullet such as more textbooks or teacher training that will improve education, or even some component such as basic schooling, everywhere. Conditions vary too much across countries, within countries and across types of educational institutions. "Best-practice" specific improvements elsewhere may suggest narrowly-defined specific changes that should be considered, but not blindly imitated. Berry (2000) discusses in greater detail some aspects of typologies that are helpful for thinking about policy alternatives for the overall project—including typologies based on structural features, on temporal conditions, and on policies. He concludes:

> "Effective use of a set of typologies to guide policy in a given country involves, first and foremost, identifying where the country fits in the typology. It also calls for recognition that the distinctions on which most typologies are built are matters of degree rather than of kind. Finally, several typologies may provide different but consistent types of policy guidance for a given country."

It is helpful, in thinking about policy options for pro-poor efficient growth in education as in other areas to follow this advice. The typologies laid out can help to place a country within a subset of the developing world that, for this purpose, are similar and from which more probably can be learned than from economies that are very different. But it must be kept in mind in thinking about educational policies in any particular context that these typologies are categorizations to help thinking about options, not straight-jackets that preclude adaptation to particular circumstances. At times the most fruitful policy options may arise not from considering the experiences of similar countries, but by thinking Aoutside of the box@ in which a country is placed in such typologies.

5. Focus on the parts of the lifecycle in which the returns are likely to be highest and include health and nutrition support among the policy tools: Though investments related to education may take many forms over the life cycle as noted in recommendation 1 above, the rates of returns to such investments seem to be relatively high often for investments early in the life cycle, including investments in better health and nutrition from conception onwards (which effectively may mean better health and nutrition for women prior to conception), particularly for poorer segments of most populations. Careful studies using longitudinal data from Bolivia, rural Guatemala, Pakistan and the Philippines all provide strong evidence (Alderman, et al. 2001, Behrman, Cheng, and Todd 2003, Glewwe, Jacoby and King 2001, Glewwe and King 2001, Martorell 1999, Ramakrishnan, et al. 1999).

There also may be high returns to investments in school-age children that improve their health and nutritional status. Miguel and Kremer (2001), for example, provide experimental evidence of such an effect with deworming programs in Kenya, including external "spillovers" to the benefits of children not treated through reducing the probability that they have problems with worms in settings where certain infectious diseases are endemic. Two years after the intervention began, the school based de-worming program had led to at least a 7 percentage point average gain in primary school participation in treatment schools representing a 25 percent decline in absenteeism due to the improved health of both treated and untreated children, because of the infectious nature of disease. Knowles and Behrman (2003a), for another example, estimate benefit-to-cost ratios of over 30 for iron supplement

programs for children age 13-15 in high anemia areas. The evidence is mixed from the few systematic studies of school feeding programs, a popular approach directed towards improving nutrition and school attendance. Such programs may not have much effect on the children concerned because they may largely substitute food obtained in school for food at home (though, of course, other household members may benefit). A school feeding experiment in the Philippines using random assignment compared no intervention, school feeding alone or school feeding combined with parent-teacher partnerships and does not find any statistically significant improvement in dropout rates after two years (Tan, Lane and Lassibille 1999). On the other hand, an experimental evaluation of the impact of free breakfasts in Kenya suggests 30 percent increases in participation (Vermeersch 2002; see recommendation 9 below).

In any case, many policies that improve health and nutrition, as well as out-of-school education, for children from poor families from conception on may have "win-win" characteristics because the poor are most likely to benefit from such programs given that they are most likely to be malnourished or at risk of most diseases and they also are most likely to be affected by imperfect information and capital markets that preclude them acting as effectively as they might in such circumstances.

6. Improve school access for basic education: Primary school access currently exists for the majority in most countries, even though not in some more remote rural areas or urban slums and squatter settlements. Secondary school access is available on a much more selective basis and in many systems, only a minority of primary school leavers can find places in public secondary schools due to supply constraints. School access policies have included a range of approaches including: (1) building new schools, (2) multiple shifts to use the same school facility more intensely, (3) development of alternative schools and (4) long-distance learning.

School construction: The impact of a major Indonesian school building program from 1973 to 1979 on grades attained and future wages for men has been estimated by exploiting differences in the number of schools constructed by region and differences in outcomes across cohorts (Duflo 2001). Comparisons of cohorts entering school before the intervention with those entering school after the intervention reveals positive and significant effects of the building program on average grades attained, the percent completing primary and adult wages. These results suggest that the program easily paid for itself in terms of increased earnings in later life under the assumption that there were not other labor market changes that were correlated with the pattern of schooling construction (a plausible assumption given the regional variations in timing unless, of course, the national government favored first the regions in which they correctly anticipated there would be relatively great economic growth—see Chapter One). Because of the scope of the program, those for whom schools became accessible tended to include many poor.

If boys and girls attend differ schools, as in a number of countries, then differential school availability may underlie educational gender gaps. A study of rural Pakistan exploited the existence of single-sex schools to estimate to what extent the considerable gender gap in cognitive skills (almost 80 percent) reflected factors on the demand side (e.g., household

behaviors, perhaps anticipating gender differences in returns to schooling) versus factors on the supply side (e.g., gender differences in the availability of schools in relatively close proximity to potential students= homes). Cognitive achievement production functions were estimated, with control for the determinants of the relevant behavioral inputs including time in school using a framework such as sketched out in Chapter One to guide the estimation by exploiting exogenous variations in prices and other factors. Conditional on this framework, the estimates suggest that about a third of the gender difference in cognitive achievement was due to differential availability of primary schools for girls than for boys (Alderman, et al. 1996b).

These two analyses suggest the possibility that subsidies for school expansion may have significant "win-win" payoffs in terms of reducing poverty, increasing growth and reducing gender gaps in situations in which many potential students do not have access to local schools. But they do not address the question of whether there is any justification for limiting such subsidies to schools operated by the public sector or whether there may be other preferable alternatives, such as long-distance learning, for attaining the same objectives.

Multiple-shifts: Multiple shifts are used in many situations in which the binding constraint is perceived to be physical infrastructure. For example, in Egypt the same building often is used twice a day but the school day is shorter, only the core subjects are taught and the teaching staff changes with each shift. There is a negative cross-sectional association between girls=, but not boys=, retention rates with these double-shift schools (Lloyd, Mete and Sathar 2002). It is not clear that this is a causal effect even though children attend local schools rather than choosing (or having their parents choose) which school to attend because it is not clear, and indeed seems unlikely, that the assignment of multiple shifts is random across various types of neighborhoods (see Chapter One). But even if there is a negative causal effect on girl=s school retention, multiple shifts may—or may not be—an effective way to increase access. To make that assessment private and social benefit-cost ratios (or internal rates of return) would have to be estimated for comparison with alternatives.

Alternatives to public schools: Some countries in some time periods have precluded private schools from operating. But in other countries, such as in East Asia and parts of Latin America and the Caribbean and Africa, they long have been options. Private schools are likely to increase the supply of schooling places unless they "crowd out" the supply of public schools (e.g., through reducing the political support for such schools). Unfortunately, there is no evidence on the extent to which such crowding out occurs. In any case, even where private schools are allowed to exist and in some cases flourish and possibly provide education of higher quality (see recommendation 7 on school quality), they generally do not qualify for some public subsidies (see recommendation 9 for some examples), a form of discrimination that is likely to increase inefficiencies.

In some contexts there has been an effort to increase school access by encouraging the development of alternative schools, generally related to NGOs, which may provide more flexibility and ease of access than standard public schools. Probably the best-known example is the non-formal primary education program in Bangladesh started by BRAC on a pilot basis in 1985 to improve access to the hard to reach poor rural girls who had missed out on school (Ahmed et al 1993). By 1993, BRAC was running over 8000 schools to provide literacy

skills to adolescent girls as well as to younger girls who still can have the opportunity to be mainstreamed back into the formal schooling system. There has been considerable expansion since then to account for about a quarter of all school enrollments in the country. Another example is the Save the Children community school project in Mali designed to develop low-cost alternatives to conventional public schools (Mushkin 1999). In Egypt, the introduction of community schools in disadvantaged areas for girls who had missed the entry age has been credited with a substantial increase in girls= enrollment. Unfortunately almost none of these types of schools have been systematically evaluated.

An exception that included an evaluation as part of the pilot phase of the program=s introduction is the Quetta Girls= Fellowship program in Balochistan, Pakistan that began in 1994 (Kim, Alderman and Orazem 1999). Subsidies were provided to private schools to establish new facilities in urban slums. The size of the subsidy was a function of the number of girls enrolled. In the start up phase of the project (1994-96) the program was introduced in ten slum areas, but with random placement in one of three sites in each area (and with the other two sites as controls). In the first two years of the program, enrollment for boys increased by 24 percent and for girls by 28 percent The authors concluded that this was more cost effective than several alternatives including income transfers to poor households or the construction of new public schools. The program has continued to expand since the pilot phase from 11 to 40 schools and from 2000 to 10,000 students.

Long-distance learning: Long-distance learning has been used increasingly, particularly to attempt to reach relatively remote areas. In Mexico, for example, even though most rural communities with more than a few hundred inhabitants have primary schools, many small rural communities do not have secondary schools—but many of these do have access to *telesecundaria*, a form or long-distance schooling using television. Once again, however, I am not aware of systematic evaluations of such efforts.

7. Consideration should be given to improving educational quality, not just to increasing the quantity of education: Improving educational quality, not only educational access, is likely to have strong positive effects for the poor. Indeed, there would be substantial gains were, for example, the quality of pre-school programs and schools used by the poor increased to the quality levels of pre-school programs and schools used by the rest of society (Behrman and Knowles 1999 document fairly significant school quality differentials associated with parental income in Vietnam despite relatively great emphasis there on equity). There are a number of studies for developed and developing countries that address questions related to the impact of quality of schooling, particularly on labor market outcomes, perhaps through cognitive achievement that can be viewed as produced by time in school, quality of school and other inputs (e.g., Alderman, et al. 1996a and Behrman, Ross and Sabot 2003 for rural Pakistan; Behrman and Birdsall 1983 and Behrman, Birdsall and Kaplan 1996 for Brazil; Boissiere, Knight and Sabot 1985 and Knight and Sabot 1990 for Kenya and Tanzania; Behrman, Rosenzweig and Taubman 1996,Card and Krueger 1992a,b Cawley, Heckman and Vytlacil 2001; Murnane, Willet and Levy 1995 for the United States; Glewwe 1996 and Glewwe, et al. 1996, for Ghana). Such studies are based on the subsample of respondents for whom wages are observed, typically modeling selectivity into such subsamples by controlling for labor supply decisions using standard methods to control for

selectivity. These studies not only find significantly positive effects of school quality, but they often find that treating school quality separately from time in school affects inferences about schooling impact, in some cases reducing considerably the estimates impact of time in school (e.g., for Brazil, not controlling for school quality results in an overestimates of the returns to time in school of about 70 percent). Those that compare the rates of return to improving schooling quality versus increasing quantity in developing countries (e.g., Brazil, Pakistan), moreover, find the rates of return as least as high to increasing quality, in part because the additional opportunity cost of students= time is basically zero for increases in quality for a given quantity—but may be considerable for increased quantity for a given quality—under the assumption that supply capacity constraints are not binding (if they are, then questions of access arise, as discussed above).

There also is a smaller literature using household survey data linked with school characteristics to investigate the role that school quality has in enrollment and attendance decisions. In Pakistan, for example, where primary girls schools are staffed by female teachers, teacher absenteeism can be a particular problem in rural areas where many teachers are not from the local area. In a recent study in rural Punjab and NWFP, the presence of a local female teacher was found to be even more important than the presence of school in encouraging parents= to enroll their girls in school, though obviously the two characteristics interact in the sense that whether there is a local female teacher is conditional on whether there is a local school (Lloyd, Mete and Sathar 2002).

There also are a number of studies that consider directly the determinants of cognitive achievement based on school samples. Most of these studies have serious limitations because they use selected samples (i.e., those in school) without controlling for what determines the selection into the subsamples, often have confusing specifications that mix production function (e.g., textbook availability, teacher characteristics) and dynamic demand decision rules (e.g., family income) on the right-hand side that makes interpretation very difficult, often ignore factors that affect learning outside of schools, and generally do not control for the endogenous choices of individual, family and school inputs. Nevertheless, they have received considerable attention, perhaps because of their direct use of school characteristics and school data, often by scholars and analysts in education. Many elements of school quality have been included in these studies: class size, the number of and qualifications of teachers, the physical infrastructure, learning materials and management structure.

Class size and teacher/student ratios have received the most research attention in this literature. Based on a review of 30 school based production functions studies in developing countries, Harbison and Hanushek (1992) and Hanuchek (1995) conclude that there is no compelling evidence to support policies to reduce class size (though Kremer 1995, in a comment on the latter, raises questions about this conclusion). More recently, Hanushek and Kimko (2000) have estimated a positive but statistically insignificant relationship between pupil/teacher ratios and international tests in math and science. Recent studies in South Africa and Israel, however, have exploited so-called “natural experiments” to raise questions about Harbison and Hanushek=s conclusion. The study in South Africa is based on a representative household sample linked with district data on schools (Case and Deaton 1999). The large variability in pupil-teacher ratios across black schools was assumed to be outside the control of students and

their families because they had little voice in school budgets and no freedom to move out of the community in search of better schools. This study estimates that decreasing the student-teacher ratio from forty to twenty (the approximate means in black and white school respectively) would result in an increase in grade attainment by 1.5 to 2.5 years and a significant increase in students= reading test scores as well (conditional on years of school attendance). However to interpret the results of this study as causal requires the additional assumption that pupil-teacher ratios are random, not purposively allocated in response to some school or community characteristics (Chapter One), which is not entirely plausible. The study on Israel takes advantage of an unusual rule (AMaimonides' Rule@) about class size that sets a firm maximum of 40 for each class (e.g., class size is immediately cut in half and another teacher is hired when total enrollment increases so that otherwise class size would exceed 40). This nonlinearity in class size reduces the chance that class size is correlated with unobserved determinants of learning. The estimates indicate a significant negative effect of class size on reading and math scores (Angrist and Levy 1999). Further evidence from a longitudinal study in Tamil Nadu supports these conclusions (Duraisamy et al., 1997); a sharp rise in enrollment triggered by numerous measures adopted by the state to encourage enrollment including free meals, uniform and books as well as special cash grants to school that retain girls and consequently an increase in student/teacher ratios have impacted negatively on the pass rate for the statewide 10^{th} grade exam. In this case, however, the result simply may reflect that the students induced to enroll by the program are less-well prepared than the average student previously enrolled, and therefore pulled down the average.

On the other hand, Behrman, Sengupta and Todd (2002) find that in rural Mexico the random assignment of scholarships to poor families in different communities increased school enrollments among those families, but did not have negative effects on the schooling attainment of children from other families in the same communities, so there is not evidence of negative crowding or spillover effects. Similarly, the lack of spillover effects on nonparticipants who attend the same schools as do program beneficiaries who increased their enrollments are reported for the Bangladesh food for education program (FFE) (Ahmed and Ninno 2002). A randomized evaluation of a program in India that provided a second teacher, where possible female, to one-teacher schools, suggests that it did not improve test scores, although it did attract more girls to school (Banerjee and Kremer, 2002). Collectively, thus, these studies provide more persuasive evidence than what existed at the time of the Hanuchek (1996) survey that increases in class size beyond some limit have negative impacts on education—though the Bangladesh, Indian and Mexican experiences suggest that the question is still open.

