

RAISING EDUCATIONAL ATTAINMENT OF THE POOR: POLICIES AND ISSUES

SEIRO ITO

Institute of Developing Economies, JETRO, Chiba, Japan

First version received March 2006; final version accepted June 2006

Despite every policymaker's recognition, enrollment rates of the low-income countries remain low. A simple framework of understanding educational outcomes is presented using a unitary model with an altruistic parent and a child. The traditional interventions, so-called supply-side policies, and recent innovation of relaxing constraints faced by households, the conditional transfer programs or so-called demand-side policies, are reviewed. In addition, recent trends on estimation technique are discussed. It has been argued that randomization is clearly the best for inference, however, one may still want to choose the optimal combination of randomized experiments and observational data, as the former requires more resources and time. This is particularly true for economics than other hard sciences, partly because of our inability to fine-tune the control, and partly because of our lack of solid microfoundation than other sciences when an experiment shows unexpected results.

Keywords: Child labor; Schooling; Experimental studies; Observational studies
JEL classification: J22, O15, I28, C81

I. INTRODUCTION

WHEN we look at countries that experienced remarkable economic growth in the past, we notice one common feature: human capital played a major role in this growth. The East Asian countries that have succeeded in greatly reducing poverty show us that education is a key component in poverty reduction and long-term welfare improvement. But it is not easy to get onto the path of economic growth and break the vicious circle of poverty: because parents are poor, they need their children to work, and because children work, they become parents who are poor.

Policymakers have long understood the importance of human capital in the growth process and have made efforts to raise the level of education attained by the poor. In the early days of development economics, there had been a consensus that a high growth sector would trickle down benefits to the backward, low-growth sectors of an economy. Later a more direct assault on poverty was initiated under programs directed at basic human needs. Although in recent years the structural adjustment programs have slashed education budgets considerably, policymakers

have been mindful of the importance of education and have sought to maintain spending levels on education. A United Nations summit on education in 1990 issued a declaration of "Education for All" that called for universal primary education to become the norm in all countries. In 2000 the UN set down Millennium Development Goals that did not just restate the need for universal primary education, but also reinforced the notion that there should be gender equality at all levels of education.

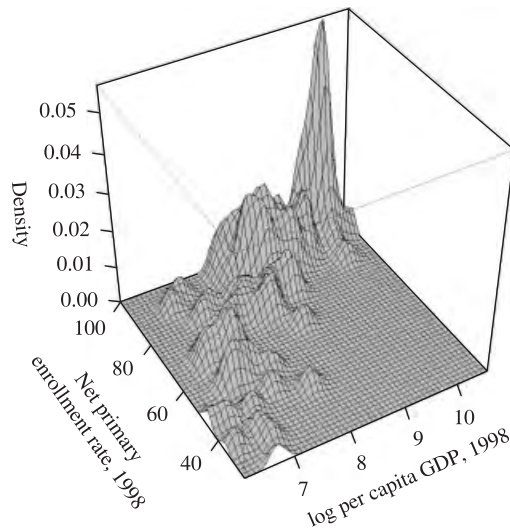
But the declaration and setting of goals has not led to noticeable improvement in educational attainment. Measurements using graphical statistical methods show considerable heterogeneity in net school enrollment rates among developing countries. Figure 1 presents the plotted (and smoothed) three-dimensional density over the space defined by the log per capita GDP and primary school enrollment rates. As expected, there is a positive correlation. One can see that the correlation is weak for low-income countries, and there is substantial heterogeneity among poor countries. Secondary school net enrollment rates, shown in Figure 2, also exhibit a positive correlation. There is little heterogeneity among the low-income countries; they are concentrated around 20%. Heterogeneity in secondary school enrollment rates can be seen among middle-income countries as a bimodality of the distribution.

Looking closer at these measurements, we can see some reasons for the heterogeneity. Figure 3 shows the density over the public primary (secondary) education expenditure per student and primary (secondary) school net enrollment rates, and Figure 4 the density over the public primary (secondary) education expenditure per student and primary (secondary) school net enrollment. There is a clear positive correlation in both figures, except for primary education in Figure 3. A similar picture can be noted in Figure 5 which plots the density for students completing four years of primary education. There is heterogeneity among low- and middle-income countries, while high-income countries look more uniform. This tendency is unchanged when we look at the student repeater rate in primary education; there is considerable bimodality among low-income countries while middle- and high-income countries are more uniform within their groups (Figure 6).

To further understand factors behind the differences in per student expenditure, the densities over ratio of student population to total population (dependency rates) has been plotted in Figures 7 and 8. We can see a significant variation in per student primary expenditure among high dependency rate countries which correspond to low- and middle-income countries. This may explain the variability in primary enrollment heterogeneity among low-income countries. For the secondary expenditure per student, there is clearly a negative correlation which partly explains why low- and middle-income countries struggle to raise the rate of secondary school enrollment.

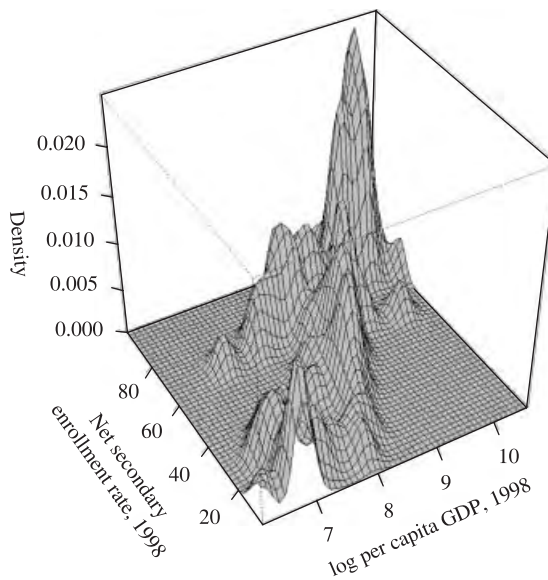
Despite the daunting task of raising educational attainment of the poor, past development gives us some hope for optimism. In the following sections we will examine the education policy intervention and its effects. In the next subsection, we will take up a simple framework for classifying the types of government interventions based

Fig. 1. Net Primary Enrollment Rate and Per Capita GDP



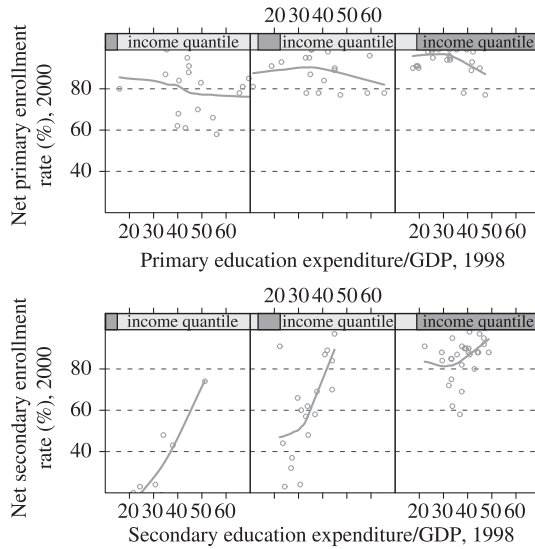
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 2. Net Secondary Enrollment Rate and Per Capita GDP



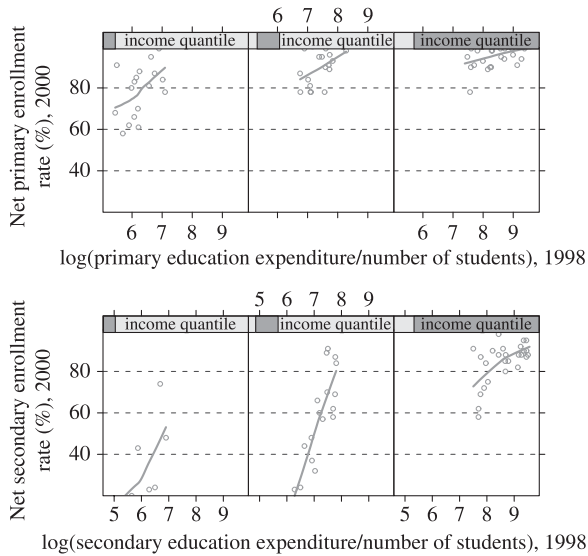
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 3. Net Enrollment Rate and Public Education Expenditure, 1998



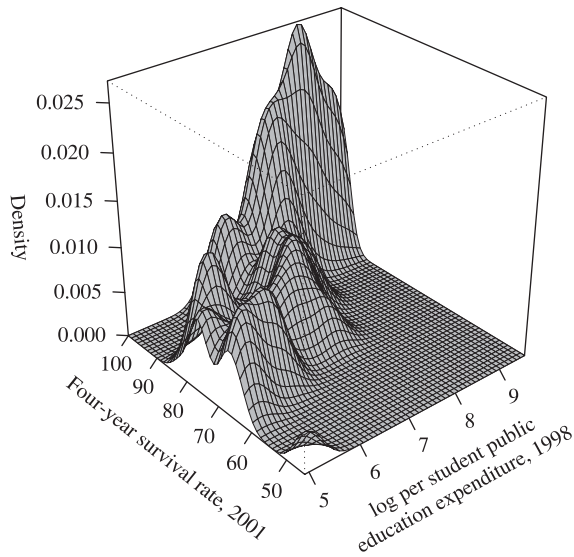
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 4. Net Enrollment Rate and Per Student Public Education Expenditure, 1998



