One Economist's Perspective on Some Important Estimation Issues

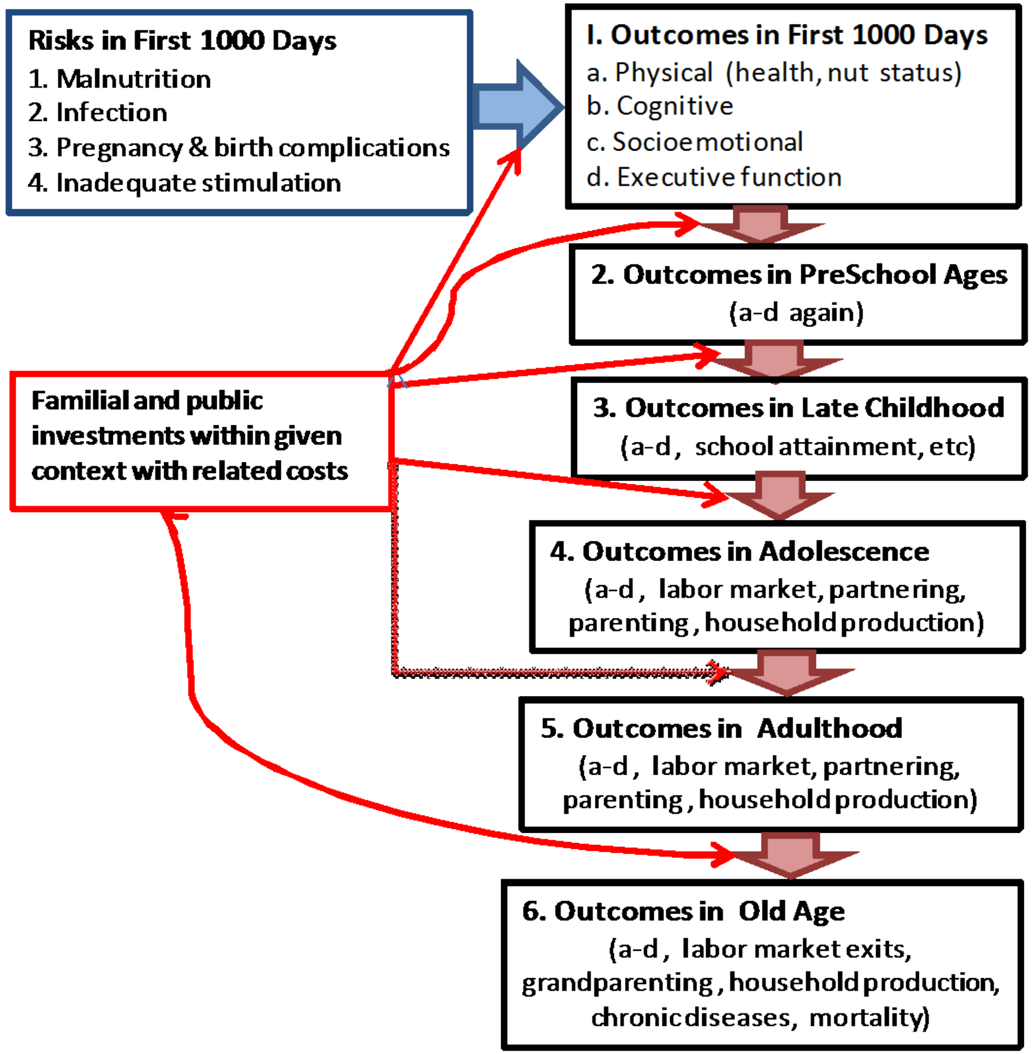
Jere R. Behrman
W.R. Kenan Jr. Professor of Economics & Sociology
University of Pennsylvania

SRCD Seattle Preconference on

"**Interventions for Children and Youth in Low- and Middle-Income Countries: New Opportunities and Challenges for Developmental Science**"

17 April 2013

I. Life-Cycle Framework



- Behavioral choices make it challenging to estimate impact of one variable (e.g., schooling, ECD program) on another.
- Measurement problems include persistent unobserved factors (“endowments” such as genetics), likely intergenerationally correlated.
- Dynamics include lagged effects but also forward-looking behaviors (and other unobservables) and dynamic complementarities or substitution.
- Heterogeneities may be critical (unobservables, parameters).
- Parameters, not variance decomposition, of primary interest.
- Generalizability from sample and from context?