Several other recent systematic studies have investigated the relevance of other aspects of school quality. Infra-structural improvements can be an important aspect of school quality in settings with difficult weather conditions and poor building materials. The existence of unusable or leaky classrooms was a statistically important factor in explaining variations in test scores and student retention (Glewwe and Jacoby 1994 on Ghana). A recent longitudinal study of the impact of an education program in Peru primarily devoted to funding community-based school renovation projects (a part of the larger Peruvian Social Fund known as FONCODES) found a sizeable impact of the program on the school attendance of younger children (Paxson and Schady 1999).

Some well-designed random experiments in Kenya have explored the effects of the provision of uniforms and textbooks (Kremer et al 1997), text books only (Glewwe, Kremer and Moulin 2000) and flip charts (Glewwe et al 2003) on various school outcomes. These studies have been small in scale and short in duration. None of these interventions found positive effects on test scores. However, the interventions themselves may have been poorly designed and chosen, although scientifically researched. The textbooks may have been too advanced for the students and the teachers may have been not properly trained to use the materials. However the lack of significant effects in these cases is a result that may be important in understanding what works and what does not work—and contrasts with retrospective studies from the area that incorrectly suggest that these interventions have significant effects, apparently because they were adopted in schools and communities in which there are important unobserved characteristics that are conducive in other ways for more education (Kremer 2003 and Chapter One)

Several studies have examined the gains of decentralization to give more attention to the special educational needs of children in disadvantaged settings through greater teacher autonomy.[35] A matched comparison design to follow a sample of treatment and non-treatment schools in Nicaragua reports no evidence that students in schools participating in the Autonomous School Program had better test scores. However, when schools, regardless of participation in the program, showed higher levels of teacher autonomy, student test scores were significantly higher (King and Özler 1998). The Escuela Nuevas were introduced in rural Colombian villages that previously did not have a primary school. These multi-grade schools were designed with flexible promotion, teacher training and instructional materials adapted to slow learners and local needs, intensive supervision, community involvement and student governance and introduced into rural villages. Improved test scores and less dropout among students attending Escuela Nuevas relative to those attending traditional schools in similar villages were found in a cross-sectional study with a quasi-experimental design (Psacharopoulos, Rojas and Velez 1993). In the random experiment in the Philippines, the introduction of multi-level learning materials combined with parent-teacher partnership resulted in a reduction in dropout rates and better language learning (Tan, Lane and Lassibille 1999). After the civil war in el Salvador in 1991, a community managed school program (EDUCO) was introduced to expand school access to remote rural areas. Communities targeted for the new schools had high rates of malnutrition and repetition rates. The program provided for community education associations composed of elected parents to manage the school, in particular to be responsible for administrative, management, hiring and firing of teacher, equipment and school maintenance and teacher supervision. An ex-post cross sectional analysis based on a sample of EDUCO and traditional schools showed that, while the student absentee rate was lower in EDUCO schools, there was no appreciable effect of participation in the program on student learning outcomes (Jimenez and Sawada, 1999). Because the evaluation was school based, it could not assess effects on enrollments—one of the intended outcomes of the program.

[35] Decentralization of the operations of schools can occur independently of decentralization of financing of schools. In the presence of substantial variation in local poverty rates, from the perspective of a pro-poor efficient growth strategy it may be desirable to decentralize school operation but to still have transfers from better-off to poorer areas to finance schools.

Female and male students may have different experiences in the same school because of differences in curricular opportunities within schools; differences in treatment by individual teachers and differences in rules, regulations and administrative practices. While some differences in gender treatment are legislated by policy (e.g. home economics for girls or required withdrawal if pregnant) and therefore uniform across schools, differences in gender treatment in the classroom, in the absence of gender training for teachers, are likely to vary both within and across schools. Few studies have attempted to quantify differences in gender treatment across schools in a way that would allow the measurement of its impact on school outcomes. Teacher attitudes toward the teachability of boys and girls was found to be a significant factor in explaining differences in exam scores in Nairobi, Kenya (Appleton 1995). Various measures of gender treatment, controlling for family and other more traditional school quality factors, were found to be affect significantly the probability of dropout for girls but not for boys in rural Kenya (Lloyd and Mensch 2000).

Thus the better of the available studies often suggest that quality improvements have a significant impact on education of students from poorer families. But many of these studies, though with notable exceptions, do not present estimates of the benefit-cost ratios or the internal rates of return to such quality improvements. Therefore, though the changes considered may have positive impacts, it is not clear where they stand policy hierarchies (Chapter One). Moreover, many of these studies other than the ones concerned with decentralization say little about the mechanisms for improving school quality. (A recent exception is Lavy-s (2002) study of Israeli programs in which he finds that, based on the same cost, monetary teacher incentives are more effective than providing additional conventional resources to schools.) Many people understand the message to be that such quality improvements should be made by fiats from Ministries of Education or other such governmental authorities. But, particularly in the presence of heterogeneous communities and needs, this may be a *non sequitur*. More attention needs to be given to mechanisms for attaining such quality improvements.

On *a priori* grounds, the best mechanisms would not seem to be ones that limited the provision of educational services or subsidies for enhancing the quality of such services to certain classes of providers. To the contrary, such prescriptions would seem to limit the incentives to improve school quality. Moreover, as noted above, there is some evidence that private suppliers more effectively provide better school quality for poor students, as is reflected in the outcome of the random assignment by lottery of poor students to private versus public schools in urban Colombia (Angrist, et al. 2002). Therefore it is important to consider how to improve incentives to elicit better educational quality for given resources. Such incentives may be improved through a variety of mechanisms, including increasing the influence of actual and potential clients through portable educational scholarships (the use of "demand-side" channels through which to channel resources; see recommendation 9 below), decentralization of many educational decisions to the local level, and creating a level playing field for alternative educational service providers rather than discriminating, for example, by whether a particular educational service provider is under public, NGO or private ownership.

8. Before using mandates to attempt to increase education, obtain more information about their effects: Most countries have compulsory schooling, though a number—mostly in sub-Saharan Africa—do not. Among those that do, the legally mandated length of compulsory education varies from 4 to 11 years (Tomasevski, 2001). However, there is little evidence that enrollment is often enforced. Effective enforcement requires that there be an adequate supply of school places for the eligible population. In many developing countries universal access at the primary level has been achieved, though in others substantial additional investment are required to attain this target. The only relatively well-designed study examining the impact of compulsory schooling laws on retention until the legal age of exit of which I am aware is based on state variations in compulsory schooling laws in the United States. This study suggests that these laws can be effective in compelling a small proportion of students who would otherwise drop out to remain in school until they attain the legal dropout age (Angrist and Krueger, 1991). But for the most part the high association between compulsory schooling attendance laws and schooling attainment across societies and over time seems to reflect that such laws endogenously respond to changing schooling incentives and norms about schooling attendance rather than being a causal mechanism through which large changes in education can be mandated.

An important administrative rule governing eligibility to enter school relates to the appropriate or required age of entry. If parents are not yet ready to send their children to school by a legally-determined age, an inflexible age of entry may work against universal schooling goals by limiting the window of time during which children are eligible to enter. Some have suggested, for example, that the enforcement of a rigid age of entry in Egypt led to the permanent exclusion of many girls from school. On the other hand, flexible age of entry rules, while more inclusive, have led to other problems. In Kenya, where the rules governing the timing of entry are very flexible, children continue to enter the system until age 11 (Montgomery and Lloyd, 1999). The resulting age heterogeneity within the classroom appears to present difficult challenges to teachers who have to deal with students at very different developmental stages.

Other administrative rules discriminate specifically against young women by prohibiting school attendance if pregnant, married and/or a parent. These administrative rules are based on assumptions about the incompatibility of roles and concerns about effects of role models. South Africa is one of the few countries in sub-Saharan Africa with a very progressive policy on school attendance for girls who are pregnant and/or mothers.

The bottom line is that, for the most part, these administrative rules have not been systematically evaluated. Therefore it is difficult to know whether they are having their intended effects and/or having important unintended effects.

9. Consider using incentives in the form of scholarships for students from poor families to increase their education: Scholarships or conditional grants for attending schools or obtaining other forms of education, particularly for poor households, have received increased attention recently. If such scholarships are directed towards basic education for the poor (e.g., pre-school programs, primary and perhaps secondary schooling, basic training), they also are likely to have "win-win" characteristics because capital and insurance markets are likely to

be particularly imperfect for human resource investments and to be more binding (in the sense that self- or family-financing is less likely) for poorer households. There is evidence from several countries that such programs can increase schooling significantly.

One of the first of the conditional grant programs was the food for education (FFE) program in Bangladesh where poor households with primary school aged children are eligible for monthly rice or wheat allotments according their number of children conditional on their children maintaining an adequate attendance at school. The program was begun on a large scale pilot basis in 1993 and has grown to the point that it covers 27 percent of all primary schools and benefits about 13 percent of all primary school students. Two ex-post evaluations, one using an existing survey (Ravallion and Wodon 2000) and one using a survey specially designed for the evaluation (Ahmed and Ninno 2002) have found strong statistically significant effects of the program on enrollment, with girls showing the strongest response.

In 1994, the government of Bangladesh also began a female secondary stipend program in rural areas to increase the enrollment of secondary school age girls, to improve their secondary school completion rates and to delay their age of marriage. This stipend program, which combines both tuition coverage and a cash grant to the girl that increases with age, is complemented by other program components designed to enhance school quality which vary according to region. A girl is eligible to receive the stipend if she maintains some minimum level of performance, has a 75 percent attendance record and remains unmarried. Using household and school data from two points in time, an ex post evaluation of the program taking advantage of differences across villages in the timing of its introduction finds, using village fixed effects, that the duration of the program had strong statistically significant effects on the enrollment of girls but no impact on the enrollment of boys and that the school enrollment benefits of the program accrue disproportionately to girls from families that owned more land (Khandker, Pitt and Fuwa 2003). As of 1997, the stipend budget alone was estimated to represent over 13 percent of the educational sector budget. Given the large expense of the program, these conclusions might suggest that the program needs to better targeted; at the moment girls are eligible regardless of family income or land ownership.

In Latin America there has been a recent expansion of conditional scholarships for school attendance of students from poor households as part of an explicit antipoverty strategy. These programs include PROGRESA in Mexico which began in 1997, the RPS program in Nicaragua begun in 2000 (Maluccio 2002) and most recently Bolsa Escola in Brazil which was expanded to a national scale in 2001. The best documented in this family of programs is the PROGRESA program in Mexico (see Appendix C). This is because the introduction of the program included a pilot evaluation phase in which 326 intervention and 180 control communities were randomly selected from 506 eligible communities and their experiences were compared over a two year period. The conditional cash grants are given to mothers for all children in grades 3-9—the higher the grade, the higher the grant with girls in secondary

receiving a slight premium over boys.[36] Thorough assessments after two years of the program how strong positive effects on schooling outcomes including less grade repetition, better grade progression (particularly from primary to secondary), lower dropout rates and higher school reentry rates among dropouts. Indeed, the program is encouraging even younger children who are not yet eligible for the subsidies (Behrman, Sengupta and Todd 2002, Schultz 2000, 2003). The long-run impact of Progresa is estimated to be about a 10 percent increase in grade attainment in poor rural communities, with private purely economic productivity benefits relative to governmental resources devoted to this program on the order of magnitude of 1.7 (with a 5 percent discount rate). Recently Progresa has been expanded to urban areas and through the secondary schooling level, and the plans are underway to introduce further incentives for successful completion of secondary school in the form of a bonus to high school graduates from poor families that can be used for tertiary schooling, purchasing health insurance, participating in housing programs, participating in micro credit programs or (with a three year lag) cashed in.

Bolsa Escola is a somewhat similar program being implemented in Brazil. A study using national data to simulate some of the potential long-term effects finds that the most critical factor in the program=s likely success is the conditionality of the grant (Bourguignon, Ferreira and Leite 2002). Unconditional grants to the poor are predicted to have little effect on enrollment; increasing the size of the grant with age is not projected to have a major incremental effect. Todd and Wolpin (2003) report a similar result for Progresa, based on a structural model of household behavior that they estimated with the Progresa data. Such results raise some questions about the conditional grant programs. If they are effective because of a capital market imperfection that precludes poor households from being able to finance education investments at the true social marginal cost of capital, then they should have the same impact whether or not they are conditional. But since the schooling results are much less if they are not conditional, it would appear that such capital market constraints are not a central problem in such schooling attendance and that the decision-makers in these poor household place higher priority on other uses of the additional resources than on sending their children to school. Perhaps the conditionality is warranted, nevertheless, on distributional grounds —that is, it assures that the additional resources that the family receives go substantially to investments in the children. But, if so, it is not clear that this is a pro-poor on intergenerational grounds since, in growing economies, at least on average children are expected to be better off than their parents.

One potential problem with conditional transfers is that the people administering the program may not enforce the conditionality in practice. School meals may therefore provide a stronger incentive to attend school than take-home rations ostensibly conditioned on school attendance (Sen 2002). School participation has been found to be 30 percent greater in 25 Kenyan pre-schools where a free breakfast was introduced than in 25 comparison schools.

[36] The rationale for giving higher scholarships to girls than to boys was that prior to the program, enrollment rates were higher for boys than for girls. Enrollment rates, however, are an imperfect measure of education. Prior to the program in fact average schooling attainment was higher for girls than for boys, but enrollments rates were higher for boys because they failed and repeated grades more than girls and therefore had to enroll for more years to reach the same level of schooling attainment (Behrman, Sengupta and Todd 2002).

The provision of meals cut into instruction time, however. Overall, test scores were 0.4 standard deviations greater in the program schools, but only if the teacher was well-trained prior to the program (Vermeersch 2002). But, as noted in recommendation 5 above, in another experimental program in the Philippines no significant effects of school feeding are found (Tan, Lane and Lassibille 1999).

A potential major limitation of many of these conditional transfer programs, whether they provide payments in terms of money or food, is that they tend to limit the use of the scholarships that are awarded to public schools. As a result, the programs do not increase incentives for the providers to improve their services to compete for students as much as they might. To the contrary, they may discourage entry of new educational providers by continuing the institutionalization of favoring a certain class of providers based on ownership. While there probably is a case for providing conditional transfers to students from poor families on redistributional pro-poverty grounds and quite possibly on efficiency grounds, there is nothing about these objectives that necessarily requires that schooling be provided only by the public sector (some of the pros and cons of such a requirement have been discussed in debates about school vouchers, e.g., see Ladd 2002 and Neal 2002).