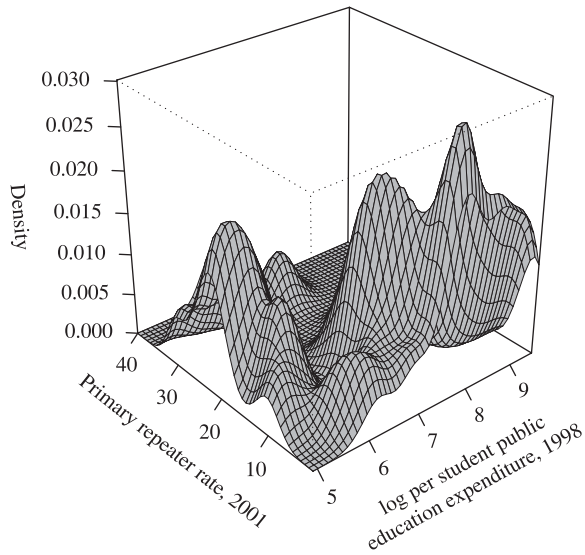
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 5. Four-Year Survival Rate and Per Student Public Primary Education Expenditure



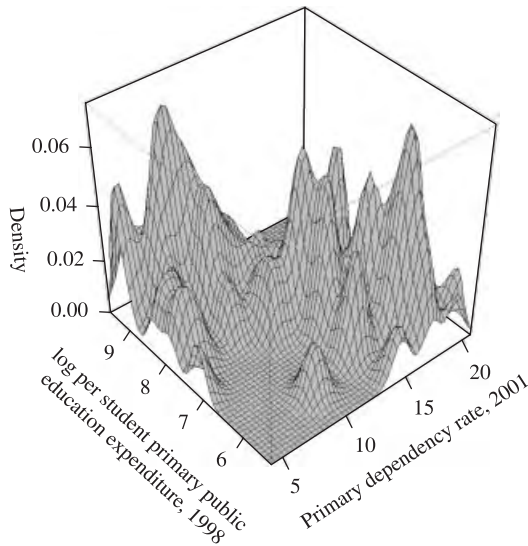
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 6. Repeater Rate and Per Student Public Primary Education Expenditure



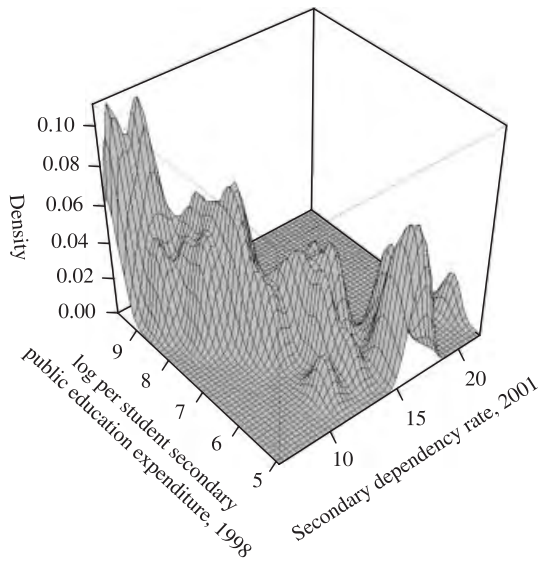
Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 7. Per Student Public Primary Education Expenditure and Primary Dependency Rate



Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

Fig. 8. Per Student Public Secondary Education Expenditure and Secondary Dependency Rate



Source: UNESCO (2004) and Heston, Summers, and Aten (2002).
 Note: Contour is fit by nonparametric kernel smoothing.

on the demand and supply of education. Section III will review recent economics literature on education policies. This has shown that some types of supply-side intervention, such as school construction along with teacher training, and demand-side intervention, such as conditional transfer programs, can be effective in inducing households to allocate their time to schooling. Section IV will look at the estimation issues and point out that more emphasis has been put on randomization of intervention while more skepticism has been put on choosing instruments.

II. FRAMEWORK

Traditional education policies have focused on what is termed the supply-side of schooling. This is directed at reducing the access costs borne by households, such as school expansion, providing safe and cheap transportation, setting a school schedule that accommodates the seasonality of local labor demand, reducing discrimination and intimidation, running schools only for girls when local custom requires. Compulsory education laws can also be considered as reducing *net* access costs, defined by the difference between gross access costs and opportunity costs of not going, as they penalize parents for not sending their children to schools. Another strand of supply-side policies is enhancement of school quality. This includes improvement of school facilities and amenities, improvement of teaching materials, improvement of school curricula, strengthening teacher quality and quantity, and setting up a feedback structure between schools and parents.

These traditional policies generally affect the entire locality, as they are mostly legislative changes decided by education officials and teachers, and they focus on the changes at the school level. Thus it is more sensible to regard these policies as community-level policies.

Because households are heterogeneous, it is not surprising that the results of community-level policy changes are not uniform among households despite the uniformity of implementation. This is well known in the program evaluation literature: household and individual heterogeneity brings about different returns to schooling, thus the participation in particular programs and their outcomes are also different. For example, when a government builds a school in a community, it is generally the rule that children from wealthy households reach higher grades than children from poor households. This can be explained in several ways: the presence of credit constraints on poor households, poor households' insufficient access to local insurance arrangements when misfortunes happen, lower returns to schooling among landless poor households when agricultural innovation rewards the landed, higher opportunity costs of schooling among poor households due to smaller domestic labor supply as mothers work longer outside.

The mixed results of supply-side or community-level policies turned policy-makers' attention to demand-side or household-level policies. This has been especially

notable in Mexico. There, a comprehensive package of policies called Progreso (Programa de Education, Salud y Alimentacion, now called Oportunidades) has been successfully introduced. It provides cash transfers and health and nutritional inputs to the poor that are primarily conditional on school attendance. It has also aimed at enhancing the quality of teachers. Mexico's conditional cash transfer program is similar in spirit to other conditional in-kind transfer programs such as food-for-education that transfer food to attending households, free school meals programs that provide food to households whose children attend schools, and voucher programs that provide subsidies in the form of vouchers that can only be used for school fees.

To understand the effects of these policies formally, one can consider the following simple model.

$$\begin{aligned} & \text{Max}_{\{x_1, l, I\}} u[x_1, -a_g \cdot 1(l > 0)] + u(x_2, 0), & (P1) \\ \text{s.t.} \quad & s = B + wT - \omega l - x_1 - P'I, \\ & x_2 = s + b_i A(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}). \end{aligned}$$

A child lives for two periods, and has the option of going to school for l days out of his/her time endowment of T days. Schooling costs wage w and tuition τ , with $\omega = w + \tau$. Since tuition does not have to be explicitly included so long as leisure does not enter in the utility function, it will be excluded from now on. Schooling, together with an education input vector \mathbf{I} , school quality vector \mathbf{Q} , child characteristics and ability vector \mathbf{C} , household characteristics vector \mathbf{H} determine educational attainment A that will be rewarded in the labor market in the second period. Inputs cost P per unit. The multiplication of attainment A by ability b_i allows either income to vary by individuals with the same attainment, or treat $b_i A$ as human capital that varies for individuals with the same educational inputs. Following Glewwe and Kremer (2005), who introduced the educational attainment function $A = A(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I})$, it is assumed that \mathbf{Q} , \mathbf{C} , and \mathbf{H} are exogenous and \mathbf{I} is endogenous. It is also assumed, based on our field work experience in south India, that there is discrimination at schools against a particular group g that damages utility. The degree of discrimination is given by $a = -a_g$, and it enters utility negatively once schooling is positive, as denoted by an index function $1(l > 0)$. Thus the intertemporal utility function is:

$$U = u(x_1, a) + u(x_2, 0).$$

The parent makes decisions about consumption in the first period x_1 , about schooling l for the child, and about education inputs \mathbf{I} taking into consideration intertemporal budget constraints:

$$\begin{aligned} s &= B + wT - \omega l - x_1 - P'I, \\ x_2 &= s + b_i A(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}). \end{aligned}$$

The first equation is period 1 budget and the second is period 2 budget. Savings s is determined by x_1 , l , \mathbf{I} , and inherited assets B , and second period consumption is determined by s and the educational attainment A of the child. The Lagrangian is:

$$\mathcal{L} = u[x_1, -a_g \cdot 1(l > 0)] + u[B + wT - \omega l - x_1 - \mathbf{P}\mathbf{I} + b_i A(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}), 0] + \lambda_h [B + wT - \omega l - x_1 - \mathbf{P}\mathbf{I}].$$

Kuhn-Tucker first-order conditions, assuming $x_1 > 0$, are:

$$\begin{aligned} \mathcal{L}_{x_1} &= u_1 - u_2 - \lambda_h = 0, \\ \mathcal{L}_l &= -a_g u_a + u_2 [-\omega + b_i A_l] - \lambda_h \omega \leq 0, l \geq 0, l \mathcal{L}_l = 0, \\ \mathcal{L}_{I_k} &= u_2 [-P_k + b_i A_{I_k}] - \lambda_h P_k \leq 0, I_k \geq 0, I_k \mathcal{L}_{I_k} = 0, \end{aligned}$$

u_i indicates partial derivative of u with respect to x_i , $u_a = \partial u / \partial a$, $A_l = \partial A / \partial l$. For a household that chooses $l > 0$, we have:

$$A(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}) = (1 + u_1 - u_2)\omega + (u_a/u_2)a_g > \omega, \quad (1)$$

if the household is credit constrained $u_1 > u_2$, or if there is discrimination $a_g > 0$ and $u_a > 0$. Thus the more the household is credit constrained, and/or the more it suffers from discrimination, the shorter and further away from optimum the child's schooling will be. Similarly, for a household with $I_k > 0$ that is credit constrained,

$$A_{I_k}(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}) = [1 + (u_1 - u_2)/u_2]P_k > P_k, \quad (2)$$

thus education input I_k becomes too small. Equations (1) and (2) can be rewritten to conform with Glewwe and Kremer (2005) as:

$$\begin{aligned} l &= l(\mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{P}, \lambda_h, a_g), \\ I_k &= I_k(\mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{P}, \lambda_h, a_g), \end{aligned} \quad (3)$$

where \mathbf{P} now includes w and τ , with $\partial l / \partial \lambda_h < 0$, $\partial l / \partial a_g < 0$, and $\partial I_k / \partial \lambda_h > 0$. Depending on credit market entitlements λ_h , discrimination at school a_g , child ability \mathbf{C} , tuition τ , and wages w , schooling and education inputs will differ.