II. Unobserved Variable Bias, Endogeneity, Simultaneity

Interested in estimating impact of nutritional status HAZ (H) on cognitive skills (C) in linear approximation:

(1) $C_t = a_0 + a_1 H_t + a_2 X + a_3 U^f + a_4 U^v_t + u_t$,

where subscripts refer to time period,
C_t = cognitive skills in the tth period,
H_t = height-for-age z score (HAZ) in the tth period,
X = other observed characteristics assumed fixed (e.g., sex, age),
U^f = unobserved fixed characteristics (e.g., innate ability, innate health),
U^v_t = unobserved time-varying characteristics (e.g., cognitive skills accumulated through interacting with others in same context),
u_t = random disturbance term (e.g, chance learning encounters),
a_i = parameters to be estimated

Some characteristics of relation (1)

(1) For impact estimate, coefficient a_1 of interest, NOT explained variance (which could be small if variance H low or variance u high)

(2) Variables (upper case letters) all may be vectors but for simplicity they are treated here as if they each have just one element.

(3) For simplicity assumed that only one observed choice variable (H_t) on right side though in more extensive specifications there may be others (e.g., home stimulation, pre-school attendance mother's time with child).

(4) If ordinary least squares (OLS) multivariate regression estimates are made of relation (1) the effective disturbance term is $a_3 U^f + a_4 U^v_t + u_t$.

(5) Except for simultaneity (below), same framework holds if C for period t + n or if H lagged.

(6) Important question is whether H_t correlated with effective disturbance term ($a_3 U^f + a_4 U^v_t + u_t$); if correlation, estimate of a_1 biased because H_t proxies in part for effective disturbance term, not only for effect of H_t alone.

Suppose H_t determined by linear relation:

(2) $H_t = b_0 + b_1 C_t + b_2 X + b_2'Z + b_3 U^f + b_4 U^v_t + v_t$,

Where variables as defined above except:

- Z represents observed variables that affect H_t. but not C_t directly (e.g., the price of nutrients, which does not directly affect cognitive skills in relation 1),
- random term is v_t (not correlated with u_t in relation 1)
- parameters are b’s instead of a’s.

In this relation H_t determined in part by current cognitive skills. That is relations (1) and (2) simultaneously determine cognitive skills and HAZ.

- Clear from relation (2) that estimation problem in using OLS for relation (1) because in relation (2) H_t depends on U^f and $U^v{}_t$ both of which are in the effective disturbance term (a_3 U^f + a_4 $U^v{}_t$ + u_t) in relation (1).
- OLS estimate of a_1 biased because H_1 in relation (1) is representing in part correlated part of these unobserved variables – and there will be a correlation because the same unobserved variables affect H_t in relation (2). ("unobserved variable bias").
- Note that can occur even if $b_1 = 0$ so that cognitive skills do NOT determine HAZ.

IF simultaneous determination of C_t and H_t so that b_1 is nonzero, relation (1) can be substituted into relation (2) :

(3) $H_t = b_0 + b_1 a_0 + [(b_2 + b_1 a_2)/(1- b_1 a_1)] X + [b_2' /(1- b_1 a_1)] Z + [(b_3 + b_1 a_3) /(1- b_1 a_1)] U^f + [(b_4 + b_1 a_4) /(1- b_1 a_1)] U^v_t + b_1 u_t + v_t$ or

(4) $H_t = c_0 + c_2 X + c_2'Z + c_3 U^f + c_4 U^v_t + z_t$,

where the definition of the c's and z_t should be clear from comparing relations (3) and (4).

Again, clear estimation problem in using OLS for relation (1) because in relations (3) and (4) H_t depends on U^f and U^v_t both of which are in the effective disturbance term ($a_3 U^f + a_4 U^v_t + u_t$) in relation (1).

What can be done to eliminate this bias or, equivalently, to eliminate the correlation between H_t and the effective disturbance term ($a_3\ U^f + a_4\ U^v{}_t + u_t$) in relation (1)?

(1) Random controlled trial (RCT) so H_t randomly distributed across individuals and NOT outcome of individual behaviors as in relation (2). A lot of attractions, but conducting a RCT not within the scope of what is often possible. (May provide Z.)