10. Improve information about what educational providers are providing: Current information is imperfect and the incentives for educational service suppliers to provide information is insufficient, in part because of the public-good characteristics of information (Chapter One). Therefore policies should require provision and publicizing of information about the value added of educational providers (often the focus is on indicators of the educational level of students/graduates, which is misleading because of variations in their initial education when starting at an institution). However at the same time care needs to be taken regarding what information is required to assure that the particular information that is required does not distort incentives in some negative way. This is tricky because of the well-known tendency to Ateach towards the tests.@ Research on the U.S. suggests, for example, that accountability systems that focus on year-to-year gains in achievement (as opposed, for example, to some combination of changes and levels or averaging out changes over more years) tend to over-reward and over-penalize small schools (simply because the noise in such measures averages out less) and that systems that have only one (Apass-fail@) or a few cutoffs tend to focus school efforts on students who otherwise would be right below the cutoff(s) (Hoxby 2002/2003, Kane and Staiger 2002). Also, an evaluation of a Kenyan program in which parent-run school committees provide gifts to teachers whose students perform well finds that teachers respond to the program primarily by increasing effort devoted to manipulating test scores, rather than by increasing effort at stimulating long-term learning; test scores for pupils who had been part of the program initially increased but then fell back to levels similar to the comparison group at the end of the program (Glewwe, Ilias and Kremer 2002).

Improvement of information, thus, probably requires ongoing evaluation of the indicators used and their effects by an independent body. But it also should be recognized that it always will be difficult to provide all of the desired information in a timely fashion and, indeed, may not be efficient to do so. Therefore it also is desirable to establish mechanisms that increase the incentives for educational systems to perform in effective ways even in the absence of perfect monitoring. This may raise some tensions between, for example, focusing on public

sector systems in which ostensibly (though not necessarily in fact) monitoring is relatively good and encouraging a diversity of educational service providers that may increase competitive pressures for effective service provision but be harder to monitor.

11. Improve information and evaluation of the rates of return to alternative strategies for increasing education: This review of what is known about education and the attainment of pro-poor and efficient growth objectives repeatedly has noted big lacunae in our knowledge—some possibly important topics have hardly been researched including those pertaining to most aspects of education other than formal schooling as well as many aspects of formal schooling, most of the research on topics that have been studied is limited to associations that may be misleading regarding true causal effects, rarely are many of the multiple effects of education incorporated into the analysis, many of the systematic studies that can with more confidence be interpreted as presenting causal effects do not integrate those estimates with cost information to obtain benefit-cost ratios or internal rates of return for comparisons with other interventions, there is very little attention to the distinction between private and social rates of return that is at the heart of the efficiency/productivity motive for policy, and all too often the sensitivities to policy, market and cultural contexts are not explored.

The rates of return to better information and better analysis on which to base policy decisions appear to be high. The resources used for data collection and evaluation of even the best-evaluated programs are a small percentage of overall program costs—or of what might be expected to be saved with improvements in programs. The relevant resources for better evaluation include financial resources, but they also include sufficient will and/or incentives to assure that critical steps in the evaluation are implemented in a timely fashion despite greater pressures—importantly, for example, collecting baseline data prior to program implementation despite pressures to proceed with programs and using as much as possible approaches that emulate good experiments with random assignments to treatment (at least for as programs are being phased in). A central point. of course, is that many programs are not very successful and should be modified or abandoned, so care need be taken that attention be paid to what is learned about unsuccessful programs and not just successful ones, though there appears to be a "publication bias" or "publicity bias" towards apparently successful programs. (This point is also made by Kremer 2003.) Appendix C summarizes some features and possible improvements in one relatively good recent experimental evaluation, Mexico-s Progresa program.

However, while improvements in information collection, monitoring, and better evaluation would have a high payoff, given the many imperfections in real world information, I should make clear that this is not a silver bullet either. To illustrate, the text box summarizes the benefit-cost ratios for selected educational programs for which information could be pieced together about various dimensions of costs and benefits and their time paths to undertake such calculations. The penultimate column gives the estimated benefit-cost ratios. The last column gives what the authors refer to as the "plausible range in estimated benefit-cost ratios." What is striking is how large these ranges are even though these examples were

limited to ones in which relatively good and complete information is available.[37] An important implication is that, while there may be significant returns to improving information about and improving analysis of educational programs and their implications for pro-poor efficient growth, it must be recognized that there are limitations to what can be learned and, also important, how quickly it can be learned. Therefore it is important to continue to attempt to improve mechanisms for making education more effective on the basis of *a priori* analysis and the imperfect existing information at the same time as efforts are made to improve such evaluations.

Educational Program	Estimated benefit-cost ratios (with 5 percent annual discount rate)	Plausible range in estimated benefit-cost ratios
Scholarship program (Colombia)	3.3	2.8 to 25.6
Adult basic education and literacy program (Colombia)	19.9	8.1to 1,764
School-based reproductive health program to prevent HIV/AIDS (Honduras)	0.5	0.1 to 4.6
Iron supplementation for secondary school students (low-income country)	32.1	25.8 to 45.2

Source: Knowles and Behrman (2003a, Table 6-1)

Policy-makers who are considering starting new programs for their countries clearly can learn from what appear to be "best-practices" elsewhere, as reflected in high benefit-cost ratios. However, as emphasized above, conditions vary so that what are good policies in one context are not necessarily good policies in another context. Therefore in thinking about introducing new policies in a new context, it is desirable to conduct benefit-cost analysis for the new context, using appropriate prices for evaluating the resource costs, eliminating transfer components that do not have resource costs, using appropriate prices to evaluate the benefits, and discounting carefully the timing of both costs and benefits. Of course there is likely to be considerable uncertainty for most of the components of such estimates—and this should be recognized. However with tools no more complicated than EXCEL or other standard spreadsheets, analysts easily can generate the implications of alternative assumptions regarding critical components of costs, benefits and discount rates. While undertaking such exercises often will result in substantial ranges of pre-project benefit-to-cost ratios, such as in the table in the previous paragraph, they also will be informative about what are the critical assumptions. This might usefully lead to attempting to improve the

[37] Three of the four examples that are given have plausible ranges of benefit-to-cost ratios that are above one, so marginal increases in these programs would seem likely to be warranted by this criterion. But it certainly should not be deduced from this fact that virtually all programs have benefit-cost ratios greater than one. Better and better-monitored programs are more likely to collect information that permits benefit-cost analysis—and most really lousy programs, on the other hand, are not likely to collect and make available information that permits serious evaluation. But the returns would be high from reducing or eliminating bad programs as well as from expanding good programs. The social rates of return almost for sure would be higher from better evaluation of all types of programs and less concentration on programs that are thought to be good.

current information about the critical factors. If the decision is made to implement the program, moreover, such analysis should help clarify what pre-program baseline data and what post-program-initiation data should be given high priority in order to be able to monitor and evaluate the program after it is implemented.

APPENDIX A

ECONOMETRIC ESTIMATES OF IMPACTS OF EDUCATIONAL INVESTMENTS

ECONOMETRIC ESTIMATES OF IMPACTS OF EDUCATIONAL INVESTMENTS

Econometric or statistical methods are used to attempt to circumvent some of the limitations of the data, including that most data that are available for evaluation of the impact of investments in education are behavioral data and not experimental data. This appendix briefly discusses some of these problems and possible resolutions.

RELATIONS TO BE ESTIMATED

Econometric analyses of impacts should be based on relations such as those implied in the discussion in Chapter One. Such relations can be used (i) to estimate directly the underlying structural relations that determine investments in education or their impact (e.g., human resource production functions) in which investments in education may be one input in the production, for example, of subsequent adult health and (ii) to estimate dynamic decision rules or demand relations for investments in education or for inputs conditional on past investments (as in relation 1).

Structural Relations—Production Functions: Structural relations are the basic underlying relations in the models of behaviors such as in Chapter One. Structural relations are not estimated nearly as often as are the reduced-form dynamic decision rules or demand relations that are discussed below. But one type of structural relations—production functions—are estimated with some frequency. A linear or log-linear approximation to a general production function with cognitive achievement (CA_i) produced by two categories (vectors) of variables relating to the i^{th} individual and his/her household (XI) and to the s^{th} school (XS) and by an explicit stochastic disturbance term (Ui) is:[38]

$$(2) \qquad CA_i = a_{XI}XI + a_{XS}XS + U_i.$$

The household vector of variables may include, for example, parent's schooling and the home learning environment, as well as the schooling to date of the ith individual. The school vector of variables includes aspects of school management and curricula that may be affected directly by policies related to the supply side of schooling. The stochastic term captures random effects that are not correlated with any of the other predetermined right-side variables. It is useful for the discussion below of estimation issues to distinguish among four

[38] Linear approximations are used here because they are the simplest forms but they still permit characterization of various estimation issues. Log-linear forms in which all of the variables are replaced by the logarithms of their values (which implies interactions among all the right-side variables) are identical in representation once the variables are redefined. In empirical studies linear and log-linear specifications are very common, but other functional forms also are used at times. For other functional forms the essence of the estimation issues is the same. If the functional form that is used is not a good approximation to the true functional form, there is misspecification error that is akin to omitted variable bias discussed below (with the unobserved variable being the variable that would have to be added to transform the assumed specification to the true functional form).

different subgroups of variables: the superscripts "o" and "u" refer to "observed in the data used" and "unobserved in the data used", the superscript "b" refers to variables that are behaviorally determined within the model used, and the superscript "p" refers to variables that are predetermined within the model used so that the variable list in the general production function relation is XI^{ob}, XI^{ub}, XI^{op}, XI^{up}, XS^{ob}, XS^{ub}, XS^{op}, XS^{up}, U_i. If these were substituted into (2) each would have its own coefficient "a" with an appropriate superscript to indicate its impact on CA_i. The distinctions among these different variable groups are important because some of the most substantial and most pervasive estimation problems arise from unobserved variables or behaviorally determined variables (see below).

The parameters (a's) in the production function give the direct impact of the right-side variables, including investments in education. With good estimates of the appropriate production functions the direct determinants of many outcomes determined by behaviors could be evaluated with considerable confidence to answer many of the questions of interest about direct policy impacts. If good estimates of production functions are embodied in overall models of optimizing behavior, moreover, simulations can be made of the impact of policies given all the relevant behavioral adjustments and of counterfactual policies changes, all conditional on the model. Note that the possibilities of simulating the impact of counterfactual policies exist with the use of structural models even though they are not possible with experiments and may be very difficult to undertake with reduced-form dynamic decision rules. Good production function estimates may be difficult to obtain, however, because of estimation problems discussed below.

Reduced-Form Dynamic Decision Rule or "Demand" Relations: A second set of relations that can be estimated to explore the determinants of investments in education and the impact of investments in education are dynamic decision rules or "demand" relations that are conditional on past investments. These relations give some behavioral outcome in the current period as dependent on all predetermined (from the point of view of the entity making the decisions) prices and resources and on the parameters in the underlying production functions and preferences. As noted above, these are the relations that are most commonly estimated. These demand functions in principle are derived explicitly from the constrained maximization behavior of families that is discussed in Chapter One. As such they incorporate all of the underlying structural parameters that are involved in that process. But all of the choice variables during the period of interest are substituted out, so the demand functions are so-called reduced-form relations because the maximizing behavior that determines such variables has been combined and "reduced" to the relations that give the behavioral outcomes as a function of purely predetermined and expected prices, resources, policies and of the underlying preferences and technologies. In some empirical studies, the underlying structural parameters can be identified from estimation of the demand relations. In most cases, however, demand functions are just posited to result from constrained maximization and the underlying structural parameters are not identified in the estimates, though the demand parameters still are some combinations of these parameters. In such cases, demand functions permit the estimation of the total effects of predetermined variables on the behavioral variables of concern, but not estimation of the exact mechanisms through which determinants act.

On a general level, demand functions can be written with a vector of behavioral outcomes (Z) dependent on a vector of prices broadly-defined (P) and a vector of resources (R). Both the prices and the resources may reflect policies. If there are uncertainties regarding relevant future prices, policies and shocks, then the characteristics known at the time of the decision of interest regarding the distributions of those outcomes should be included. A linear approximation for a family or individual facing prices PF and with resources RF_f and a vector of stochastic terms (V_f) is:

(1A) $Z_f = b_{PF}PF + b_{RF}RF_f + V_f$.

The resources include predetermined schooling and other human resources. The stochastic term in each relation includes all the effects of all the stochastic terms in all of the production activities in which the family/individual is engaged (i.e., all of the elements of the vector U_i), plus perhaps other chance events. Both prices and resources may be observed or unobserved in the data, so it is useful to indicate that distinction here as above in the discussion of production function inputs (again, using superscripts "o" and "u"). There is one such demand relation (or one element in the vector Z_f) for every behavioral outcome of the individual/family (and similarly for firms or other entities). Each of these demand relations conceptually includes the same identical right-side predetermined variables so that any predetermined variable that affects any one behavioral outcome may affect all other behavioral outcomes. Good estimates of these relationships, with predetermined investments among the right-side variables, would inform us about the impacts of these investments – an essential component for evaluating the rates of return to these investments.

ESTIMATION PROBLEMS

There are a number of possible problems in obtaining good estimates of the impact of investments in education on different outcomes such as those considered in Chapter Three. Therefore what are presented as estimates of relations such as those that are discussed in Chapter Two often may be biased. These estimation problems share a common characteristic: the disturbance term in the relation actually estimated is <u>not</u> simply an element in U_i or V_f that is distributed independently of all the right-side variables in the relation being estimated, but instead is correlated with right-side variables (e.g., because it is a compound disturbance term that includes unobserved variables as well as U_i or V_f or because of the way that U_i or V_f is defined for the sample used in the estimates).