As Glewwe and Kremer (2005) note, supply-side interventions can be expressed as variations in \mathbf{P} and \mathbf{Q} . We can also include a variation in a_g . This gives:

$$\begin{aligned} dl &= \frac{\partial l}{\partial \mathbf{Q}'} d\mathbf{Q} + \frac{\partial l}{\partial \mathbf{P}'} d\mathbf{P} + \frac{\partial l}{\partial a_g} da_g, \\ dI_k &= \frac{\partial I_k}{\partial \mathbf{Q}'} d\mathbf{Q} + \frac{\partial I_k}{\partial \mathbf{P}'} d\mathbf{P} + \frac{\partial I_k}{\partial a_g} da_g, \end{aligned}$$

for $k = 1, \dots, K$. Similarly, the effects of demand-side interventions that entail transfers t_h to household h that reduce credit constraints λ_h , and changes in local prices \mathbf{P} can be expressed as:

$$dl = \frac{\partial l}{\partial \mathbf{P}'} d\mathbf{P} + \frac{\partial l}{\partial \lambda_h} \frac{\partial \lambda_h}{\partial t_h} dt_h,$$

$$dI_k = \frac{\partial I_k}{\partial \mathbf{P}'} d\mathbf{P} + \frac{\partial I_k}{\partial \lambda_h} \frac{\partial \lambda_h}{\partial t_h} dt_h,$$

for $k = 1, \dots, K$, where we denoted transfers to household h as t_h and its impact on credit constraint as $\partial \lambda_h / \partial t_h$. For example, supply-side interventions that weaken discrimination increase l . Demand-side policies that increase household income increase l and I . Demand-side policies that relax credit constraints of households increase l and I .

If we rewrite equation (1) as

$$\frac{1}{1 + u_1 - u_2} [A_l(l, \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}) - (u_a/u_2)a_g] = \omega,$$

or

$$\delta(l | \mathbf{Q}, \mathbf{C}, \mathbf{H}, \mathbf{I}, \lambda_h, a_g) = \omega, \text{ where } \delta(l | \cdot) = \frac{1}{1 + u_1 - u_2} [A_l(\cdot) - (u_a/u_2)a_g],$$

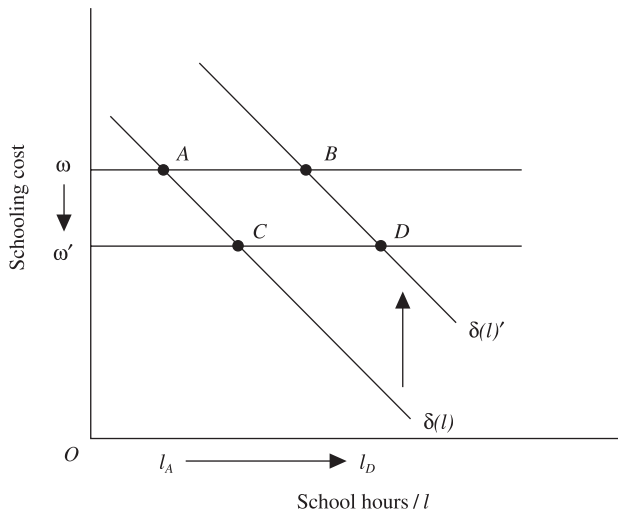
the LHS can be understood as net marginal benefits of schooling. We can depict choice l as an equality between net marginal benefits and marginal costs. A cut in tuition shifts down the marginal cost curve. An increase in transfer t_h or a decrease in credit constraint λ_h results in an increase in δ , or an upward shift in the net marginal benefit curve in Figure 9. A comprehensive intervention, such as Progresa in Mexico, can be considered as simultaneous demand- and supply-side interventions that shift both the ω and δ curves.

III. RECENT EXPERIENCES

Given that education is considered important in development, there have long been policies to promote it. Government initiatives to supply schooling opportunities, termed in this paper as supply-side policies, are seen as traditional policies. In recent years, some countries have started to assist the demand of households for education, what have been known as the demand-side policies.¹ This section will review the econometric studies on them, as summarized in Table 1.

¹ Progresa-like programs include: Bolsa Familia (Bolsa Escola, Brazil), Familias en Accion, PACES, Columbia, Programa A de Familias (Honduras), PATH (Jamaica), Red de Proteccion Social (Nicaragua), and Social Solidarity Fund (Turkey). See Skoufias (2005); Rawlings and Rubio (2005); Das, Do, and Özler (2005).

Fig. 9. Effects of Pro-Schooling Interventions



Source: UNESCO (2004) and Heston, Summers, and Aten (2002).

Note: Contour is fit by nonparametric kernel smoothing.

A. Demand-Side Policies

Most of the recent demand-side policies take the form of conditional transfer programs. These programs that disburse resources when certain conditions, set by the government, are met. Examples of conditions include school attendance rate, household income, and location of residence. These conditions serve the dual role of providing assistance and as a self-targeting device that directs resources toward the poor.

1. *Progresa (Oportunidades)*

The conditional transfer program that has attracted most attention is Progresa in Mexico. The reasons for such attention are the magnitude of geographical and household coverage, the novelty of the idea, and the program's well-established effectiveness as suggested by numerous academic articles that have been produced mainly in collaboration with IFPRI (International Food Policy Research Institute) economists.

The coverage of the program was phased in geographically due to logistical inability of across-the-board implementation throughout the country. Following its inception in 1997, Skoufias (2005) notes that by 2000, the program included nearly 2.6 million families in 72,345 localities in all 31 states, constituting around 40% of all rural families and one-ninth of all families in Mexico. The program was the flagship of the government's poverty alleviation efforts, and its annual budget in 1999 was around US\$777 million, which was about 20% of the federal poverty alleviation budget or 0.2% of GDP.

TABLE 1

Summary of Recent Demand- and Supply-Side Interventions

Name	Country	Period	Policies	Estimated Impacts
Demand-side policies:				
Progresa	Mexico	1997–2000	Conditional cash transfers to mothers, health care, etc	Added 0.7 year of schooling, no change in achievements, lower morbidity
FFE	Bangladesh	1995–96	Conditional in-kind transfers (rice)	Schooling probabilities increased 17% for boys and 16% for girls
School meals	Kenya	2000–2002	Conditional in-kind transfers (breakfast)	Schooling probabilities increased by 30% points, score improvements
PACES	Columbia	1991–present	Conditional in-kind transfers (vouchers)	Secondary completion probabilities increased by 25%, lower rates of repetition, higher scores
Supply-side policies:				
SD INPRES	Indonesia	1973–78	School expansion in lower enrollment areas, increase of teachers	Added 0.12 and 0.19 year of schooling, greater labor participation, wage increase from 1.5% to 2.7%, wage decrease of older cohorts
Compulsory schooling	Taiwan	1968–86	Free tuition and textbooks, increased teachers and schools	Schooling increased 0.4 year for boys and 0.25 year for girls
Operation Blackboard	India	1987–94	Teacher redistribution from large to small schools	Increased primary completion rates by 2% points for boys and 4% points for girls, mostly from poor households
Balsakhi	India	2001–3	Remedy schools using unlicensed teachers	Test scores increased 0.15 standard deviations in the first year and 0.25 standard deviations in the second year

Source: Compiled from various studies on these projects. See References at the end of this paper.

In a nutshell, Progresa gave cash to poor households conditional on the previous bimonthly and cumulative attendance rates being no less than 85%. Eligibility was set using information from a preprogram census designed specifically for the program. The amount of cash was worth around 20% of the average preprogram consumption of recipient households. Progresa also consolidated health and nutrition support with eligibility, and made eligibility conditional on complying with timely health checks at local clinics and take up of nutritional supplements. It also augmented the supply side by retraining teachers and health care workers.