(2) "Natural experiment" in which some people beneficiaries of options that improved HAZ and others were not e.g. policy or weather changes that affected people differentially depending on age or geography or both. (May provide Z.)

(3) Fixed effects (FE) estimates in which dichotomous ("dummy") variable added for each individual that control for ALL individual fixed characteristics, including U^f and thus removes U^f from compound disturbance term in relation (1):

- Need multiple observations at level for which FE used
- Related to first-differencing & sibling (twins) estimates
- Still biases if unobserved time-varying factors that enter into both relation (1) and relation (2).

(4) Instrumental variable (IV) estimates: correlation between H_t and disturbance term ($a_3\ U^f + a_4\ U^v_t + u_t$) in relation (1) broken by replacing H_t by predicted value of H based on relation (4) (which depends on X and Z, but not U^f and U^v_t).

- Two-stage process (two-stage least squares, 2SLS): in first-stage relation (4) estimated and then second-stage relation (1) estimated using predicted value of H rather than actual value.
- For Z to be good instrument it must (i) predict well H and (ii) NOT be included in relation (1) (or, equivalently, not correlated with disturbance term ($a_3\ U^f + a_4\ U^v_t + u_t$)). Standard statistical tests.
- # of instruments must be >= # of right-side behavioral variables.
- Perhaps in combination with FE estimates (IV-FE).
- Local average treatment effects "LATE" (e.g., minimum schooling laws)

(5) Regression discontinuity (compare those right above and right below eligibility cutoffs)

(6) Propensity score matching (compare based on similar propensities to receive treatment given observed variables)

$$(1') \; S^c = a_0 + a_1 S^m + a_2 X + a_3 U^f + a_4 U^v + u,$$

Where S^c = child schooling, S^m is mother's schooling, U^f is family endowments for mother, U^v is individual-specific endowments for mother.

Probable bias because S^m correlated with unobserved endowments (parallel relation for S^m).

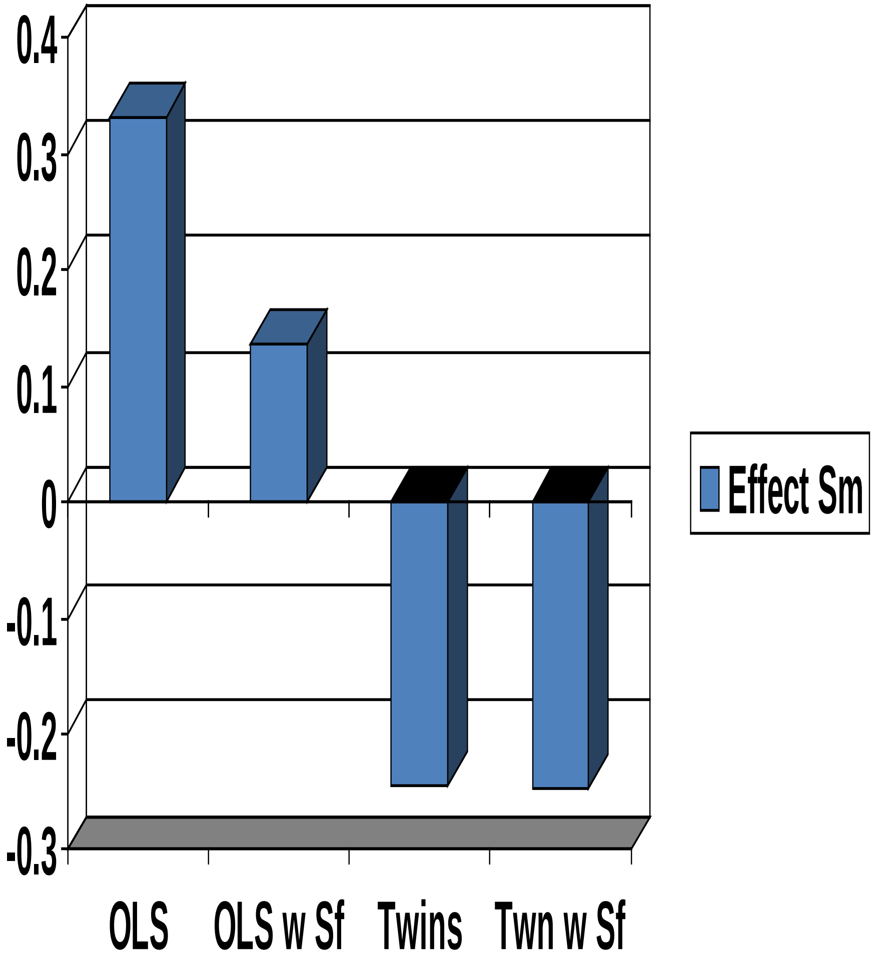
Adult sisters (mothers) sibling estimates.