Measurement error: Measurement error may contaminate any of the observed variables used for estimates of the relations (1A) and (2). Random measurement error occurs if what is observed is not the true variable, but the true variable plus a random error.[39] As is well known, random measurement error in a right-side variable tends to cause bias in the coefficient estimate of that variable towards zero. Intuitively, if the observed investments in education are a noisy measure of the true value of investments in education, the true

[39] Random measurement error is what usually is emphasized, and is what we discuss here. Measurement error also may be systematically related to the true variable, with implications that depend on the exact nature of the systematic relation.

dependence of outcomes such as cognitive achievement on investments in education is masked, and the result is an underestimate of the effect of the investment on the outcome. The bias is greater the larger is the variance in the measurement error relative to the variance in the true value. Random measurement error can be reduced with better measurements of the desired concept. Random measurement error can be controlled with instrumental variable estimates if the error in the instrument used is independent of the error in the observed variables of interest, as may occur, for example, with different measures of the same concept.[40]

Omitted variables: In both production function estimates and demand function estimates there may be variables that should be included among the right-side variables but that are not observed and therefore not included. For the production function estimates, for example, there may be unobserved inputs such as inherent ability, motivation, and school management capabilities. In terms of relation (2A) with the subcategories of variables, the basic estimation problem is that the observed right-side variables (XI^{ob}, XI^{op}, XF^{ob}, XF^{op}) may be correlated with the unobserved variables (XI^{ub}, XI^{up}, XF^{ub}, XF^{up}) that are included in the compound disturbance term with U_I.[41] Therefore the estimates of the impact of the observed variables include not only their true effects but also part of the effects of any correlated unobserved variables. For the demand relations (1A), the compound disturbance term includes, in addition to V_f, the other unobserved variables (PF^u, RF^u). If any of the observed variables on the right-side of relation (1A) is correlated with any of the unobserved variables, its coefficient estimate is biased because, in addition to its own effects, it is representing in part the effect(s) of the correlated unobserved variable(s). If, say, ability affects the outcome of interest and investments in education are correlated with ability because individuals with greater ability tend to obtain more schooling, then the usual estimate of the impact of investments in education on the outcome is likely to be biased (and likely to make investments in education appear to have a larger effect than they really do) because in the usual estimates investments in education in part are representing the effects of unobserved ability, not only of investments in education per se.[42]

As is well known, the sign and magnitude of omitted variable bias depends on the effect of the omitted variable(s) and on its correlation with included variables and their true coefficients. Five means of dealing with omitted variable bias are (i) to measure variables that often are unmeasured (e.g., "ability" could be measured using Raven's tests, as in Knight and Sabot 1990), (ii) to use fixed effects to control for unobserved variables (which requires multiple observations at the level of aggregation at which the fixed effects are used) as in Ashenfelter and Krueger 1994, Behrman and Rosenzweig (1999, 2002), Behrman, Rosenzweig and Taubman (1994, 1996), Behrman, et al. (1999), Foster and Rosenzweig

[40] For example, schooling reports from adult siblings or from adult children could be used for this purpose if schooling is a right-side variable as in some recent estimates of the impact of schooling on wages (Ashenfelter and Krueger 1994, Behrman, Rosenzweig and Taubman 1994).

[41] The compound disturbance term includes all the unobserved variables unless their effects are controlled in some way.

[42] While we indicate the likely direction of biases in the text, the actual biases depend on the exact difference between the correct specification and the specification used and the covariances among all the variables in both specifications, so the actual biases may not be in the directions that we posit are likely on the basis of first-order effects.

(1995, 1996), Miller, Mulvey and Martin (1995, 1997), Rosenzweig and Wolpin (1986).[43] (iii) to replace right-side variables with their instrumented values by using identifying instruments that do not appear in the relation being estimated and are not correlated with the disturbance term in the relation being estimated as attempted by Alderman, et al. (1996b, 2001), Angrist (1990), Angrist and Lavy (1999), Card (1995, 1999), (iv) to compare the behaviors or the "before and after" changes in behaviors of beneficiaries of policies with those of individuals who would be eligible but are in the control group (perhaps with matching on observed characteristics as in Heckman, Ichimura and Todd 1997), and (v) to use experimental data as in Angrist, et al. (2002), Behrman, Sengupta and Todd (2002), Glewwe, Ilias and Kremer (2002), Kremer (2003), Kremer *et al.* (1996), Miguel and Kremer (2001), Skoufias (2001) and Schultz (2000, 2003). All of these approaches have their limitations. Some variables are very hard to measure at reasonable costs. Fixed effects do not permit the estimation of the linear effects of observed variables at the same level of aggregation as the fixed effects (though the effects of interactions among observed variables, such as family background and program characteristics, can be estimated), exacerbate the impact of measurement errors, do not control for unobserved variables, and control for the unobserved fixed variables perfectly only if the true relationship being estimated can be manipulated so that the unobserved fixed effect appears only as an additive linear term . It often is difficult to find identifying instruments that (a) do not appear in the relation being estimated, (b) are independent of the compound disturbance term in the relation being estimated (which includes all of the unobserved variables), and (c) are sufficiently correlated with the observed right-side behavioral variables (though lagged price and other shocks are candidates for panel data). The control group comparisons depend on good assignments of households to the actual and potential beneficiary groups versus those who are not eligible; if there are incorrect assignments, misleading comparisons may be made (though matching may reduce this problem considerably). Matching controls only for observed variables but not unobserved ones. Experiments often are costly, hard to maintain (i.e., keeping control and treatment groups separate), and in some cases not politically possible or ethical. (see Chapter One).

Endogeneity: Endogeneity bias occurs when a variable that is determined within the model appears as a right-side variable in some other relation. Among the relations discussed above, production functions are the ones for which endogeneity most obviously might be a problem because the right-side variables include some behavioral inputs (e.g., health and nutritional status in relation 2). Included on the right side of that relation is a stochastic term (V_f) that, as noted in Chapter One, includes the stochastic terms from all of the production function relations in the model—including those for health and nutritional status. This results in a correlation between health and nutritional status and the stochastic term in the cognitive achievement production function that causes biases in the estimated impact of health and nutritional status on cognitive achievement production. The sign and the magnitude of the bias depend upon the exact structure of the model. If there are no unobserved behavioral inputs in the production function, prices and any other variables that enter into the reduced-

[43] The studies that use identical twins to control for unobserved endowments (including genetic endowments and common family background) generally find upward biases of from 12 to 100 percent in the estimated impact of schooling on wages (see the summary in Behrman and Rosenzweig 1999).

form demand relations in (1A), but not directly in the production function, can serve as identifying instruments for controlling for endogeneity.

Selectivity: Selectivity bias may result if observations are available only for a selected subset of the sample. A relevant example is for estimating the impact of school characteristics on cognitive achievement tests given at the secondary school level. Such test scores are not observed for everyone in most samples from most developing countries because not everyone attends secondary school. These test scores are only observed for individuals whose expected gains from attending secondary school exceed the cost of attending secondary school, which is likely to differ by socioeconomic class because those who are better off are more likely to be able to self-finance such investments in the presence of imperfect or absent capital markets for investments in education. The problem is that this subsample is not randomly selected. The subsample selection procedure, with its systematic relation between the disturbance term in the true relation and test scores, creates a correlation between the disturbance term and test scores for the subsample for which estimates of the relation can be made. As a result, if the relation is estimated using only this subsample, a biased estimate of the true relation between test scores and school characteristics is obtained. The standard means of controlling for selectivity is to incorporate the behaviors that cause selectivity explicitly into the model, though in some cases finding observed variables that determine the selectivity but that do not enter into the relation of interest (which is necessary to identify the coefficients of interest) may be difficult.

APPENDIX B

PATTERNS IN RECENT DECADES BETWEEN SCHOOLING ENROLLMENTS AND REGION, TIME, INCOME AND ADULT WOMEN'S SCHOOLING

PATTERNS IN RECENT DECADES BETWEEN SCHOOLING ENROLLMENTS AND REGION, TIME, INCOME AND ADULT WOMEN'S SCHOOLING

This appendix provides additional information that underlies the summary in Chapter Two, with attention to three measures of schooling and the gender gaps in each; primary schooling enrollment rates, secondary schooling enrollment rates and expected schooling for a synthetic cohort.[44] These descriptions are based on aggregate data for the 1960s through the 1990s from the World Bank (2001) and Barro and Lee (from the World Bank website).[45] The tables at the end of the appendix give the regression estimates that underlie these verbal summaries —with estimates based on the categories discussed in this appendix (e.g., region, decade for time, income-country group for income, low or middle or high adult female schooling) and, in addition, continuous measures for time, income and adult female schooling.

REGION (TABLE B.1)

Table B.1 Schooling by Region

Regions	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected Years of Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
East Asia & Pacific	-12.427	-3.740	7.514	5.140	-43.530	-16.000	4.270	4.590	-5.170	-12.700	0.760	5.030
Europe & C. Asia	-5.571	-1.720	1.307	0.920	-3.235	-1.190	2.216	2.390	-0.623	-1.540	0.306	2.040
Latin America & Caribbean	-5.361	-2.010	1.569	1.340	-37.036	-16.340	-2.674	-3.450	-4.278	-12.920	-0.174	-1.410
Middle East & N. Africa	-31.078	-9.970	20.585	15.010	-44.132	-16.950	11.550	12.960	-6.298	-16.400	2.244	15.740
South Asia	-56.474	-13.610	24.863	13.610	-58.677	-16.300	10.486	8.510	-8.947	-17.000	2.323	11.890
Sub-Saharan Africa	-46.298	-19.760	15.953	15.480	-65.765	-33.400	5.262	7.810	-9.085	-31.390	1.367	12.720
Constant	102.292	58.060	1.754	2.260	76.877	52.210	1.169	2.320	13.310	61.630	0.184	2.290
	N	1063	N	1063	N	1054	N	1054	N	1010	N	1010
	F(6, 1056	102.9	F(6, 1056	91.77	F(6, 104	232.79	F(6, 1047	55.84	F(6, 1003	206.47	F(6, 1003	84.94
	Adj R^2	0.3654	Adj R^2	0.339	Adj R^2	0.5691	Adj R^2	0.2381	Adj R^2	0.5499	Adj R^2	0.333
	Root MSE	26.834	Root MSE	11.808	Root MSE	22.526	Root MSE	7.7099	Root MSE	3.2539	Root MSE	1.208

Region 9 = Developed Countries from All Regions is the omitted group.

[44] Expected schooling for a synthetic cohort is calculated based on enrollment rates at a point of time. If primary school is six grades, secondary school is six grades and tertiary school is four years, for example, the expected schooling for a synthetic cohort with enrollment rates of P, S and T for the three schooling levels is 6*P + 6*S + 4*T. These measures have the advantage of summarizing the enrollment rates. But they incorporate at least two additional sources of error beyond those in the enrollment rates. First, the number of grades at schooling levels varies across countries and over time within countries, but data on these variations are not readily available, so we assume the same number of grades in each level for the whole sample. Second, the World Bank (2001) does not give tertiary enrollment rates by sex, but only the total tertiary enrollment rate. Therefore we assume that the share of females in tertiary schooling is the same fraction of the share of females in secondary schooling as is the share of females in secondary schooling relative to the share of females in primary schooling. That is, if females are 40 percent of primary students and 20 percent of secondary students, we assume they are 10 percent of tertiary students (= 20 percent*20 percent/40 percent). For many developing countries for most or all of the relevant time period, the total tertiary enrollment rate is low enough that the latter assumption probably does not introduce great error, but if, for example, gender differentials in tertiary schooling enrollments have tended to decline more rapidly than those in primary and secondary schooling over time, this procedure probably understates somewhat the extent to which the gender gap in expected schooling has declined.

[45] This section draws on and summarizes much more extensive research in Behrman and Sengupta (2002), to which the interested reader is referred for further details.

Female gross primary school enrollment rates have been significantly lower in each of the six developing country regions in comparison with developed countries. The difference in gross female primary enrollment rates between developed and developing countries, however, has varied considerably among developing regions. It has averaged the largest for South Asia (-56 percentage points), followed by sub-Saharan Africa (-46 percentage points), and the Middle East and North Africa (-31 percentage points). It has been much smaller for East Asia and the Pacific (-12 percentage points), Europe and Central Asia (-6 percentage points) and Latin America and the Caribbean (-5 percentage points).

The male-female gaps in gross primary enrollment rates have been significantly greater in four of the developing regions than in the developed countries. The gap has been largest in South Asia (25 percentage points larger than in developed countries), but also fairly large in the Middle East and North Africa (21 percentage points) and sub-Saharan Africa (16 percentage points) and somewhat smaller though still significant in East Asia and the Pacific (8 percentage points). Gender differences in gross primary enrollment rates were not significantly different in Europe and Central Asia or in Latin America and the Caribbean from those for developed countries.

Female gross secondary school enrollment rates have been significantly and, in most case, substantially below the rates for the developed economies – and substantially more so than for female gross primary school enrollments. As in the case of gross primary enrollment, the two regions with female gross secondary enrollments most below the developed countries have been sub-Saharan Africa (-66 percentage points below the developed countries) and South Asia (-59 percentage points), though at the secondary level, sub-Saharan Africa is at the bottom of the list instead of South Asia. Surprisingly countries in East Asia and Pacific (-44 percentage points) and Latin America and the Caribbean (-37 percentage points) have been approximately as far below the developed countries as countries in the Middle East and North Africa (-44 percentage points), a result that is striking when compared to the gap between the first two of these regions and the third in female primary gross enrollment rates. This suggests that the developing countries in East Asia and Pacific and Latin America and the Caribbean have been relatively successful in universalizing primary education, but have a large drop off in the transition from primary to secondary school. Developing countries in Europe and Central Asia have had the highest secondary level enrollment rates among the six developing country regions, not significantly below the developing countries. Further, this is the only one of the developing country regions for which the difference in enrollment rates in comparison with developed countries is smaller for female gross secondary enrollment rates than for female gross primary enrollment rates.

For five of the developing country regions the estimated gender gaps for gross secondary schooling enrollment rates are less than those for gross primary schooling rates. One of these five cases merits special note, however, because it is the only case for either primary or secondary gross enrollment rates in which the gender gap significantly favors females more than in developed countries – Latin America and the Caribbean. Only for three regions are there larger differences than 5 percentage points between the gender gaps for secondary versus primary gross enrollments: for Middle East and North Africa, South Asia and sub-Saharan Africa. For each of these regions, interestingly, the gender gaps in gross secondary

school enrollment rates (12, 10 and 5 percentage points) are substantially smaller than those in gross primary school enrollment rates (21, 25 and 16 percentage points). The lessening of gender differences for secondary as compared with primary enrollments in most of these regions may reflect that as boys age the opportunity costs of them attending school rather than working increase more than for girls.

The enrollment differences discussed above in developing countries in comparison with developed economies imply substantial differences in expected years of female schooling on the average for most of the developing world in comparison with the developed countries: Sub-Saharan Africa and South Asia have averaged about than nine years less, the Middle East and North Africa over six years less, East Asia and the Pacific over five years less and Latin America and the Caribbean over four years less (Europe and Central Asia is not significantly different). The gender gap in expected schooling favoring males has been largest in South Asia (2.3 years), the Middle East and North Africa (2.2 years) and sub-Saharan Africa (1.4 years), much smaller but still statistically significant in East Asia and the Pacific (0.8 years) and Europe and Central Asia (0.3 years) – and insignificant in Latin America and the Caribbean.