Evaluation of the program found Progresa to be effective, that it added 0.72 years for girls, 0.64 years for boys. With an average of 6.2 years of completed schooling in the population below 18, Progresa extended the schooling of poor rural children in Mexico by at least 10%.²

Targeting is considered to be more precise (when compared with census information) at the lower end of poverty, but not at relatively well-off households (Behrman and Todd 1999). As the sum that Progresa transferred was more for older children and girls, it had more effects on secondary school enrollment of girls (Schultz 2004). There was also a fall in the rate of child labor participation (Parker and Skoufias 2000), implying some substitutability between schooling and work. Work incentives (hours) for adults were not affected, and children's time use indicates that their work hours decreased for boys, while leisure hours decreased for girls (Parker and Skoufias 2000). However, the qualitative level of educational attainment, as measured by test score, showed no significant changes (Schultz 2004).³

2. *Food-for-education, free school meals*

Food-For-Education (FFE) programs are popular policies in low-income countries. These programs disburse food, mostly staples, to households whose children's school attendance rate is above a set standard. FFE is a type of conditional (in-kind) transfer program. Most FFE schools do not charge schooling fees.

Ravallion and Wodon (2000) evaluate the impact on school attendance and child work hours of an FFE program implemented in Bangladesh. Every month, rice equivalent to about 13% of a boy's average monthly earnings was provided if attendance is above 85%. Eligibility for rice was confirmed by attendance records submitted by school headmasters who computed the necessary amount of rice. Each school submitted requests to the FFE office at the local district headquarters. There was no explicit targeting.

² It has been pointed out that the program was highly cost effective relative to other interventions (Coady and Parker 2002).

³ The health and nutritional portion of the program was also successful. There are fewer child stunting (Behrman and Hoddinot 2001) and adult morbidity (Gertler and Boyce 2001). One observes better diet with more vegetables and fruits (Hoddinot and Skoufias 2004).

Ravallion and Wodon (2000) find that, based on data for 1995–96, at the sample mean, an annual transfer of 100 kg of rice increased schooling probabilities by 17% points for boys, and 16% points for girls. There was a limited substitutability between schooling and leisure, as schooling did not completely replace labor, implying that leisure had fallen. Average foregone income was about 19% of transfer benefits (about Tk 119 a month) for a household with one boy and one girl. While the results are indicative of a successful program, the authors raise a concern that the overcrowding of classrooms could lower school quality were it not accompanied with supply-side measures.

Programs providing free school meals are similar in spirit to FFE programs. Vermeersch (2003) evaluates the Kenyan program of 2000–2002 for preschoolers. The program offered a fully subsidized, in-school breakfast. This automatically satisfied the attendance conditionality. The author finds increased participation for both children who were already attending school and for those who were not attending, and the effect on the former was larger. It was also noted that the test scores for children whose teachers were better trained at the onset of the program improved. However, there was no evident impact on the cognitive abilities and anthropometrics of the children. The latter suggests the possibility of intrahousehold substitution of nutritional intake that partially offsets the impact on targeted preschoolers, which was tested in Jacoby (2002).

3. *Vouchers*

Voucher programs are also conditional (in-kind) transfer programs. As summarized in Gauri and Vawda (2004), these programs focus on increasing competition among schools to promote public and private school efficiency. Implicit in them is the idea that private schools are more responsive to parental demands. Given the right to choose schools, parents effectively decide which schools will be subsidized, leading to increased efficiency. The idea is that chosen schools are better schools which are rewarded with increased school enrollment and subsidy.

Households can use vouchers as an instrument for financing private school tuition. Unlike Progresa or FFEs, there are no extra subsidies other than payments used for tuition. The ability of voucher programs to enhance the quality of schooling depends on the ability of parents to monitor schools, and the ability of schools to respond. Moreover, a necessary precondition is the existence of an active private education sector with school enrollment being open to all. Other typical fears about voucher programs include bloated school size, and polarization as schools choose students to earn better performance in terms of scores while parents choose schools where they expect human capital externalities or peer effects among students.

Angrist *et al.* (2002) evaluate the Columbian voucher program in 1995 and 1997 in selected urban areas. From 1991 onward, the program covered over 125,000 students and was geographically targeted at low enrollment areas. Vouchers were randomly assigned by lottery to already-enrolled secondary students. Vouchers

covered about half of tuition (in 1995) and could be accepted only by private secondary schools whose tuition fees were in the bottom-half of the entire private school system. Vouchers were renewable conditional on promotion. Angrist *et al.* (2002) find a 25% point increase in secondary school completion, lower rates of repetition, and higher test scores equivalent to an additional year of schooling. They also find increased educational spending by the lottery-winner households.⁴ The findings suggest that vouchers can be a promising tool to enhance school quality, if public schools have lower quality, and if private schools are well developed.

B. *Supply-Side Policies*

Supply-side policies are the traditional policies whose aim has been to provide schooling opportunities. Experience with these policies has been long and numerous. This subsection looks mainly at recent studies that employ rigorous statistical methodologies.

1. *School expansion*

At the initial stages of development, budgetary and teacher supply constraints limit the geographical scope for building schools. As the governments expand the schooling systems to unserved areas, this creates a natural experiment for measuring the effects of school availability in a neighborhood.

Duflo (2001) studies the impact of Indonesian primary school expansion between 1973 and 1978 by examining pre-program and post-program cohorts. During the period, the Indonesian government constructed more than 61,000 primary schools, and the number of teachers per district increased from 1,530 to 2,082. The government built more schools in low enrollment rate areas. The effect was an increased enrollment rate among 7–12-year-olds from 69% to 83%. Additional schooling ranged from 0.12 to 0.19 years, and wage increased from 1.5% to 2.7% among the graduates (Duflo 2001). The expansion also had effects on older cohorts, increased labor participation in the formal sector and reduced wages (Duflo 2004).

Another similar intervention was implemented in Mozambique (1991–94), where provincial directorates of education allocated funds to their districts for school expansion (Handa 2002). The allocation was uniform in the sense that it was not affected by observable village-level variables. Using census data, the study finds increased schooling for boys, but not for girls.

2. *Compulsory schooling*

Compulsory schooling in Taiwan beginning from 1968 extended free tuition and textbooks to junior high schools. This was backed by a significant increase in education budget during 1967/68–69/70, which raised real expenditures by 20.5%,

⁴ They also note the cost-effectiveness of the program: three-year society (government + household) costs US\$195/student, while the mean wage increased by between US\$36–300 a year.

12.7%, 16.8%, respectively in each year. The series of expenditure growth corresponded roughly to a 1% increase in education's share of GDP. The number of secondary schools increased by 33.1% in the northwest, 58.1% in the southwest, and the number of classes increased by 85% in the country as a whole. The number of teachers was also significantly increased which enabled a rise, rather than a drop, in teacher-student ratio despite the sharp rise in enrollment. Spohr (2003) finds the increase in schooling between 1968 and 1986 was 0.4 years for boys and 0.25 years for girls (for both below 12-year-olds in 1968), compared to no reform.

3. *Teacher reallocation*

Operation Blackboard in India (1987–94) originally was intended to augment the number of teachers in small schools. However, what happened was a redistribution of teachers from large schools to small schools, as the total number of teachers remained roughly the same. This provided a natural experiment on measuring the impact of teacher reallocation: from large to small schools. Using state-level data and the number of potentially affected schools, Chinn (2005) finds that the program increased the primary school completion rate for 4% points for girls and 2% points for boys, mainly for those from poor households.

4. *Remedial schooling*

A remedial schooling program in northern India (2001–3), studied by Banerjee *et al.* (2005), sets up remedial schools for urban underachievers using unlicensed teachers. Data were collected in the cities of Mumbai and Vadodara. Preliminary evaluation finds the program is effective in increasing test scores, particularly for underachievers. It also notes the cost-effectiveness, as it is 12–16 times cheaper than hiring licensed teachers in schools.

5. *Providing second chances: The MVF*

A South Indian non-government organization (NGO), the Mamidipudi Venkatarangaiya Foundation (MVF),⁵ sets up residential bridge camps (RBCs) for out-of-school children and trains them for 6 to 12 months before sending them to formal schools. It also promotes schooling from the regular age of five.

The strategy of the MVF programs can be summarized roughly as threefold. One aspect is appraisal. Working with local volunteers, MVF activists regularly visit households whose children are child laborers to persuade the parents to send their children to school or to the nearest RBC, whichever is found to be appropriate for the child. They do not explicitly give away any resources such as cash or food. They

⁵ The MVF was founded in 1981 and began its mission of eradicating child labor in 1991 in the state of Andhra Pradesh. In recent years, it has been active in more than 6,000 villages in 137 mandals of 13 districts in AP. There are currently about 4,000 local activists working with the MVF.

also run occasional rallies and street plays condemning child labor and its employers. A second aspect is a personalized catching-up process. If the child needs remedial education before attending a formal school, MVF puts the child in an RBC and trains him/her in a class assigned on the basis of learning ability. The cost of enrollment in an RBC is free, and local youths with high school diplomas or higher, who do not necessarily have official teacher licenses, teach classes of between 20 and 25 students. Third is monitoring and community involvement. The local MVF activists monitor school attendance almost daily. A network of youth groups and village leaders alerts MVF if they find households with child laborers. Collaboration with these existing networks, which is crucial for obtaining information and village-wide support, also helps MVF recruit volunteers when it expands into a new area.

Since 1991, the MVF removed over 250,000 children from work and sent them to schools full time (Mukherjee, Sarkar, and Sudarshan 2005). MVF works with local networks of officials, PRIs (Panchayati Raj Institutions),⁶ SECs (School Education Committees), village youth organizations, community-based organizations (CBOs), the BKVV (Bala Karmika Vimochana Vedika),⁷ and CRPFs (Child Rights Protection Forums), to implement their policies more effectively. Despite being a major organization running programs in a huge number of villages to eradicate child labor and sending a great number of children to school, the MVF has never had time and resources to undertake a statistically rigorous evaluation of its work, in part due to the overwhelming prevalence of child labor in India, and in part to the lack of funds. A randomized experiment involving 480 households in 32 villages (17 treatment and 15 control villages) is under way in the backward area of Andhra Pradesh (AP) which is expected to provide insights on the need for resource transfers, constraint relaxation, and intertemporal changes in parental behavior.