TIME (TABLE B.2)

Table B.2: Schooling Over Time

Time	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected years of Female Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Year	1.344	4.030	-0.176	-1.210	1.179	3.650	0.143	1.570	0.153	3.250	0.012	0.750
$(Year)^2$	-0.014	-1.600	-0.003	-0.750	0.005	0.630	-0.008	-3.240	0.001	0.630	-0.001	-2.860
Constant	62.355	23.650	15.276	13.210	16.169	6.220	5.469	7.460	5.254	14.050	1.302	10.690
	N 1063		N 1063		N 1054		N 1054		N 1010		N 1010	
	F(2, 1060) 46.91		F(2, 1060) 27.61		F(2, 1051) 132.71		F(2, 1051) 22.82		F(2, 1006) 109.86		F(2, 1006) 33.83	
	Adj R^2 0.0796		Adj R^2 0.0477		Adj R^2 0.2001		Adj R^2 0.0398		Adj R^2 0.1775		Adj R^2 0.0611	
	Root MSE 32.316		Root MSE 14.173		Root MSE 30.691		Root MSE 8.6551		Root MSE 4.3987		Root MSE 1.4332	
Decade 2 (1970, 1975)*	9.120	3.200	-1.748	-1.400	10.816	3.960	0.241	0.310	1.448	3.650	-0.068	-0.530
Decade 3 (1980, 1985)	20.948	7.520	-5.450	-4.460	28.579	10.670	-1.170	-1.550	3.822	9.790	-0.387	-3.050
Decade 4 (1990, 95, 99)	23.534	8.380	-8.318	-6.760	39.555	14.440	-4.159	-5.400	5.201	13.120	-0.916	-7.100
Constant	67.164	33.650	14.563	16.660	20.671	10.560	5.769	10.470	5.821	20.740	1.318	14.450
	N 1063		N 1063		N 1054		N 1054		N 1010		N 1010	
	F(3, 1059) 30.3		F(3, 1059) 18.39		F(3, 1050) 84.56		F(3, 1050) 14.01		F(3, 1006) 69.7		F(3, 1006) 21	
	Adj R^2 0.0764		Adj R^2 0.0468		Adj R^2 0.1923		Adj R^2 0.0357		Adj R^2 0.1696		Adj R^2 0.0561	
	Root MSE 32.371		Root MSE 14.18		Root MSE 30.84		Root MSE 8.6734		Root MSE 4.4197		Root MSE 1.4369	

Female gross primary schooling enrollment rates have increased significantly over time, but at diminishing rates, at least after the 1980s. The female gross enrollment rates in the 1970s averaged 9 percentage points above those in the 1960s, those in the 1980s averaged 21 percentage points above those in the 1960s, and those in the 1990s averaged 24 percentage points above those in the 1960s. This implied an increase from an average female gross

primary school enrollment rate of 67 percent in the 1960s to 91 percent in the 1990s (and 93 percent in 2000).

Male-female gross primary schooling enrollment gaps significantly decreased over time. The male-female enrollment gap in gross primary enrollment rates in the 1970s averaged –2 percentage points below those in the 1960s, those in the 1980s averaged –5 percentage points below those in the 1960s, and those in the 1990s averaged –8 percentage points below those in the 1960s. This implied a decrease from an average male-female gross primary school enrollment rate gap of 15 percentage points in the 1960s to 7 percentage points in the 1990s. That is, this gender gap on average was more than cut in half over this period.

Female gross secondary schooling enrollment rates increased significantly over time at close to a linear rate of about 1.2 percentage points per year. The female gross secondary enrollment rates in the 1970s averaged 11 percentage points above those in the 1960s, those in the 1980s averaged 29 percentage points above those in the 1960s, and those in the 1990s averaged 40 percentage points above those in the 1960s. This implied an increase from an average female gross secondary school enrollment rate of 21 percent in the 1960s to 61 percent in the 1990s – almost a tripling. Though female gross secondary enrollment rates on average started far below primary enrollment rates in the 1960s (21 versus 67 percent on average) and remained substantially below primary enrollment rates in the 1990s (61 versus 91 percent on average), the stronger association between secondary than primary enrollment rates over time meant that the female gross primary-secondary enrollment rate gap fell considerably (from 56 percent on average in the 1960s to 30 percent on average in the 1990s).

The male-female gross secondary schooling enrollment rate gaps decreased significantly over time, at somewhat increasing rates. The male-female enrollment gap in gross secondary enrollment rates in the 1970s was not significantly different than that in the 1960s, but that in the 1980s averaged –1 percentage point below those in the 1960s, and those in the 1990s averaged –4 percentage points below those in the 1960s. This implied a decrease from an average male-female gross secondary school enrollment rate gap of 6 percentage points in the 1960s to 2 percentage points in the 1990s. These reductions were not as large as those for the primary school gross enrollment gender gap, but because the gap was much lower at the secondary level throughout the four decades, even with the smaller absolute declines the secondary gender gap on average was cut almost to zero over this period.

Female expected schooling for a synthetic cohort increased significantly over time at close to a linear rate of about a sixth of a grade per year. Female expected schooling in the 1970s averaged 1.4 years above that in the 1960s, that in the 1980s averaged 3.8 years above that in the 1960s, and that in the 1990s averaged 5.2 years above that in the 1960s. This implied an increase from an average female expected schooling of 5.8 years in the 1960s to 11.0 years in the 1990s. These are considerable gains—a 90 percent increase.

The male-female expected schooling gap declined significantly over time, probably at somewhat increasing rates. The male-female expected schooling gap in the 1970s was not significantly different than that in the 1960s, but that in the 1980s averaged –0.4 years below

that in the 1960s, and that in the 1990s averaged –0.9 years below that in the 1960s. This implied a decrease from an average male-female expected schooling gap of 1.3 grades in the 1960s to 0.4 grades in the 1990s.

PER CAPITA INCOME (TABLE B.3)

Table B.3 Schooling by Income

Income	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected Years of Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Low Income Countries*	-48.585	-21.770	17.046	15.880	-61.012	-28.500	5.968	8.400	-8.804	-29.400	1.511	13.140
Lower Middle Income Countries	-13.790	-5.800	8.863	7.750	-38.977	-17.190	4.009	5.330	-4.921	-15.520	0.869	7.130
Upper Middle Income Countries	-4.711	-1.770	3.191	2.490	-26.770	-10.370	1.010	1.180	-3.058	-8.540	0.232	1.680
Constant	102.292	58.320	1.754	2.080	76.877	45.910	1.169	2.100	13.310	57.110	0.184	2.050
	N 1063		N 1063		N 1054		N 1054		N 1010		N 1010	
	F(3, 1059) 209.69		F(3, 1059) 98.98		F(3, 1050) 279.99		F(3, 1050) 28.6		F(3, 1006) 306.38		F(3, 1006) 69.07	
	Adj R^2 0.3709		Adj R^2 0.2168		Adj R2 0.4428		Adj R2 0.0729		Adj R2 0.4759		Adj R2 0.1683	
	Root MSE 26.717		Root MSE 12.854		Root MSE 25.614		Root MSE 8.5046		Root MSE 3.5113		Root MSE 1.3489	
Per Capita GNI (PPP)	0.006	11.790	-0.001	-10.110	0.004	25.040	-0.001	-8.710	0.001	22.590	0.000	-11.340
[Per Capita GNI (PPP)]^2	0.000	-9.420	(dropped)		(dropped)		(dropped)		(dropped)		0.000	8.350
Constant	70.990	40.740	12.839	18.730	29.225	21.560	6.199	12.710	7.060	35.890	1.763	18.170
	N 575		N 575		N 569		N 569		N 548		N 548	
	F(2, 572) 84.18		F(2, 573) 102.25		F(1, 567) 627.23		F(1, 567) 75.86		F(1, 546) 510.33		F(2, 545) 91.23	
	Adj R^2 0.2247		Adj R^2 0.1499		Adj R-squared 0.5244		Adj R^2 0.1165		Adj R^2 0.4822		Adj R^2 0.2481	
	Root MSE 23.956		Root MSE 12.183		Root MSE 24.094		Root MSE 8.6708		Root MSE 3.4099		Root MSE 1.305	
Per Capita GNI (Atlas Method)	0.004	9.820	-0.002	-9.000	0.008	23.860	-0.001	-7.530	0.001	20.330	0.000	-9.280
[Per Capita GNI (Atlas)]^2	0.000	-7.440	0.000	6.210	0.000	-14.650	0.000	5.000	0.000	-12.690	0.000	6.330
Constant	75.131	55.170	14.104	22.400	24.154	22.960	6.498	16.070	6.286	38.670	1.366	20.440
	N 750		N 750		N 744		N 744		N 721		N 721	
	F(2, 747) 58.26		F(2, 747) 56.93		F(2, 741) 461.49		F(2, 741) 40.61		F(2, 718) 340.13		F(2, 718) 61.14	
	Adj R^2 0.1326		Adj R^2 0.1299		Adj R^2 0.5535		Adj R^2 0.0963		Adj R^2 0.4851		Adj R^2 0.1431	
	Root MSE 29.185		Root MSE 13.494		Root MSE 22.697		Root MSE 8.7229		Root MSE 3.4225		Root MSE 1.4068	
Per Capita GDP (PPP)	0.006	11.630	-0.003	-11.330	0.010	24.360	-0.002	0.001	21.190	39.890	0.000	-10.970
[Per Capita GDP (PPP)]^2	0.000	-9.260	0.000	8.440	0.000	-15.070	0.000	0.000	-13.380	-24.620	0.000	8.040
Constant	70.005	38.960	17.773	20.580	14.492	10.170	8.957	5.014	23.040	43.360	1.795	17.910
	N 556		N 556		N 549		N 549		N 529		N 529	
	F(2, 553) 83.87		F(2, 553) 91.46		F(2, 546) 590.77		F(2, 546) 64.8		F(2, 526) 452.16		F(2, 526) 88.6	
	Adj R^2 0.23		Adj R^2 0.2458		Adj R^2 0.6828		Adj R^2 0.1889		Adj R^2 0.6308		Adj R^2 0.2491	
	Root MSE 24.163		Root MSE 11.614		Root MSE 19.528		Root MSE 8.4138		Root MSE 2.8667		Root MSE 1.32	
Per Capita GDP (Constant 1995 $)	0.004	13.020	-0.002	-11.400	0.006	22.330	-0.001	-8.410	0.001	20.840	0.000	-11.200
[Per Capita GDP (Const 95 $)]^2	0.000	-9.910	0.000	7.960	0.000	-13.230	0.000	6.380	0.000	-12.830	0.000	8.380
Constant	72.009	56.480	14.960	25.250	21.939	20.180	6.431	16.900	6.003	37.580	1.397	22.460
	N 872		N 872		N 856		N 856		N 833		N 833	
	F(2, 869) 104.42		F(2, 869) 92.86		F(2, 853) 467.78		F(2, 853) 43.5		F(2, 830) 386.67		F(2, 830) 79.35	
	Adj R^2 0.1919		Adj R^2 0.1742		Adj R^2 0.522		Adj R^2 0.0904		Adj R^2 0.4811		Adj R^2 0.1585	
	Root MSE 27.903		Root MSE 12.968		Root MSE 23.65		Root MSE 8.2801		Root MSE 3.4128		Root MSE 1.3294	

***The category 'Higher Income Countries (OECD and non-OECD)' has been used as the omitted group**

Female gross primary schooling enrollment rates significantly increase with higher incomes. The female gross enrollment rates for low-income countries have averaged -49 percentage points below those in the high-income countries, those in the lower middle-income countries have averaged -14 percentage points below and those in the upper middle-income countries have averaged -5 percentage points below. This implies an average female gross primary school enrollment rate of 53 percent in the low-income countries and of 88 percent in the lower-middle-income countries as compared with 97 percent and 102 percent in the upper-middle-income and high-income countries.

The male-female gross primary schooling enrollment rate gap decreases significantly as income increases, at somewhat decreasing rates. The male-female enrollment gap in gross primary enrollment rates in the low-income countries averaged 17 percentage points above the high-income countries, while those in the lower-middle-income countries and in the upper-middle-income countries, respectively, averaged 9 and 3 percentage points above those in high-income countries. This pattern of estimates implies a decrease from an average male-female gross primary school enrollment rate gap of 19 percentage points in low-income countries to 11 percentage points in lower-middle-income countries to 5 percentage points in upper-middle-income countries to 2 percentage points in high income countries. Thus, this gender gap is fairly strongly inversely associated with income.

Female gross secondary schooling enrollment rates increase significantly with income. The female gross enrollment rates in low-income countries averaged -61 percentage points below those in high-income countries, those in the lower-middle-income countries averaged -39 percentage points below, and those in the upper-middle-income countries averaged -27 percentage points below. This implies an increase from an average female gross secondary school enrollment rate of 16 percent in the low-income countries to 38 percent in the lower-middle-income countries to 50 percent in the upper-middle-income countries to 77 percent in the high-income countries. For female gross secondary enrollment rates, in contrast to female gross primary enrollment rates, there is a significant and substantial difference between the upper-middle-income country group and the high-income country group

Male-female gross secondary schooling enrollment rate gaps decrease significantly as income increases. The male-female enrollment gap in gross secondary enrollment rates in low-income countries averaged 6 percentage points above and the gap in lower-middle-income countries averaged 4 percentage points above those in high-middle-income countries and high-income countries. This implies a decrease from an average male-female gross secondary school enrollment rate gap of 7 percentage points in low-income countries to 5 percentage points in low-middle-income countries to 1 percentage point in high-middle-income countries and in high-income countries. The reductions with movement from one to another income group are not as large as those for the primary school gross enrollment gender gap in important part because the gap was much lower at the secondary level for the low-income group.

Female expected schooling for a synthetic cohort significantly increases as income increases, though at a diminishing rate. The expected years of female schooling averaged –8.8 grades less for the low-income group than for the high-income group, -4.9 grades less for the lower-

middle-income group and –3.1 grades less for the upper-middle-income group than for the high-income group. These are considerable differences across income groups.

The male-female expected schooling gap decreases significantly as income increases, probably at somewhat increasing rates. The estimated male-female expected schooling gap on average has been 1.5 years greater for low-income countries than in high-income countries, 0.9 years greater for lower-middle-income countries than in high-income countries and 0.2 years greater for upper-middle-income countries than in high-income countries. This implies a decrease from an average male-female expected schooling gap of 1.7 grades in low-income countries to 1.1 grades in lower-middle-income countries to 0.4 in upper-middle-income countries to 0.2 grades for high-income countries.

ADULT FEMALE SCHOOLING (TABLE B.4)

Table B.4 Schooling by Average Adult Female Schooling

Average Adult Female Schooling	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected years of Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Avg. Fem. Scl. 6 - 9 yrs	21.616	8.160	-12.285	-9.290	53.592	24.190	-6.940	-8.510	6.742	20.110	-1.323	-9.710
Avg. Fem. Scl. 9 yrs and more	21.782	4.950	-11.960	-5.440	64.649	17.640	-7.464	-5.530	8.195	15.020	-1.399	-6.310
Constant	79.307	67.560	12.671	21.620	28.648	29.100	6.392	17.640	7.084	47.890	1.259	20.940
	N 741		N 741		N 732		N 732		N 716		N 716	
	F(2, 738) 41.34		F(2, 738) 52.69		F(2, 729) 402.86		F(2, 729) 46.55		F(2, 713) 282.95		F(2, 713) 60.63	
	Adj R^2 0.0983		Adj R^2 0.1226		Adj R^2 0.5237		Adj R^2 0.1108		Adj R^2 0.4409		Adj R^2 0.143	
	Root MSE 27.803		Root MSE 13.883		Root MSE 23.151		Root MSE 8.5219		Root MSE 3.4434		Root MSE 1.3993	
Avg. Adult Fem. Scl.	21.997	31.060	-9.576	-23.210	12.629	18.860	-3.157	-9.750	2.269	25.770	-0.811	-17.340
(Avg. Adult Fem. Scl.)^2	-1.680	-23.090	0.694	16.390	-0.289	-4.220	0.177	5.340	-0.097	-10.810	0.053	11.180
Constant	39.809	29.400	30.038	38.090	1.075	0.840	12.563	20.340	2.419	14.320	2.783	31.010
	N 741		N 741		N 732		N 732		N 716		N 716	
	F(2, 738) 717.75		F(2, 738) 451.78		F(2, 729) 1248.24		F(2, 729) 135.99		F(2, 713) 1392.39		F(2, 713) 310.73	
	Adj R^2 0.6595		Adj R^2 0.5492		Adj R^2 0.7734		Adj R^2 0.2697		Adj R^2 0.7956		Adj R^2 0.4642	
	Root MSE 17.084		Root MSE 9.9511		Root MSE 15.969		Root MSE 7.723		Root MSE 2.0821		Root MSE 1.1064	

*** The category 'Avg. Adult Fem. Schooling<=6yrs' has been used as the omitted group**

The variations in female enrollment rates and expected schooling and in gender gaps in these variables are more consistent with variations in adult female schooling than with region, time or income with the single exception that variations in female secondary school enrollment rates are more consistent with regional variations. Thus there are quite strong intergenerational female schooling associations.