C. *Interpretations*

1. *Conditional transfers*

The focus of conditional transfer programs is on relaxing credit constraints on poor households, rather than being concerned with other constraints induced by other market failures. This follows since, with other market failures such as limited labor substitutability between hired and family labor, giving more money now to the household would not solve underschooling. If it is a lack of finances that is to be resolved, it is not clear from theory that households should be subsidized rather than given loans. It also does not immediately follow why a subsidy has to be conditional on school attendance when credit access is granted, because a household should optimally allocate resources when there is no credit constraint.

⁶ A panchayat is an elected village council that acts as a conduit between the local government and villagers. The leader of panchayats is called the Sarpanch.

⁷ Teachers' forum against child labor, or a teacher group working against child labor.

The choice of providing subsidy rather than loans needs to be understood in the context of political economy and the severity of market failures. In a democracy where the majority of voters are poor, it is not difficult to see why the leadership chooses subsidy programs for the poor. Even in wealthier countries, the choice might fall to universal subsidization like voucher programs rather than granting loans.

The severity of market failures involves the enforcement and monitoring of contracts. Apart from political economy arguments, it may be technically infeasible or too costly to implement school loan programs geared toward low-income households. The enforceability of contracts becomes difficult in low-income countries because the government does not usually keep and update household residency records. The notion of targeting the poor makes enforceability more problematic because they tend to be landless and are thus more mobile across regions. Another source of difficulty in enforcement is risks. Poor households are more prone to health shocks which may hamper human capital investment in children, either a shock to the children themselves or to family members that require children to quit school to contribute to household income. This is parallel to the default on loan contracts in banking theory under an assumption that the poor face greater risks, and because the poor have few assets, a government may not be able to collect their loans. Given such difficulties arising from market failures, a government may simply opt for subsidies.

Given that conditional transfer programs show successful results while unconditional transfer programs generally do not, one needs to have more realistic models to understand the successes based on conditionality. Conditionality can serve two purposes: it can be a self-targeting device and a commitment device, and these can induce changes in intrahousehold resource allocation. Conditionality requiring frequent and timely visits to field offices or local health clinics can be more painful for the wealthy. This will help self-targeting of school subsidies to lower income households, an argument also made for workfare programs. However, the empirical evidence from Progresa shows that targeting is less well achieved at the household level (Behrman and Todd 1999).

The commitment aspect of conditionality involves direct transfers to mothers and conditionality based on school attendance. This is done in the hope that such conditionality will work against the misuse of transfers, such as spending on alcohol and cigarettes by male household members. Clearly the economic models to justify this contention formally will require collective models, as such a notion of conditionality is unnecessary and thus cannot be understood in unitary models. Or, if the unitary model is correct, one needs to incorporate the problem of self-control (Mullainathan 2006).

As popular as these programs are, however, there is only a limited number of theoretical works that examine conditional transfer programs using a non-unitary framework.⁸ A directly applicable model of conditional cash transfer in education

⁸ The literature on targeting, or income maintenance programs, typically assumes unitary models.

is studied by Martinelli and Parker (2003a) who explicitly examine the effects of targeted transfers conditional on schooling on households with fixed bargaining powers. Holding the bargaining powers constant, a marginal conditional transfer is better (than an unconditional transfer) for a parent with a binding bequest constraint if the parent's own consumption and that of child are complements, or they are substitutes but he/she cares for the child more than the other parent does. A marginal conditional transfer gives lower welfare to all members if the household is bequest-unconstrained, as it leads to overinvestment in education, lower net returns, hence lower wealth.

Another model that can provide insights on the role of conditionality is the separate-spheres bargaining model of Lundberg and Pollak (1993).⁹ The underlying model is cooperative in nature; however, its distinctive character is that it endogenizes the threat points by assuming they correspond to the outcomes of noncooperative games within a household (not the outside options as assumed in Manser and Brown 1980 or MacElroy and Horney 1981) when the couple decides not to cooperate. Separate-spheres bargaining model is particularly useful for areas like rural India where divorce is a practical impossibility.

Endogenization of threat points naturally explains that the recipient of transfers matters in the bargained outcomes. If some of the household public goods, such as the welfare of children and security of the household, are not provided jointly, the noncooperative game leads to lower-than-optimal provision, which decreases the threat point utility. It is important to recognize that the less-than-optimal provision of public goods by female members, child rearing, for example, also has an effect on the utility of males. We should consider not just the direct effects of transfers on the mother's utility, but also the cross-effects on the father's utility. In the context of Progresa, it is likely that directed transfers to mothers increase their threat points and induce relative intrahousehold allocation in favor of mothers *vis-à-vis* fathers. However, the utility level of fathers does not necessarily decrease if total household resources are increased by the transfer. In light of this, transfers, rather than loans, may be advisable for inducing fathers to cooperate with pro-schooling decisions, as they compensate for the submergence of the father's relative bargaining power.

The models with endogenous threat points suggest that it is important to know empirically how school attendance conditionality affects the bargaining power of mothers and fathers. Martinelli and Parker (2003b, tab. 5) on Progresa is the only study that has dealt with this problem. They show that Progresa increased the share of expenditure on women's and boy's clothing and on girl's clothing, but only the last was (weakly) significant. It is also worth pointing out that the successes of conditional transfer programs are also consistent with models that do not assume neoclassical rationality in individuals.

⁹ The extension of Chen and Woolley (2001) incorporates caring (*à la* Becker) members, and the generalized Nash bargaining.

Apart from conditionality, another related point of interest on transfers is intra-household substitution of resources. Using a term in Jacoby (2002), a “flypaper effect” may be absent, or worse, a transfer targeted to a child may induce intrahousehold substitution of resources away from the child. This happens if household member preference is altruistic, because they try to smooth (consumption) benefits across individuals. While Jacoby (2002) finds such substitution effects were minimal in his data on the Philippines, substitution is an empirical possibility that should be considered when implementing targeted transfers. In South Africa, Duflo (2003) used difference-in-differences estimates to show that the height-for-age of girls increased when pensions are received by women. This suggests that there may be an empowerment effect in cash transfers directed to mothers or grandmothers, as suggested by Lundberg and Pollak (1993). If so, the results in Martinelli and Parker (2003b, tab. 5) suggest that a negative substitution effect of conditionality on child consumption could have canceled out the empowerment effect of transfers.

When we focus on the welfare of the individual, not of the household, it is unclear how to evaluate transfer policies. As long as there are income and price effects caused by conditional transfers within households and corresponding substitutions of resources among household members, an outcome of longer and continued schooling for some members may result in lower utilities for other members. For example, longer schooling of a child may be offset by shorter home teaching by parents, smaller home education inputs, longer work hours for other members, or loss in leisure hours for targeted members. It is only at the household level without addressing which member’s welfare the transfer program is to improve that we can claim conditionality is beneficial to poor households. After all, conditionality in transfer programs assumes some form of conflict of interests, but evaluation of such programs tends to ignore the possible negative effects on other members and focuses only on the education outcomes of particular, mostly the targeted, members. Such details are not always well scripted in policies, and negligence of individual welfare can result in the lower welfare of targeted and/or other members.

2. *Supply-side policies*

It is not easy to interpret the effects of supply-side policies because most of the studies are based on large, census-like data with little information on household background. Therefore, without specific surveys that are designed to evaluate the programs, supply-side policy evaluation becomes an estimation of unconditional averages of each cohort.

However, some studies have revealed the heterogeneous impacts of policies on different quantiles. By running quantile regressions, Bedi and Edwards (2002) show that quality enhancement in Honduras has resulted in greater returns for low-income earners. Banerjee *et al.* (2005) also show that remedial schooling in urban India is more effective for students with lower grades. Chinn’s analysis (2005) suggests that teacher reallocation to single-teacher schools results in greater primary school

completion rates for children from poorer households. These studies are suggestive but do not tell why these programs worked and what lessons we can learn when applying similar programs to other regions.

3. *MVF interventions*

There are a number of reasons for the effectiveness of MVF programs. First is a lowered access cost. Examples of this include assistance with paperwork upon admission and continuation of formal schools, and avoidance of discrimination based on caste and sex by continued monitoring on school attendance. Second is giving dropouts second chances. This is done through reductions in transition costs by attending bridge camps, and continuation through camp residency that keeps parents from changing their minds. Third is enhancement of learning capacity by investing in physical human capital at RBCs, and by personalized catch-up processes. The latter is particularly important for children with longer absences from schools. Fourth is implicit subsidization of complying households. The claim that MVF does not give any cash is true, but there are implicit cash transfers made through the negotiation of debt write-offs for bonded laborers and the release of children from workplaces. In addition, RBCs are free of charge and parents are exempt from paying for tuition and boarding fees. And last, learning and re-discovering by parents and children of the value of education certainly plays some role.

These points, along with suffering little loss of resources from corruption, are made possible by the MVF's working principles. Corruption is averted by the centrally administered budgeting process. MVF activists are motivated because they not only care about education but also because of the incentives of job promotion. MVF programs of deterring discrimination through attentive monitoring, of hiring local youth, of persuading of creditors, employers, and parents to reject child labor, and of identifying targeted households are all done not only due to the efforts of the MVF but also by working with local leaders and youth groups who have better local networks and authority.

IV. RECENT ESTIMATION STRATEGIES

The necessary condition for the linear regression estimator to be consistent and unbiased is zero correlation between regressors and disturbance terms.¹⁰ This simple requirement is rarely satisfied in the observational data that economists mostly work with. The source of the problem can be stated in two ways: either it is because economists cannot collect a complete set of information to be added in regressions, or it is because economists find it difficult to implement controlled experiments to ensure orthogonality between regressors and disturbances.

¹⁰ Sufficiency requires the correct specification of functional forms and distributional assumptions.

In the past (and present), the problem of non-orthogonal regressors automatically led researchers to seek instrumental variables, often in the form of natural experiments (Rosenzweig and Wolpin 2000). The problem is particularly severe when examining policy impacts if participation in policies is voluntary, and if policy-makers have (and they mostly do have) a choice of where to implement policies. Acknowledging the difficulty, some economists have started to design randomized experiments specifically for the hypothesis under examination. This new way of estimation is backed not just by the notion of methodological orthodoxy in scientific research, but also because of increased usage of multipurpose survey data which led to a loss in novelty of using microeconomic data, and the increased collaboration between researchers and program implementers. This section discusses recent trends in estimation strategies related to educational attainment.

A. Randomization

The most prominent feature of recent estimation strategies of policy impacts is the use of explicit randomization in policy implementation. One of the leading examples is Miguel and Kremer (2004) who evaluate randomized deworming intervention in Kenya. The program was run by a Dutch NGO and, due to their administrative and logistical constraints, they could select only a fraction of schools in each year for intervention. They did so through randomization, making program assignment orthogonal to disturbance terms by definition. A similar design was also used in Mexico's Progresa, except that randomization was at the village level and targeting encompassed much broader areas. Randomization seemed to have worked well at the village level, but not so well at the household level (Behrman and Todd 1999).

With randomized assignments, one can consistently estimate the causal effects of policies.

$$\text{Causal effects on } A = \varepsilon[A \mid \textit{treated}] - \varepsilon[A \mid \textit{control}] = \varepsilon[A \mid \mathbf{P}_{\textit{treated}}] - \varepsilon[A \mid \mathbf{P}_{\textit{control}}].$$

Under the most widely used assumptions, effects of a policy \mathbf{P} are given as linear:

$$\varepsilon[A \mid \mathbf{P}_{\textit{treated}}] = a_{0i} + a_t + \mathbf{b}'\mathbf{P}_{\textit{treated}},$$

where a_{0i} is a heterogeneous intercept for individual i and a_t is a time effect. The virtue of randomized experiments is that, by definition, the estimated parameters \mathbf{b} are guaranteed to be consistent. One does not need any instruments. Since $\mathbf{P}_{\textit{treated}}$ and $\mathbf{P}_{\textit{control}}$ are assigned randomly, heterogeneity expressed in a_{0i} or a_t does not affect the consistency of \mathbf{b} .

B. Panel Data Estimation

When $\mathbf{P}_{\textit{treated}}$ and $\mathbf{P}_{\textit{control}}$ are not randomly assigned, we would have to control for the unobserved heterogeneity. The oft used method, where panel data are available,

is to use the difference of “within” variation across groups. To get rid of individual intercept terms, one takes a first difference:

$$\varepsilon[\Delta A \mid \Delta \mathbf{P}_{treated}] = \Delta a_t + \mathbf{b}' \Delta \mathbf{P}_{treated}.$$

To further take away common changes in A , without a loss of generality, we assume that $\Delta \mathbf{P}$ is comprised of two parts, $\Delta \mathbf{P} = \Delta \mathbf{P}_0 + \Delta \mathbf{P}_1$ where $\Delta \mathbf{P}_0$ is a common component vector both for the treated and the control, and $\Delta \mathbf{P}_1$ is a unique component vector for the treated. The crucial identification assumption is that the effects of $\Delta \mathbf{P}_0$ are common for all individuals, or, if not, they are uncorrelated with $\Delta \mathbf{P}_1$. We take a difference of differences to get only the policy-induced changes in ΔA :

$$\varepsilon[\Delta A \mid \Delta \mathbf{P}_{treated}] - \varepsilon[\Delta A \mid \Delta \mathbf{P}_{control}] = \mathbf{b}_1' \Delta \mathbf{P}_1.$$

This provides consistent estimates under fairly reasonable set of assumptions, namely, heterogeneity is additive and not time-varying. If one worries about the presence of individual fixed-growth heterogeneity $a_{1i,t}$ which results in a_{1i} in the first-differenced equation, then one can run a difference-in-difference-in-differences or triple-difference estimation, which requires at least three period observations. However, even the triple-difference estimator is a partial solution for the omitted variable problem, as one cannot get rid of the biases arising from the time-varying omitted variables that change at varying rates through time.¹¹

C. *Natural Experiments and Regression Discontinuity Design*

Another source of orthogonal regressors is natural experiments. The studies with this identification strategy compare the effects under different exposure to the natural experiments, which are deemed to be random by nature. Most of these studies use large scale, census-like data to get as much variation in exposure as possible, and typically estimate cohort- or region-specific effects by comparing over cohorts/regions with different exposure to experiments. Duflo (2001) uses an Indonesian census to examine the effects of nation-wide school construction efforts on labor market participation and wages, by comparing pre- and post-intervention cohorts. A similar design has been used in Spohr (2003) where pre- and post-intervention cohorts are compared to examine the impacts of compulsory schooling laws on attendance. Cross-sectional comparison was used in Chinn (2005) in assessing teacher reallocation policy in India, by looking at the number of potentially targeted schools in each state.

The sources of data these studies use are censuses and administrative records on intervention rules. If the intervention rules are based only on observables, or the

¹¹ That is not the only concern. If the errors are serially correlated, Bertrand, Duflo, and Mullainathan (2003) point out that there may be a substantial bias in standard errors in difference-in-differences estimators if one does not use one of their proposed remedies, including robust standard errors.

interventions are such that allows us to consider them as uniform, one can convincingly show that intervention is orthogonal to disturbance terms. One example of the former is that the assignment determined at the central government level and followed rigorously by the local government is orthogonal to individual-level variables, because they are unobservable to the central government, thus can be used to identify the treatment effects (Ravallion and Wodon 2000). This is a rare case that is justified to use the assumption of *selection on observables*. Ravallion and Wodon (2000) use the village-level intervention indicator to explain the exposure of individuals to a program. A uniform-intervention example is well-enforced policies such as the compulsory schooling law in Taiwan.

Another large number of studies use weather variables as the source of natural experiments. These are random in nature; however, in order to use spatial and temporal variations in weather to identify transitory and permanent shocks, one must use the correct measure of net income, and also include local prices, as weather can affect locale-specific prices, hence choices of households (Rosenzweig and Wolpin 2000).¹²

The regression discontinuity design is similar to natural experiments. This exploits the sharp change in policy variables under some predetermined policy rules. A very good example is Maimonides' rule of Israel used in Angrist and Lavy (1999). The rule dictates class size to be below 40. This creates a sharp difference in class size between a potential pool of students of 40 and 41. Under the condition that no one intentionally migrates from one pool to another, Maimonides' rule creates an exogenous variation in class size which will serve in identifying the effect of class size on education outcomes. Regression discontinuity design provides an exogenous variation similar to natural experiments when there is a threshold in certain programs and there is no intentional migration from one regime to another.

D. *Propensity Score Matching*

A popular estimation strategy under a stronger assumption (on a shakier foundation) of selection on observables is the propensity score matching estimators. If this assumption is correct, then one can consistently estimate the participation probability of the treatment group. Using the estimated parameters, one can compute the participation probabilities of the control group. Propensity score matching compares the individuals with similar probabilities: they are considered to be the "same" individuals and they give the factual and counterfactual outcomes, with treatment and without treatment of the "same" individual.

¹² Otherwise, if we are to use gross income, one must assume perfect substitutability of family and hired labor (which is often rejected). If we do not include prices, one has to assume perfect market integration of all agricultural inputs, including labor, for all regions, or perfect separability of consumption from production.

The pioneers of the method, however, point out that precision in matching is crucial for obtaining consistent estimates, which in turn requires sufficiently close propensity scores between the treated and the controls, in other words, a large, cross-sectional data set with rich individual information (Heckman, Ichimura, and Todd 1997). If one is left with a large data set with detailed individual information, then propensity score matching estimator is a reasonable choice.

E. *Instrumental Variables*

In econometrics courses, we have been taught that when no explicit randomization is available, one must seek for instrumental variables; and it is argued that one should satisfy three criteria for treating variables as instruments.

The first is relevance. This is the requirement of a strong correlation between endogenous regressors and instruments. Without strong instruments, standard errors of estimates will follow nonstandard distributions (Staiger and Stock 1997). But this is a relatively innocent requirement because one can demonstrate the strength of correlation, say, with F statistics (Stock, Wright, and Yogo 2002).

The second criterion is validity. This is the requirement of a lack of correlation between instruments and disturbance terms. Since the disturbance terms are not directly observable, it is difficult to show if validity holds, and there is a possibility of subjective and unwarranted use of IVs in estimations. It is, however, not impossible to show the validity under a proper context.