Female gross primary schooling enrollment rates increase significantly and substantially for higher adult female schooling levels (at diminishing rates). The female gross enrollment rates for medium- and high-adult-female-schooling countries averaged 22 percentage points above those in the low-adult-female-schooling countries (where the three country groups are defined by average adult female schooling below six years, between six and nine years, and above nine years). This implies an average female gross primary school enrollment rate of 79

percent in the low-adult-female-schooling countries and of 101 percent in the middle-adult-female-schooling countries and in the high-adult-female-schooling countries.

The male-female gross primary schooling enrollment rate gaps decrease significantly as adult female schooling increases, at somewhat decreasing rates, The male-female enrollment gap in gross primary enrollment rates in the medium-adult-female-schooling countries and those in the high-adult-female-schooling countries averaged -12 percentage points below those in low-adult-female-schooling countries. This pattern of estimates implies a decrease from an average male-female gross primary school enrollment rate gap of 13 percentage points in low-adult-female-schooling countries to 1 percentage point in middle- and high- adult-female-schooling countries. Thus, this gender gap is fairly strongly inversely associated with adult female schooling.

Female gross secondary schooling enrollment rates increase significantly with adult female schooling. The female gross enrollment rates in medium-adult-female-schooling countries averaged 54 percentage points above those in low-adult-female-schooling countries and those in high-adult-female-schooling countries averaged 65 percentage points above those in low-adult-female-schooling countries. This implied an increase from an average female gross secondary school enrollment rate of 29 percent in the low-adult-female-schooling countries to 83 percent in middle-adult-female-schooling countries to 94 percent in the high-adult-female-schooling countries. For female gross secondary enrollment rates, in comparison with female gross primary enrollment rates, the gap between the middle-adult-female-schooling country group and the high-adult-female-schooling country group of 11 percentage points is much larger. Thus increasing average female adult schooling beyond six years has little association with increases in female gross primary schooling enrollment rates (that are almost universal in such cases) but does have positive associations with female gross secondary enrollment rates.

The male-female gross secondary schooling enrollment rate gaps decrease significantly as adult female schooling increases. The male-female enrollment gap in gross secondary enrollment rates in middle- and high-adult-female-schooling countries averaged -7 percentage points below that in low-adult-female-schooling countries. This implies a decrease from an average male-female gross secondary school enrollment rate gap of 6 percentage points in low-adult-female-schooling countries to -1 percentage points in middle- and high-adult-female-schooling countries (i.e., favoring females over males). The reductions with movement from the low to higher adult-female-schooling group are not as large as those for the primary school gross enrollment gender gap.

Expected female schooling for a synthetic cohort increases significantly with adult female schooling (at a diminishing rate). The female expected schooling in medium-adult-female-schooling countries averaged 6.7 years above those in low-adult-female-schooling countries and those in high-adult-female-schooling countries averaged 8.2 years above those in low-adult-female-schooling countries.

The male-female expected years of schooling gap decreases significantly as adult female schooling increases. The male-female expected years of schooling gaps in middle- and high-

adult-female-schooling countries averaged respectively –1.3 and –1.4 grades below that in low-adult-female-schooling countries. This implies a decrease from an average male-female expected years of schooling gap of 1.3 grades in low-adult-female-schooling countries to basically zero in middle- adult-female-schooling countries and to –0.1 (i.e., favoring females over males) in high-adult-female-schooling countries.

MULTIVARIATE ASSOCIATIONS (TABLES B.5A AND B.5B)

Table B.5a Multivariate Relations with Categorical Variables

Independent Variables	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected Years of Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
East Asia & Pacific	25.419	6.020	-12.959	-5.540	-2.597	-0.750	-4.450	-2.580	-0.155	-0.300	-0.632	-2.360
Europe & C. Asia	25.463	4.600	-9.485	-3.090	(dropped)	(dropped)	(dropped)	(dropped)	(dropped)	(dropped)	(dropped)	(dropped)
Latin America & Caribbean	25.262	6.430	-15.863	-7.280	1.112	0.350	-11.307	-7.150	0.236	0.500	-1.448	-5.920
Middle East & N. Africa	9.346	2.090	2.427	0.980	-0.206	-0.060	3.610	2.080	-0.677	-1.300	0.956	3.560
South Asia	(dropped)	(dropped)	(dropped)	(dropped)	-1.832	-0.450	-1.015	-0.510	-1.405	-2.340	0.352	1.140
Sub-Saharan Africa	13.193	3.950	-11.157	-6.030	-8.982	-2.510	-6.912	-3.900	-1.331	-2.500	-0.764	-2.780
Decade 2 (1970, 1975)	8.372496	4.03	-1.961129	-1.7	12.85378	8.58	-0.302	-0.41	1.613	7.200	-0.122	-1.060
Decade 3 (1980, 1985)	17.722	8.460	-4.864	-4.190	24.218	16.010	-1.766	-2.350	3.301	14.610	-0.461	-3.950
Decade 4 (1990-1999)	20.721	9.450	-7.067	-5.810	32.685	20.410	-3.657	-4.600	4.439	18.620	-0.799	-6.490
Low Income Countries	-63.950	-17.250	27.918	13.580	-45.021	-12.490	9.607	5.370	-6.621	-12.340	1.914	6.900
Lower Middle Income Countries	-38.902	-8.740	18.838	7.630	-28.611	-8.980	6.357	4.020	-3.831	-8.090	1.112	4.550
Upper Middle Income Countries	-25.789	-5.890	12.149	5.000	-16.962	-5.470	3.610	2.340	-2.026	-4.390	0.480	2.020
Avg. Adult Fem. Scl. 6 - 9 yrs	-7.449	-2.930	-1.363	-0.970	21.532	11.640	-4.929	-5.370	2.121	7.700	-0.563	-3.950
Avg. Adult Fem. Scl. 9 yrs or more	-10.954	-2.870	0.274	0.130	24.333	8.880	-5.061	-3.720	2.424	5.960	-0.512	-2.430
Constant	95.837	40.330	5.482	4.160	45.608	26.550	6.321	7.420	9.677	37.940	0.907	6.880
	N 741		N 741		N 732		N 732		N 716		N 716	
	F(13, 727) 64.85		F(12, 727) 44.75		F(12, 718) 249.68		F(12, 718) 34.7		F(12, 702) 201.75		F(12, 702) 49.12	
	Adj R^2 0.5287		Adj R^2 0.4346		Adj R^2 0.8156		Adj R^2 0.3747		Adj R^2 0.7849		Adj R^2 0.4667	
	Root MSE 20.101		Root MSE 11.145		Root MSE 14.405		Root MSE 7.1461		Root MSE 2.1356		Root MSE 1.1038	

Region 9, all developed countries, is used as the omitted group
Decade 1 (1960, 1965) has been used as the omitted group
The category 'Avg. Adult Fem. Schooling<=6yrs' has been used as the omitted group
The category 'Higher Income Countries (OECD and non-OECD)' has been used as the omitted group

Table B.5b Multivariate Relations with Continuous Variables

Independent Variables	Female Gross Primary Enrollment (%)		Male-Female Gap in Gross Primary Enrollment (%)		Female Gross Secondary Enrollment (%)		Male-Female Gap in Gross Secondary Enrollment (%)		Expected years of Female Schooling		Male Female Gap in Expected Years of Schooling	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
East Asia & Pacific	-7.282	-1.860	0.987	0.430	-2.303	-0.780	-1.080	-0.650	-0.387	-0.940	-0.057	-0.230
Europe & C. Asia	-8.373	-1.660	3.039	1.030	-2.543	-0.660	6.113	2.780	-0.471	-0.870	0.764	2.370
Latin America & Caribbean	-9.494	-2.870	-0.803	-0.420	-2.996	-1.190	-7.573	-5.320	-0.572	-1.630	-0.784	-3.740
Middle East & N. Africa	-2.739	-0.710	4.618	2.050	8.355	2.870	3.145	1.910	0.811	2.000	0.595	2.450
South Asia	-21.832	-4.820	10.421	3.930	1.985	0.570	1.777	0.910	-0.810	-1.680	0.665	2.310
Sub-Saharan Africa	-12.956	-3.370	-1.353	-0.600	-8.119	-2.780	-4.370	-2.640	-1.166	-2.860	-0.485	-1.990
Year	0.935	4.410	-0.119	-0.960	1.089	6.780	0.082	0.900	0.132	5.880	0.004	0.300
Year^2	-0.019	-3.410	0.004	1.200	-0.009	-2.090	-0.004	-1.650	-0.001	-1.800	0.000	-0.970
Per Capita GDP (const. 95 prices)	-0.001	-2.840	0.000	1.590	0.003	7.670	0.000	-1.740	0.000	5.170	0.000	-0.600
[Per Capita GDP (const. 95 prices)]^2	0.000	1.990	0.000	-1.580	0.000	-6.370	0.000	2.560	0.000	-4.130	0.000	1.370
Avg. Adult Fem. Scl.	19.901	22.510	-9.522	-18.440	9.999	14.920	-2.348	-6.180	1.891	20.170	-0.711	-12.660
(Avg. Adult Fem. Scl.)^2	-1.460	-18.120	0.653	13.880	-0.440	-7.240	0.087	2.530	-0.106	-12.460	0.041	8.100
Constant	47.832	11.890	29.137	12.400	-9.159	-3.000	14.913	8.600	1.658	3.890	2.955	11.580
	N 669		N 669		N 660		N 660		N 648		N 648	
	F(12, 656) 110.25		F(12, 656) 78.33		F(12, 647) 386.9		F(12, 647) 40.21		F(12, 635) 354.17		F(12, 635) 65.94	
	Adj R^2 0.6625		Adj R^2 0.5814		Adj R^2 0.8754		Adj R^2 0.4165		Adj R^2 0.8676		Adj R^2 0.5464	
	Root MSE 15.845		Root MSE 9.2558		Root MSE 11.911		Root MSE 6.7569		Root MSE 1.6538		Root MSE 0.99	

Region 9, all developed countries, is the omitted group

When all four variables—region, time, income and adult female schooling—are included, the overall relations are more consistent with the variations in female schooling enrollments and expected years of schooling and related gaps than if only one of these four variables are considered at a time. In almost all cases, however, the associations with each of these four individual variables are lessened somewhat in comparison with those when only one is included at a time. But the associations with adult female schooling tend to be lessened least, again reinforcing the strength of the intergenerational female schooling associations. The associations with income also remain fairly strong. But those with region and time tend to become much weaker than when they are included alone in the estimates—suggesting that in substantial part they when they alone are included, they are representing primarily differences in adult female schooling and in income. The gender gaps in schooling enrollment rates associated with region, for example, are much smaller or are not significant or even are reversed to favor females (e.g., for secondary school enrollments in Latin American and the Caribbean) once there is control for time, income and adult female schooling.

APPENDIX C

EXPERIMENTAL EVALUATION: THE EXAMPLE OF MEXICO'S PROGRESA

EXPERIMENTAL EVALUATION: THE EXAMPLE OF MEXICO'S PROGRESA

In 1997, the federal government of Mexico introduced the Programa de Educación, Salud y Alimentación (the Education, Health, and Nutrition Program), known by its Spanish acronym, PROGRESA, as part of an effort to break the intergenerational transmission of poverty. PROGRESA has a multiplicity of objectives, primarily aimed at improving the educational, health and nutritional status of poor families, and particularly of children and their mothers. PROGRESA provides cash transfers linked to an individual's enrollment and regular school attendance and to clinic attendance. The program also includes in-kind health benefits and nutritional supplements for children up to age five, and pregnant and lactating women. By the end of 1999, PROGRESA covered approximately 2.6 million families or about 40 percent of all rural families and one-ninth of all families in Mexico.[46] At that time the program operated in almost 50,000 localities in more than 2,000 municipalities and 31 states. PROGRESA's budget of approximately $777 million in 1999 was equivalent to 0.2 percent of Mexico's GDP. In early 1998, the International Food Policy Research Institute (IFPRI) was asked to assist the PROGRESA administration to "determine if PROGRESA is functioning in practice as it is intended to by design." The evaluation is based on longitudinal data collected from 24,000 households from 506 localities in seven states who were interviewed periodically between November 1997 and November 1999. Of the 506 localities, 320 localities were assigned to the treatment group and 186 localities were assigned as controls. Specifically, the 320 treatment localities were randomly selected using probabilities proportional to size from a universe of 4,546 localities that were covered by phase II of the program in seven states. Using the same method, the 186 control localities were selected from a universe of 1,850 localities in these seven states that were to be covered by PROGRESA in later phases. Statistical tests confirm that indeed assignment to treatment versus control was random at the community level (Behrman and Todd 1999).[47] As originally planned the localities serving the role of a control group started receiving PROGRESA benefits by December 2000. A number of evaluation studies have been carried out using these data, in most cases exploiting some dimension of the experimental design of the study (Skoufias 2001 provides a synthesis of many of these studies and of the overall program).

PROGRESA has several striking and unusual features from the perspective of social scientists and policy analysts.

(1) PROGRESA based some of its essential components on the outcomes of social science research in the literature. Transfers were given to women, for example, because previous

---

[46] PROGRESA has evolved into OPORTUNIDADES under the Fox administration, and has expanded into semi-urban and urban areas (cities up to one million inhabitants) with some modifications such as covering higher grades of schooling and self-selection for applying for enrollments. In 2002 the target is to enroll a million families. If that target is achieved the program will cover over 20 million individuals.

[47] The same study reports that there are somewhat more significant differences in some variables than would be expected by chance at the household and individual level. But most of these differences, though significant due to the large sample size, are not very substantial.

research on intrahousehold allocations suggested that income directed towards mothers had larger associations with investments in children than income directed towards fathers.

(2) PROGRESA used modern social science tools in order to guide its decisions. For instance discriminant analysis was used on census data in the initial stages of identifying target communities and households within those communities and GIS systems were used to systematize information on location of schooling and health services relative to the communities.