This leads to the third criterion of refutability (Angrist and Krueger 1990). This is the requirement of having an ability to show the validity of instruments using subpopulations. One can satisfy the refutability criterion if one can show testable predictions on subpopulations where “treatment effects” should not be observed.

For example, if an experience of military service affects wages but not draft eligibility, then the proposed instrument for military service, draft eligibility, should not affect the wages for cohorts who did not experience the draft due to the end of war. Another example is nonexistence of birth month effects on a college graduate’s earnings if the birth month is a valid IV for high school compulsory schooling in the earnings regressions. This follows because if there is no birth month effect for college graduates, then, by an analogy, there should not be any birth month effects for high school students in the absence of compulsory schooling laws for high school students. Or better, one can directly test birth month effects on high school dropouts in the state with no compulsory schooling laws to see if instrument validity is refuted. Refutability in the program evaluation context is that, if an instrument explains the participation in the programs, there needs to be a refuting possibility that can examine whether IV is correlated with outcome variables of interest in the nontargeted (control) group. If one can examine such a correlation, the study satisfies refutability, and if it is indeed uncorrelated, the IV is likely to satisfy the validity requirement.

In some programs, such as the effects of compulsory schooling laws studied by Angrist and Krueger (1990), eligibility to participate in the program, or the date of birth, is considered to be randomly assigned and is used as an IV. Imbens and Angrist (1994) show how to estimate the causal effects on potentially eligible subpopulations, an estimator known as the local average treatment effects (LATE) or Wald estimator. Provided that the instruments are valid, LATE gives a simple and powerful way to estimate the causal effects. Because LATE uses instruments which do not require a perfect correlation with participation, it has an advantage that it can incorporate targeting errors.

Not having refuted the validity is particularly damaging in the program evaluation context, where the omitted variables problem is considered to be severe. In other words, one of the main thrusts of randomization can be considered as a skepticism or a pessimism over the possibility of not refuting the instrument validity. However, as shown below, if eligibility to participate in a program is randomly assigned, then eligibility is a valid instrument for program participation. The reason why some researchers only use OLS in randomized trials even when targeting errors are not negligible, when IV estimator has supposedly an advantage over OLS in this regard, is that LATE identifies ATE on unidentified subpopulation (Rosenzweig and Wolpin 2000).

F. *Randomized Experiments and Observational Data*

There is no doubt that a randomized experiment is preferred over collecting observational data in estimating causal impacts, because random assignment ensures consistency of estimates. Thus, whenever feasible, it is always preferable to randomize program assignment when one wants to evaluate causal effects.

Randomization is not free of flaws, however. As Basu (2005) notes, a randomized trial reveals a causal relationship that is specific to the area under the trial. The absence of “external validity,” to use the psychometrics terms, or in econometrics terms, the absence of out-of-sample predictions on samples drawn from a potentially different population, limits the usefulness of estimated results. This is true, but as Duflo (2005) points out, this does not diminish its important role in falsifying a theory. In fact, this is exactly the way it is with observational data: if one finds some interesting (IV estimation) results in a data set from a particular area, then, by the same token as the “external validity” argument, it is naïve to expect to find a similar result in other data sets that are taken from different areas with potentially distinct populations.

If the evidence cannot be explained in the existing theories, a new theory is called for, which will serve as a new base for designing a randomized trial for further testing. It is only through fruitful correspondence between a theory and randomized trials that we can learn about the reality; and through forming a norm, it will ultimately affect the policy design implemented elsewhere in other geographical areas. If the instance of nonfalsification grows, so does the trust and belief in the theory.

There is nothing objective about such beliefs, as Basu (2005) notes, but this is exactly how scientific knowledge evolves: nonfalsification, or rather, a sequence of falsification, theory correction, and further tests. As scientists, we must use the best prior belief about the problem and thus have no choice but to stick to the Bayesian norm. When applying the unfalsified beliefs to geographically distant areas, a theory should provide an educated guess on what will work.

Despite these potential benefits in increasing knowledge and orthodoxy in the method of scientific exploration, one may still have to consider a problem of choosing between two competing means: one is to learn about causal effects correctly, and feed them back to a theory to come up with better understanding, although this can take some time. The other is to ignore the potential bias of estimators but to use the observational data to get more casual but potentially erroneous interpretations, which may not take as long as the first option. If there is not a good chance of getting an approximately correct interpretation, in the sense that the median interpretation is an unbiased estimator of correct interpretation and its standard errors are reasonably small, one should take the first option. If there is a good chance, then there is a meaningful trade-off between obtaining a consistent interpretation and immediacy of policy implication. The possibility of having a meaningful trade-off, and the severity of trade-off, depend on urgency in the need for policy changes and how badly the estimates from observational studies (using instrumental variables) are biased. As an economist, one may want to acknowledge the relative costs of focusing on randomized experiments and of holding onto observational studies in improving social welfare through better policies.

This is particularly true if one wants to evaluate a project that requires voluntary participation of households. Unlike deworming medicines that are deemed to be beneficial and are assigned to children at parents' will, adults can opt out from a program, more so if the benefits are uncertain or if it entails some costs at the onset. The large number of observations in Progesa may be more than one could hope for, but the minimum number of observations can be far larger and more geographically dispersed than the observational data.

Another drawback is that the estimates in randomized trials only measure total effects or total derivatives of policies (Glewwe and Kremer 2005). They do not give partial effects. To illustrate, recall the simple model of household education choice of equation (P1). We derived that households choose schooling l and education inputs I , and attainment A will be:

$$A = A[l(Q, C, H, P, \lambda_h, a_g), Q, C, H, I(Q, C, H, P, \lambda_h, a_g)].$$

For simplicity, suppose that a government intervention changes P_i only, and keeps other exogenous variables unchanged. Further assume that there is only one education input I . Then, the total derivative of the policy on outcome variable A is:

$$\frac{dA}{dP_i} = \frac{\partial A}{\partial I} \frac{dI}{dP_i} + \frac{\partial A}{\partial I} \frac{dI}{dP_i}. \quad (4)$$

Thus one cannot know the partials $\partial A/\partial I$ or $\partial A/\partial I$, the structural parameters. As Glewwe and Kremer (2005) note, this may be enough for the policymakers who just want to see if the policy worked or not. However, knowing the values of structural parameters is an important policy concern because it gives how much each argument marginally contributes in raising educational attainment. Policymakers may want to know which arguments to strengthen for the policy to be most effective. If the policy did not work, is it because of little impacts of dI/dP_i or dI/dP_i , or is it because these forces are canceling out each other? If it does work, what has made the policy successful? Is it $(\partial A/\partial I)(dI/dP_i)$ or $(\partial A/\partial I)(dI/dP_i)$, or both? If we can measure $\partial A/\partial I$ or $\partial A/\partial I$, we can see which argument must be made more responsive, or where the problem lies.¹³

Putting it differently, while we can measure dI/dP_i or dI/dP_i , we cannot know why and how they changed. This is in a sense a theory version of the “external validity” argument, as it reveals if I or I has been induced to change by policies in the trial area, but cannot tell if the same will happen elsewhere. If dI/dP_i or dI/dP_i were large (small), why and how so?

The answers to the last questions need more detailed models of the household that reveal determination of dI/dP_i and dI/dP_i , and more detailed data that enable statistical tests of competing hypotheses. But this would most probably require orthogonal regressors beyond random assignment, or require the program placement design to be finely tuned to discriminate among competing hypotheses.

A related point is a comparison with clinical trials in medicine, or other controlled experiments in the hard sciences. It is true that these fields also progress through theorizing and controlled experiments. For example, the reasons that a particular drug is effective (or not effective) can be explained by invoking rigorous theories based on the solid microfoundations of molecular biology. Given the greater ability to implement controlled experiments and more direct observability of key variables, or the greater ability to control more variables, theorization is less troublesome and interpretation is less ambiguous. Economics being less able to control all the factors affecting the outcome of interest, theorization is more problematic and there can be more competing theories under a given set of estimated results. This leads to the notion of costs that we discussed earlier: randomized experiments are preferable in identifying the causal effects, while theorizing a causal relationship and implementing it as a policy may entail substantial time and energy. As scientists, economists may want to collect the consistent estimates, while as practitioners, they may want to get a quick fix that may be less accurate or less effective. The analogy can

¹³ Of course, if the policymaker does not have any means to independently control both I and I , knowing the structural parameters may not help make policy more effective.

be similar to the intake of aspirin when one has a headache: it was not well known until as recently as the 1980s why it works, but people use it anyway because it relieves the pain to some extent.

V. CONCLUSION

Though recognized by policymakers everywhere as a problem, school enrollment rates in low-income countries remain low. This study has shown that enrollment rates and per student budget are positively correlated, and the latter is further correlated with the dependency ratio. This indicates the need for the better use of the limited funds available.

This study presented a simple framework for understanding educational outcomes using a unitary model with an altruistic parent and a child. It reviewed the traditional interventions, i.e., supply-side policies, and the recent innovation of relaxing constraints faced by households, i.e., the demand-side policies that have introduced conditional transfer programs. This review showed that, in studies using census-like and school-level data, supply-side policies such as school expansion and compulsory schooling are effective in raising enrollment, but these studies fail to show why. The same is true for teacher reallocation policies and remedial schooling programs. The inability to show the reasons behind the success is due to the lack of detailed household-level data. Without this, one cannot estimate the behavioral relationship of the students and their parents. Therefore a future evaluation of supply-side interventions may be rewarding if one collects both school- (or district-) and household-level data.

On the other hand, randomized demand-side interventions such as Progresia are necessarily accompanied by household-level time, health, and consumption data, because of the need to identify the poor households. This is likely to give an unusually rich possibility for deeper inquiries into questions that have not been answered by studies relying on natural experiments. Given Progresia is both demand- and supply-side intervention, one may want to exploit the detailed household-level information against the supply-side aspects of intervention to understand the heterogeneity in responses. One should, nonetheless, keep in mind that evaluation of demand-side policies is also not straightforward. So long as conditional and demographically targeted transfers affect the distribution of bargaining powers within households, and hence the household-specific prices, they will change the resource allocation which may lower the welfare of some members relative to the absence of such programs. Due primarily to the lack of data, most evaluation studies focus on household-level welfare, or household-level averages, not individual-level welfare. Further inquiries and data collection efforts on who will benefit/lose in net terms are warranted in future research to better understand the consequences.

This study finished with the review on recent trends on estimation techniques. While arguing that randomization is clearly the best for drawing inferences, one may still

want to choose the optimal combination of randomized trials and observational data, as the former requires more resources and time as inputs. This is particularly true for economics rather than for other hard sciences, partly because of economists' inability to fine-tune the controls, and partly because of the lack of a solid microfoundation in economics compared with other sciences when an experiment shows unexpected results.

REFERENCES

- Angrist, Joshua D.; Eric Bettinger; Erik Bloom; Elizabeth King; and Michael Kremer. 2002. "Vouchers for Private Schooling in Columbia: Evidence from Randomized Natural Experiment." *American Economic Review* 92, no. 5: 1535–58.
- Angrist, Joshua D., and Alan B. Krueger. 1990. "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106, no. 4: 979–1014.
- . 1998. "Empirical Strategies in Labor Economics." *Handbook of Labor Economics*, vol. 3A, ed. Orley Ashenfelter and David Card. Amsterdam: Elsevier.
- Angrist, Joshua D., and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114, no. 2: 533–75.
- Banerjee, Abhijit; Shawn Cole; Esther Duflo; and Leigh Linden. 2005. "Remedying Education: Evidence from Two Randomized Experiments in India." Working Paper, Massachusetts Institute of Technology. http://econ-www.mit.edu/faculty/download_pdf.php?id=677.
- Basu, Kaushik. 2005. "New Empirical Development Economics: Remarks on Its Philosophical Foundations." *Economic and Political Weekly*, October 1.
- Bedi, Arjun Singh, and John H. Y. Edwards. 2002. "The Impact of School Quality on Earnings and Educational Returns: Evidence from a Low-income Country." *Journal of Development Economics* 68, no. 1: 157–85.
- Behrman, Jere R., and John Hoddinot. 2001. "An Evaluation of the Impact of Progresa on Preschool Child Height," FCND Discussion Paper no. 104, International Food Policy Research Institute.
- Behrman, Jere R., and Petra E. Todd. 1999. "Randomness in the Experimental Samples of Progresa." Working Paper, International Food Policy Research Institute. http://www.ifpri.org/themes/progresa/pdf/behrmantodd_random.pdf.
- Bertland, Marianne; Esther Duflo; and Sendhil Mullainathan. 2003. "How Much Should We Trust Difference-in-Differences Estimators?" Working Paper, Massachusetts Institute of Technology.
- Chen, Zhiqi, and Frances Woolley. 2001. "A Cournot-Nash Model of Family Decision Making." *Economic Journal* 111, no. 474: 722–48.
- Chinn, Aimee. 2005. "Can Redistributing Teachers across Schools Raise Educational Attainment? Evidence from Operation Blackboard in India." *Journal of Development Economics* 78, no. 2: 384–405.
- Coady, David P., and Susan W. Parker. 2002. "A Cost-effectiveness Analysis of Demand- and Supply-side Education Interventions: The Case of PROGRESA in Mexico." FCND Discussion Paper no. 127, International Food Policy Research Institute. <http://www.ifpri.org/divs/fcnd/dp/papers/fcndp127.pdf>.

- Das, Jishnu; Quy-Toan Do; and Berk Özler. 2005. "Reassessing Conditional Cash Transfers Programs." *World Bank Research Observer* 20, no. 1: 57–80.
- Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91, no. 4: 795–813.
- . 2003. "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa." *World Bank Economic Review* 17, no. 1: 1–25.
- . 2004. "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics* 74, no. 1: 163–97.
- . 2005. "Field Experiments in Development Economics." Paper prepared for World Congress of the Econometric Society.
- Gauri, Varun, and Ayesha Vawda. 2004. "Vouchers for Basic Education in Developing Countries: An Accountability Perspective." *World Bank Research Observer* 19, no. 2: 259–80.
- Gertler, Paul J., and Simone Boyce. 2001. "An Experiment in Incentive-Based Welfare: The Impact of PROGESA on Health in Mexico." Discussion Paper, University of California, Berkeley. http://faculty.haas.berkeley.edu/gertler/working_papers/PROGRESA%204-01.pdf
- Glewwe, Paul, and Michael Kremer. 2005. "Schools, Teachers, and Education Outcomes in Developing Countries." Second draft for *Handbook of Economics of Education*.
- Handa, Suhanshu. 2002. "Raising Schooling Enrollment in Developing Countries: The Relative Importance of Supply and Demand." *Journal of Development Economics* 69, no. 1: 103–28.
- Heckman, James J.; Hidehiko Ichimura; and Petra Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Econometrica* 64, no. 4: 605–54.
- Heston, Alan; Robert Summers; and Bettina Aten. 2002. *Penn World Table*. Version 6.1. Center for International Comparisons at the University of Pennsylvania (CICUP), October.
- Hoddinot, John, and Emmanuel Skoufias. 2004. "The Impact of PROGESA on Food Consumption." *Economic Development and Cultural Change* 53, no. 1: 37–61.
- Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62, no. 2: 467–76.
- Jacoby, Hanan. 2002. "Is There an Intrahousehold 'Flypaper Effect'? Evidence from a School Feeding Programme." *Economic Journal* 112, no. 476: 196–221.
- Lundberg, Shelly, and Robert A. Pollak. 1993. "Separate Spheres Bargaining and the Marriage Market." *Journal of Political Economy* 101, no. 6: 988–1010.
- MacElroy, Marjorie B., and Mary Jean Horney. 1981. "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand." *International Economic Review* 22, no. 2: 333–49.
- Manser, Marylin, and Murray Brown. 1980. "Marriage and Household Decision-Making: A Bargaining Analysis." *International Economic Review* 21, no. 1: 31–44.
- Martinelli, César, and Susan W. Parker. 2003a. "Should Transfers to Poor Families be Conditional on School Attendance? A Household Bargaining Perspective." *International Economic Review* 44, no. 2: 523–44.
- . 2003b. "Do School Subsidies Promote Human Capital Investments among the Poor?" ITAM Discussion Paper Series no. 03–06, Instituto Tecnológico Autónomo de México.

- Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72, no. 1: 159–217.
- Mukherjee, Aditya; Urmila Sarkar; and Ratna M. Sudarshan. 2005. "MV Foundation: An Evaluation of the Programme 'Elimination of Child Labour through the Universalization of Elementary Education'." A report submitted for MVF-EU Donor Consortium, 2005.
- Mullainathan, Sendhil. 2006. "Development Economics through the Lens of Psychology." Working Paper, Department of Economics, Harvard University.
- Parker, Susan W., and Emmanuel Skoufias. 2000. "The Impact of Progresa on Work, Leisure, and Time Allocation." Discussion Paper, International Food Policy Research Institute. http://www.ifpri.org/themes/progresa/pdf/parkerskoufias_timeuse.pdf.
- Rawlings, Laura B., and Gloria M. Rubio. 2005. "Evaluating Impacts of Conditional Cash Transfer Programs." *World Bank Research Observer* 20, no. 1: 29–55.
- Ravallion, Martin, and Quentin Wodon. 2000. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy." *Economic Journal* 110, no. 462: C158–C175.
- Rosenzweig, Mark R., and Kenneth I. Wolpin. 2000. "Natural 'Natural Experiments' in Economics." *Journal of Economic Literature* 38, no. 4: 827–74.
- Skoufias, Emmanuel. 2005. "Progresa and Its Impact on the Welfare of Rural Households in Mexico." Research Report no. 139, International Food Policy Research Institute. <http://www.ifpri.org/pubs/abstract/139/rr139.pdf>.
- Schultz, T. Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74, no. 1: 199–250.
- Spohr, Chris A. 2003. "Formal Schooling and Workforce Participation in a Rapidly Developing Economy: Evidence from 'Compulsory' Junior High School in Taiwan." *Journal of Development Economics* 70, no. 2: 291–327.
- Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regressions with Weak Instruments." *Econometrica* 65, no. 3: 557–86.
- Stock, James H.; Jonathan H. Wright; and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business, Economics, and Statistics* 20, no. 4: 518–29.
- United Nations Educational, Scientific and Cultural Organization (UNESCO). 2004. *The Global Education Digest 2004: Comparing Education Statistics across the World*. Montreal: UNESCO Institute of Statistics.
- Vermeersch, Christel. 2003. "School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation." Photocopy. University of Oxford. <http://www.nuff.ox.ac.uk/users/vermeersch/schoolmeals.pdf>.