(3) PROGRESA recognized that baseline data and longitudinal household and service-provider data with treatment and control groups were essential to serious evaluation of the program and implemented the collection of baseline data (prior to the introduction of the program, which always is the stated intent but often not the realization in various developing country contexts) with longitudinal follow-up and random assignment to treatment and control communities.

(4) PROGRESA, as noted, contracted an outside research agency, IFPRI (a member of the CIGAR group of international agricultural research institutions, with a strong history of data analysis and evaluation in developing countries), to undertake an extensive evaluation of the program.

These are considerable and important features of PROGRESA. Moreover PROGRESA and the key individuals behind PROGRESA not only incorporated such features into their program plans, but – which is much more difficult – gave them sufficient priority that they were carried out reasonably well during the very difficult time of program development, implementation and rapid expansion, all within an environment with not inconsiderable political pressures. PROGRESA already has been a model for development of related programs and evaluation strategies elsewhere, for example, in Brazil Colombia, Honduras, Nicaragua, . Hopefully these and other programs will build not only on the substantive aspects of PROGRESA, but also learn from and improve upon the program evaluation that has been given considerable thought and high priority by PROGRESA.

Other programs, first of all, usefully could emulate the strengths of PROGRESA in terms of building on existing social science research, data collection and evaluation that are noted above. These are considerable strengths and require substantial commitments to give sufficient priority to these matters in the presence of all of the great pressures and unanticipated problems that a new program inevitably faces. But the result is the potential for much better evaluation of the program and how it or other programs can be modified to improve the attainment of the objectives.

But beyond emulating the strong points of the PROGRESA in the above regards, other programs could improve upon PROGRESA in some respects that would enhance more the evaluation possibilities. Of course some of these improvements from the point of view of evaluation may have political and other costs. We recognize that there are such costs, but have no special insight regarding them, so we do not discuss them here.

(1) The evaluation design could explore a number of aspects that are difficult to explore with the present PROGRESA data or that require imposing a lot of structure to do so. There could be randomly assigned variations (perhaps across communities to avoid invidious comparisons within communities), for example, in the payment schedule for attending different grades of school by gender; in whether the payments are made to mothers, fathers or the children themselves; in whether payments are made to the demand side (households, individuals) or to suppliers (e.g., schools) directly; in whether other components of the PROGRESA package were included; in whether these payments are conditional on attending schooling;

(2) The data collection/evaluation process could be made more independent of the implementing agency. The evaluating agency could be contracted by some other part of the government (not by the implementing agency) and have direct responsibility for collecting the data to be used (rather than having the implementing agency in charge of data collection). Data could be made available for public use earlier. These changes would increase credibility regarding the degree of independence of the evaluation.

Both of these changes are being implemented in part for the evaluation of the expansion of PROGRESA in urban areas in the current OPORTUNIDADES program, with the JOVENES CON OPORTUNIDADES supplementary incentives for secondary school completion.

APPENDIX D

REFERENCES CITED AND OTHER RELEVANT STUDIES

REFERENCES CITED AND OTHER RELEVANT STUDIES

Amin, S, I Diamond, R T Naved, M Newby, 1998, "Transition to Adulthood of Female Garment-factory Workers in Bangladesh," *Studies in Family Planning* 29, 185-200.

Ahmed, Akhter U. and Carlo del Ninno, 2002, "The Food for Education Program in Bangladesh: an Evaluation of its Impact on Educational Attainment and Food Security," Washington, DC: International Food Policy Research Institute, http://www.cgiar.org/ifpri/divs/fcnd/dp.htm.

Alderman, Harold, 2001, "Multi-Tier Targeting of Social Assistance: The Role of Intergovernmental Transfers," *World Bank Economic Review* 15:1,33-54.

Alderman, Harold, Jere R. Behrman, David Ross, and Richard Sabot, 1996a, "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labor Market," *Oxford Bulletin of Economics and Statistics* 58:1 (February), 29-55.

Alderman, Harold, Jere R. Behrman, David Ross, and Richard Sabot, 1996b, "Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy," *Journal of Human Resources* 31:1 (Winter), 229-254.

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon, "Child Health and School Enrollment: A Longitudinal Analysis," *Journal of Human Resources* 36:1 (Winter 2001), 185-205.

Alderman, Harold, Peter Orazem and Elizabeth M. Paterno. 2001. "School Quality, School Cost and the Public/private School Choices of Low-income Households in Pakistan," *Journal of Human Resources* 36(2): 304-326.

Angrist, Joshua D., 1990, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review* 80, 313-336.

Angrist, Joshua D., Eric Bettinger, Erik Bloom, Elizabeth King and Michael Kremer, 2002, "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review* 92:5 (December), 1535-59.

Angrist, Joshua D. and Victor Lavy, 1997, "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco," *Journal of Labor Economics* 15:1 Part 2 (January), S48-S76.

Angrist, Joshua D., and Victor Lavy, 1999, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *The Quarterly Journal of Economics* 114:2 (May): 533-76.

Ashenfelter, Orley and Alan Krueger, 1994, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review* 84:5 (December), 1157-74.

Attanasio, Orazio, Costas Meghir and Ana Santiago, 2002, "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to evaluate Progresa," London: University College London, mimeo.

Bacolod, Marigee and Elizabeth M. King, "The Effects of Family Background and School Quality on Low and High Achievers: Determinants of Academic Achievement in the Philippines," Washington, DC: World Bank, mimeo, 2000.

Banerjee, Abhijit and Kremer, Michael; with Lanjouw, Jean and Lanjouw, Peter. "Teacher-Student Ratios and School Performance in Udaipur, India: A Prospective Evaluation." Mimeo, Harvard University, 2002.

Barro, Robert J. and Jong-Wha Lee, 1993, "International Comparison of Educational Attainment," *Journal of Monetary Economics* 32:3 (December), 363-394.

Barro, Robert and X. Sala-I-Martin, 1995, *Economic Growth*. New York: McGraw-Hill.

Basu, Kaushik and James E. Foster, 1998, "On Measuring Literacy," *Economic Journal* 108:451 (November), 1733-49.

Basu, K, A Narayan, M Ravallion, 2002, "Is Literacy Shared Within Households? Theory and Evidence for Bangladesh," *Labour Economics* 8.

Becker, Gary S., 1967, "Human Capital and the Personal Distribution of Income: An Analytical Approach," Ann Arbor: University of Michigan, Woytinsky Lecture, republished in Gary S. Becker, *Human Capital*, New York: NBER, 2nd edition 1975, 94-117.

Behrman, Jere R., 1997, "Women's Schooling and Child Education: A Survey," Philadelphia, PA: University of Pennsylvania, mimeo.

Behrman, Jere R. and Nancy Birdsall, 1983, "The Quality of Schooling: Quantity Alone is Misleading," *American Economic Review* 73, 928-946.

Behrman, Jere R. and Nancy Birdsall 1985, "The Quality of Schooling: Reply," *American Economic Review* 75, 1202-1205.

Behrman, Jere R. and Nancy Birdsall, 1988a, "Implicit Equity-Productivity Tradeoffs in the Distribution of Public School Resources in Brazil," *European Economic Review* 32:8 (October), 1585-1601.

Behrman, Jere R. and Nancy Birdsall, 1988b, "The Reward for Good Timing: Cohort Effects and Earnings Functions for Brazilian Males," *Review of Economics and Statistics* 70:1, 129-135.

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 Foregone: Education in Brazil*, Baltimore, MD: The Johns Hopkins University Press for the Inter-American Development Bank, 245-266.

Behrman, Jere, Yingmei Cheng and Petra Todd, 2003, "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach," Philadelphia, PA. University of Pennsylvania, mimeo.

Behrman, Jere R. and Anil B. Deolalikar, 1991a, "The Poor and the Social Sectors During a Period of Macroeconomic Adjustment: Empirical Evidence for Jamaica," *World Bank Economic Review* 3:2 (May), 291-313.

Behrman, Jere R. and Anil B. Deolalikar, 1991b, "School Repetition, Dropouts and the Returns to Schooling: The Case of Indonesia," *Oxford Bulletin of Economics and Statistics* 53:4 (November), 467- 80.

Behrman, Jere R. and Anil B. Deolalikar, 1993, "Unobserved Household and Community Heterogeneity and the Labor Market Impact of Schooling: A Case Study for Indonesia," *Economic Development and Cultural Change* 41:3 (April), 461-488.

Behrman, Jere R. and Anil B. Deolalikar, 1995, "Are There Differential Returns to Schooling by Gender? The Case of Indonesian Labor Markets," *Oxford Bulletin of Economics and Statistics* 57:1 (February), 97-118.

Behrman, Jere R., Suzanne Duryea and Miguel Székely, 1999, "Schooling Investments and Aggregate Conditions: A Household-Survey-Based Approach for Latin America and the Caribbean," Washington, DC; Inter-American Development Bank, Background Paper for IPES 2000 for IDB, mimeo.

Behrman, Jere R, Suzanne Duryea and Miguel Székely, 2003, "Aging and Economic Opportunities: Major World Regions around the Turn of the Century," in Orazio Attanasio and Miguel Székely, eds., *A Dynamic Analysis of Household Decision-Making in Latin America*, Washington, DC; Inter-American Development Bank, 2002.

Behrman, Jere R., Andrew Foster, Mark R. Rosenzweig and Prem Vashishtha, 1999, "Women's Schooling, Home Teaching, and Economic Growth," *Journal of Political Economy* 107:4 (August), 682-714.

Behrman, Jere R. and Elizabeth M. King, 2001, "Household Schooling Behaviors and Decentralization," *Economics of Education Review* 20:4 (2001), 321-341.

Behrman, Jere R. and James C. Knowles, 1998a, "Population and Reproductive Health: An Economic Framework for Policy Evaluation," *Population and Development Review* 24:4 (December), 697-738.

Behrman, Jere R. and James C. Knowles, 1998b, "The Distributional Implications of Government Family Planning and Reproductive Health Services in Vietnam," prepared for the Rockefeller Foundation, Philadelphia: University of Pennsylvania, mimeo.

Behrman, Jere R. and James C. Knowles, 1999, "Household Income and Child Schooling in Vietnam," *World Bank Economic Review* 13:2 (May 1999), 211-256.

Behrman, Jere R. and Mark R. Rosenzweig, 1994, "Caveat Emptor: Cross-Country Data on Education and the Labor Force," *Journal of Development Economics* 44:1 (June), 147-172.

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 (1999), 159-167.

Behrman, Jere R. and Mark R. Rosenzweig, 2002,"Does Increasing Women's Schooling Raise the Schooling of the Next Generation?", *American Economic Review* 92:1(March 2002), 323-334.

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 (December), 1131-1174.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman, 1996, "College Choice and Wages: Estimates Using Data on Female Twins," *Review of Economics and Statistics* 73:4 (November), 672- 685.

Behrman, Jere R., David Ross and Richard Sabot, 2003, "Improving the Quality Versus Increasing the Quantity of Schooling: Evidence from Rural Pakistan," *Topics in the Economics and Growth of Developing Areas of The B.E. Journals of the Economics and Growth of Developing Areas* (2003, forthcoming).

Behrman, Jere R. and Ryan Schneider, 1994, "An International Perspective on Schooling Investments in the Last Quarter Century in Some Fast-Growing East and Southeast Asian Countries," *Asian Development Review* 12:2, 1-50.

Behrman, Jere R. and Piyali Sengupta, 2002, "The Returns to Female Schooling In Developing Countries Revisited," Philadelphia, PA: University of Pennsylvania, prepared for the World Bank.

Behrman, Jere R., Piyali Sengupta and Petra Todd, 2002, "Progressing through PROGRESA: An Impact Assessment of Mexico's School Subsidy Experiment," Philadelphia, PA: University of Pennsylvania,, processed.

Behrman, Jere R., and Petra E. Todd, 1999, "Randomness in the experimental samples of PROGRESA (education, health, and nutrition program)," February. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, D.C.

Berry, Albert, 2002, "Poverty-Problem Country Typologies," Bethesda, MD: DAI and BIDE.

Boissiere, Maurice, John B. Knight, and Richard H. Sabot, 1985, "Earnings, Schooling, Ability and Cognitive Skills," *American Economic Review* 75, 1016-1030.

Bommier, Antoine, and Sylvie Lambert. 2000. "Education Demand and Age at School Enrollment in Tanzania," *Journal of Human Resources* 35(1): 177-203.

Bourguignon, François, Francisco H. G. Ferrerira and Phillippe G. Leite, 2002, "Ex-ante Evaluation of Conditional Cash Transfer Programs: The Case of Bolsa Escola," Washington, DC: World Bank, processed.

Card, David E., 1995, "Earnings, Schooling, and Ability Revisited," *Research in Labor Economics* 14, 23-48.

Card, David E., 1999, "New Developments in the Economics of Schooling," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4, Amsterdam: North-Holland Publishing Co.

Card, David and Alan B. Krueger, 1992, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100:1 (February), 1-40.

Card, David and Alan B. Krueger, 1992, "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107:1 (February), 151-200.

Caribbean Group for Cooperation in Economic Development (CGCED), 2002, Youth development in the Carribbean, Draft Report No. 24163-LAC, Washington, DC: The World Bank (June).

Case, Anne and Angus Deaton, 1999, "School Inputs and Educational Outcomes in South Africa," *Quarterly Journal of Economics* 114:3 (August), 1047-1084.

Cawley, John, James Heckman and Edward Vytlacil, 2001, "Three Observations on Wages and Measured Cognitive Ability," *Labour Economics* 8 (September), 419-442.

Coale, Ansley J. and Edgar M. Hoover, 1958, *Population Growth and Economic Development in Low- Income Countries: A Case Study of India's Prospects*, Princeton, NJ: Princeton University Press.

de la Fuente, Angel and Rafael Doménech, 2002, "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make? An Update and Further Results," Barcelona: Universidad de Valencia, mimeo.

Devarajan, S., L. Squire and S. Suthiwart-Narueput, 1997, "Beyond Rate of Return: Reorienting Project Appraisal, *The World Bank Research Observer* 12(1):35-46.

Duflo, Esther, 2001, "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review* 91:4 (September), 795-813.

Feldstein, Martin, 1995, "Tax Avoidance and the Deadweight Loss of the Income Tax," Cambridge, MA: National Bureau of Economic Research Working Paper No. 5055.

Foster, Andrew and Mark R. Rosenzweig, 1995, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy* 103:6 (December),1176-1209.

Foster, Andrew D. and Mark R. Rosenzweig, 1996, "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *American Economic Review* 86:4 (September), 931-953.

Foster, Andrew W. and Mark R. Rosenzweig, 2002, "Household Division and Rural Economic Growth," *Review of Economic Studies* (October).

Gershberg, Alec Ian and Til Schuermann, 1999, "Education Finance in a Federal System: Changing Investment Patterns in Mexico," *Economics of Education Review*.

Glewwe, Paul, 1996, "School Quality, Student Achievement, and Fertility in Developing Countries: An Assessment of Current Knowledge and Recommendations for Future Research," Washington, DC: World Bank, mimeo (prepared for the National Academy of Science Committee on Population.

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 (December), 267-290.

Glewwe, Paul with contributions by Jaikishan Desai, Dean Jolliffe, Raylynn Oliver and Wim Vijverberg, 1996, *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*, London: Macmillan.

Glewwe Paul, Hanan Jacoby and Elizabeth King. 2001. "Early Childhood Nutrition and Academic Achievement: a Longitudinal Analysis," *Journal of Pubic Economics* 81 (September 2001), 345-368

Glewwe Paul and King Elizabeth. 2001. "The Impact of Early Childhood Nutrition Status on Cognitive Achievement: Does the Timing of Malnutrition Matter? " *World Bank Economic Review* 15:1 (May), 81-113.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin, 2000, "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya," Washington, DC: World Bank, mimeo.

Glewwe, Paul, Michael Kremer, Sylvie Moulin and Eric Zitzewitz, 2003, "Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya." *Journal of Development Economics* (forthcoming).

Gundlach, Erich and Ludger Woesmann, 2000, "The Fading Productivity of Schooling in East Asia," Kiel Institute of World Economics, Kiel, Germany.

Handa, Sudhanshu and Kenneth R. Simler. 2000. "Quality or Quantity? The Supply Side Determinants of Primary Schooling in Rural Mozambique," Food Consumption and Nutrition Division Discussion Paper no. 83. Washington, D.C.: International Food Policy Research Institute.

Hanushek, Eric A., 1995, "Interpreting Recent Research On Schooling in Developing Countries," The *World Bank Research Observer* 10:2 (August), 227-246.

Hanushek, Eric A., Joao Batista Gomes-Neto, and Ralph W. Harbison, 1996, "Efficiency-Enhancing Investments in School Quality," in *Opportunity Foregone: Education in Brazil*, Nancy Birdsall and Richard Sabot, eds., Washington, D.C.: Inter-American Development Bank; distributed by Johns Hopkins University Press, Baltimore, 385-424.

Hanushek, Eric A., and Dennis D. Kimko. 2000. "Schooling, Labor-Force Quality, and the Growth of Nations." *American Economic Review* 90 (5)1184-1208.

Harberger, Arnold C., 1997, "New Frontiers in Project Evaluation? A Comment on Devarajan, Squire and Suthiwart-Narueput," *The World Bank Research Observer* 12(1):73-79.

Heckman, J., H. Ichimura, and P. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies* 64:4, 605-654.

Hoxby, Caroline M., 2002/2003, "Economics of Education," *NBER Reporter* (Winter 2002/2003), 1-5.

Jacoby, Hanan, 1994, "Borrowing Constraints and Progress Through School: Evidence from Peru," *Review of Economics and Statistics* 76:1 (February), 151-160.

Jalan, Jyotsna and Elena Glinskaya, 2002, "Improving Primary School Education in India: an Impact Assessment of Dpep-Phase I," New Delhi: Indian Statistical Institute and Washington: World Bank, processed.

James, Estelle, Elizabeth M. King and Ace Suryadi, 1996, "Finance, Management, and Costs of Public and Private Schools in Indonesia," *Economics of Education Review* 15:4, 387-398.

Jimenez, Emmanuel, Marlaine E. Lockheed, and Vicente Paqueo, 1991, "The Relative Efficiency of Private and Public Schools in Developing Countries," *The World Bank Research Observer* 6:2 (July), 205- 218.

Jimenez, Emmanuel and Vicente Paqueo, 1996, "Do Local Contributions Affect the Efficiency of Public Primary Schools," *Economics of Education Review* 15:4, 377-386.

Kane, Thomas J. and Douglas O. Staiger, 2002, "The Promise and Pitfalls of Using Imprecise School Accountability Measures," *Journal of Economic Perspectives* 16:4 (Fall), 91-114.

Kim, Jooseop, Harold Alderman, and Peter F. Orazem. 1999. "Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program," *The World Bank Economic Review* 13(3): 443-465.

King, Elizabeth M. and M. Anne Hill, eds, 1993, *Women's Education in Developing Countries: Barriers, Benefits, and Policies*, Baltimore and London: The Johns Hopkins University Press, Published for the World Bank.

King, Elizabeth M. and L.A. Lillard, 1987, "Education Policy and Schooling Attainment in Malaysia and the Philippines," *Economics of Education Review* 6:2 (November), 167-181.

King, Elizabeth M. and Andrew Mason, 2000, *Engendering Development: Gender Equality in Rights, Resources, and Voice,* New York: Oxford University Press for the World Bank.

King, Elizabeth M, and K. McLean. "Decentralization of the education sector" in J. Litvack and J. Seldon (eds.), *Decentralization Briefing Notes*, World Bank Institute Working Papers, Washington DC, 1999.

King, Elizabeth M. and Berk Özler, 1998, "What's Decentralization Got to Do With Learning? The Case of Nicaragua's Education Reform," Washington, DC: World Bank, Development Research Group, Impact of Education Reform s Working Paper Np. 9.

Kingdon, Geeta and Francis Teal, 2002, "Does Performance Related Pay for Teachers Improve Student Performance? Some Evidence from India," Oxford, UK: University of Oxford. Department of Economics, processed.

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

Knowles, James C. and Jere R. Behrman, 2003b, "The Economic Returns to Investing in Youth in Developing Countries: A Review of the Literature," Bangkok/Philadelphia, PA: University of Pennsylvania, mimeo, for World Bank Human Development Network.

Knowles, James C. and Jere R. Behrman, 2003a, "Assessing the Economic Benefits of Investing in Youth in Developing Countries," Bangkok/Philadelphia, PA: University of Pennsylvania, mimeo, for World Bank Human Development Network.

Knowles, James C. and Jere R. Behrman, 2003b, "Assessing the Economic Benefits of Investing in Youth in Developing Countries: Literature Review," Bangkok/Philadelphia, PA: University of Pennsylvania, mimeo, for World Bank Human Development Network.

Kremer, Michael R., 1995, "Research on Schooling: What We Know and What We Don't: A Comment on Hanushek," *World Bank Research Observer* 10:2 (August), 247-254.

Kremer, Michael, 2003, "Randomized Evaluation of Educational Programs in Developing Countries: Some Lessons," *American Economic Review* (May, forthcoming).

Kremer, M., Moulin, S., Myatt, D. and Manunyu, R., 1996, "Textbooks, Class-Size, and Test-Scores: Evidence from a Prospective Evaluation in Kenya," Cambridge, MA: MIT, mimeo.

Ladd, Helen F., 2002, "School Vouchers: A Critical View," *Journal of Economic Perspectives* 16:4 (Fall), 3-24.

Lam, David and Deborah Levison, 1991, "Declining Inequality in Schooling in Brazil and its Effect on Inequality in Earnings," *Journal of Development Economics* 37:1/2 (November), 199-226.

Lam, David and Robert F. Schoeni, 1993, "Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil," *Journal of Political Economy* 101:4 (August), 710-740.

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

Lavy, Victor, 2002, "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," *Journal of Political Economy* 110:61, 1286-1317.

Lloyd, Cynthia B., Barbara S. Mensch, and Wesley H. Clark. 2000. "The Effects of Primary School Quality on School Dropout among Kenyan Girls and Boys," *Comparative Education Review* 44(2): 113-147.

Lloyd, Cynthia B., Cem Mete and Zeba A. Sathar, 2002, "The Effect of Gender Differences in Primary School Access, Choice and Quality on the Decision to Enroll in Rural Pakistan," New York: Population Council, mimeo.

Martorell, R. 1999. The Nature of Child Malnutrition and its Long-Term Implications," *Food and Nutrition Bulletin* 20, 288-292.

Miguel, Edward and Michael Kremer, 2001, "Worms: Education and Health Externalities in Kenya," Harvard University and University of California-Berkeley, mimeo.

Miller, Paul, Charles Mulvey and Nick Martin, 1995, "What Do Twins Studies Tell Us About The Economic Returns to Education? A Comparison of US and Australian Findings," *American Economic Review* 85:3, 586-599.

Miller, Paul, Charles Mulvey and Nick Martin, 1997, "Family Characteristics and the Returns to Schooling: Evidence on Gender Differences from a Sample of Australian Twins," *Economica* 64:253 (February), 119-136.

Montgomery, Mark and Cynthia—Lloyd, 1999, "Excess Fertility, Unintended Births, and Children's Schooling," in C Bledsoe, JB Casterline, JA Johnson-Kuhn, and JG Haaga (eds), *Critical Perspectives on Schooling and Fertility in the Developing World*: Washington, DC, National Academy Press, 216-266.

Murnane, Richard J., John B. Willet and Frank Levy, 1995, "The Growing Importance of Cognitive Skills in Wage Determination," *Review of Economics and Statistics* 77:2 (May), 251-266).

Muskin, J, 1999, "Including Local Priorities to Assess School Quality: the Case of Save the Children Community Schools in Mali," *Comparative Education Review* 43, 36.

Neal, Derek, 2002, "How Vouchers Could Change the Market for Education," *Journal of Economic Perspectives* 16:4 (Fall), 25-44.

Page, John, et al., 1993, *The East Asian Miracle: Economic Growth and Public Policy*, Washington, D.C.: The World Bank.

Paxson, Christina and Norbert R. Schady, 2002, "The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru," *World Bank Economic Review* 16:2, 297-319.

Pitt, Mark M., Mark R. Rosenzweig, and Donna M. Gibbons, 1993, "The Determinants and Consequences of the Placement of Government Programs in Indonesia," *The World Bank Economic Review* 7:3 (September), 319-348.

Pritchett, Lant, 2003, "It Pays to Be Ignorant," *Journal of Policy Reform* (forthcoming)..

Psacharopoulos, George, 1994, "Returns to Investment in Education: A Global Update," *World Development* 22:9 (September), 1325-1344.

Psacharpoulos, George, Rojas and Velez, 1993

Ramakrishnan, U., H. Barnhard, D.G. Schroeder, A.D. Stein and R. Martorell. "Early Childhood Nutrition, Education and Fertility Milestones in Guatemala." *Journal of Nutrition* 129(12):2196-2202, 1999.

Rao, Vijayendra, 1993, "The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India," *Journal of Political Economy* 101:4 (August), 666-677.

Rosenzweig, Mark R., 2000, "Schooling, Learning and Economic Growth," in *Back to Shared Prosperity: The Growing Inequality of Wealth and Income in America*, Ray Marshall, ed., Economic Policy Institute, Washington, DC.

Rosenzweig, Mark R., 1995, "Why Are There Returns in Schooling?" *American Economic Review* 85:2 (May), 153-158.

Rosenzweig, Mark R., 1982, "Educational Subsidy, Agricultural Development, and Fertility Change," *Quarterly Journal of Economics* 97:1.

Rosenzweig, Mark R. and Kenneth J. Wolpin, 1986, "Evaluating the Effects of Optimally Distributed Public Programs," *American Economic Review* 76:3 (June), 470-487.

Rosenzweig, Mark R. and Kenneth J. Wolpin, 1993, "Credit Market Constraints and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks," *Journal of Political Economy* 101:2 (April), 223-245.

Sánchez-Páramo, Carolina and Norbert Schady, 2003, "Off and Running? Technology, Trade, and the Rising Demand for Skilled Workers in Latin America," Washington, DC: World Bank, processed.

Sawada, Yasuyuki and Michael Lokshin, 2000, "Sequential Schooling Decisions of Households: An Analysis of Field Survey Data from Rural Pakistan," Tokyo and Washington: University of Tokyo and World Bank, processed.

Schultz, T. Paul, "The Impact of PROGRESA on School Enrollments," Washington, DC: International Food Policy Research Institute, processed, 2000.

Schultz, T. Paul, 2003,. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program," *Journal of Development Economics* (forthcoming).

Sen, Amartya, 2002, "The Pratichi Report," Pratichi India Trust.

Skoufias, Emmanuel, 2001, "PROGRESA and its Impacts on the Welfare and Human Capital of Adults and Children in Rural Mexico: A Synthesis of the Results of an Evaluation by the International Food Policy Research Institute (IFPRI)," Washington, DC: International Food Policy Research Institute, Food Consumption and Nutrition Division, mimeo.

Skoufias, Emmanuel and Susan Parker, 2001, "Conditional Cash Transfers and Their Impact on Child Work and Schooling: Evidence from the PROGRESA Program in Mexico," *Economia* 2:1 (Fall 2001), 45-96.

Srinivasan, T. N., 1994, "Data Base for Development Analysis: An Overview," *Journal of Development Economics* 44:1 (June), 3-26.

Summers, Lawrence H., 1992, "Investing in All the People," *Pakistan Development Review* 31:4 (Winter), 367-406.

Summers, Lawrence H., 1994, "Investing in All the People: Educating Women in Developing Countries," Washington, DC: World Bank, Economic Development Institute Seminar Paper No. 45.

Tan, J-P, J Lane, and G Lassibille, 1999, "Student Outcomes in Philippine Elementary Schools: an Evaluation of Four Experiments," *The World Bank Economic Review*. 13, 493-508.

Todd, Petra and Kenneth Wolpin, 2003, "Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico," Philadelphia, PA" University of Pennsylvania, processed.

Tomasevski, K, 2001, "Free and Compulsory Education for All Children: The Gap between Promise and Performance" (Right to Education Primers No. 2), Lund and Stockholm, Sweden, Raoul Wallenberg Institute and Swedish International Development Cooperation Agency (Sida).

Truong Si Anh, John Knodel, David Lam and Jed Friedman, 1996, "Family Size and Children's Education in Vietnam," *Demography* 35:1 (February), 57-70.

Truong Si Anh, John Knodel, David Lam and Jed Friedman, 1996, "Family Size and Children's Education in Vietnam," *Demography* 35:1 (February), 57-70.

Vermeersch, Christel, 2002, "School Meals, Educational Achievement and School Competition: Evidence from a Randomized Experiment," Cambridge, MA: Harvard University, mimeo.

World Bank, 1995, *Viet Nam: Poverty Assessment and Strategy*, Report No. 13442-VN, East Asia and Pacific Region, Washington, DC: The World Bank.

World Bank, 2001, *World Bank Development Indicators on CD-ROM*, Washington, DC: World Bank.

Yang, Dennis Tao and Xiaodong Zhu, 2000, "Family Investments in Education in Economic Development," Durham, NC: Duke University, mimeo.

The goal of the USAID-funded Pro-Poor Economic Growth Research Studies and Guidance Manual Activity is to identify and disseminate policies, reforms, and activities that USAID decision makers can incorporate into their programs and that they can recommend to countries wishing to pursue strongly pro-poor, poverty-reducing, economic growth objectives.

The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of USAID.

DAI

Development Alternatives, Inc.
7250 Woodmont Avenue, Suite 200
Bethesda, Maryland 20814 USA

301-718-8699 info@dai.com www.dai.com

BIDE

Boston Institute for Developing Economies, Ltd.
4833 West Lane, Suite 100
Bethesda, Maryland 20814

301-652-9740 manage@bide.com www.bide.com

U.S. Agency for International Development
Bureau for Economic Growth, Agriculture, and Trade
1300 Pennsylvania Avenue, N.W.
Washington, D.C. 20523

www.usaid.gov