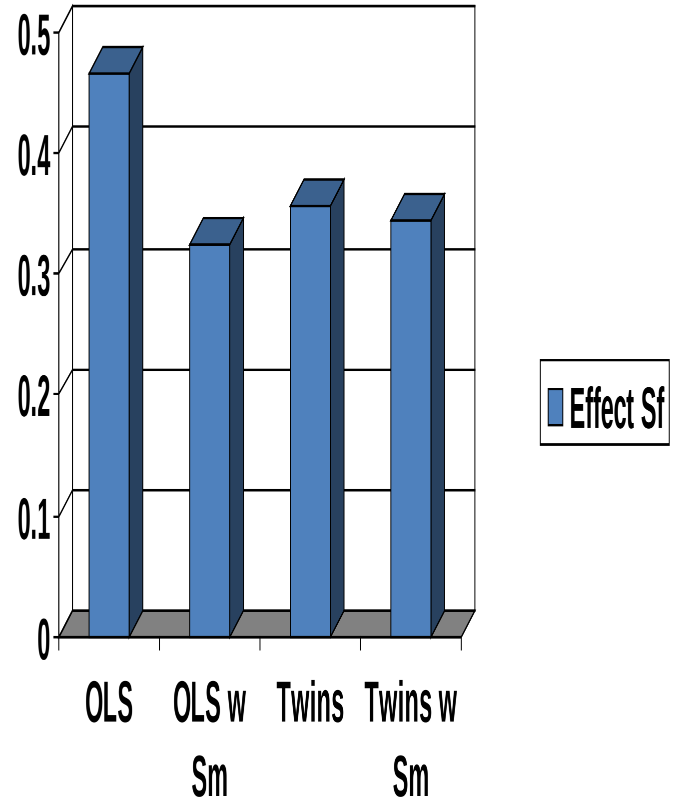
Behrman, J. R. and B. L. Wolfe (1987). "Investments in Schooling in Two Generations in Pre Revolutionary Nicaragua: The Roles of Family Background and School Supply." Journal of Development Economics 27:12(October): 395-420.

Adult sisters identical twins estimates.

Behrman, J. R. and M. R. Rosenzweig (2002). "Does increasing women's schooling raise the schooling of the next generation?" American Economic Review 92(1): 323-334

Mothers' (left) and Fathers' (right) Additional Grade of Schooling and Child Schooling (vertical)





$$(1'')\ Y = a_0 + a_1 BW + a_2 X + a_3 U^f + a_4 U^v + u,$$

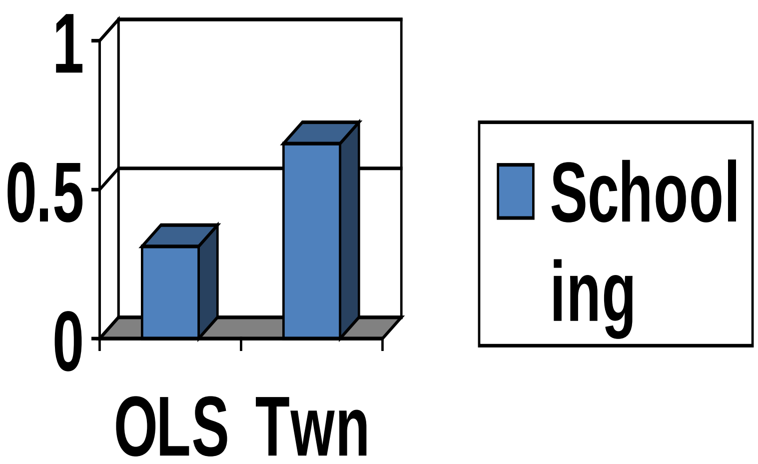
Where Y is child outcome, BW is birth weight, U^f is family endowments for child, U^v is individual-specific endowments for child.

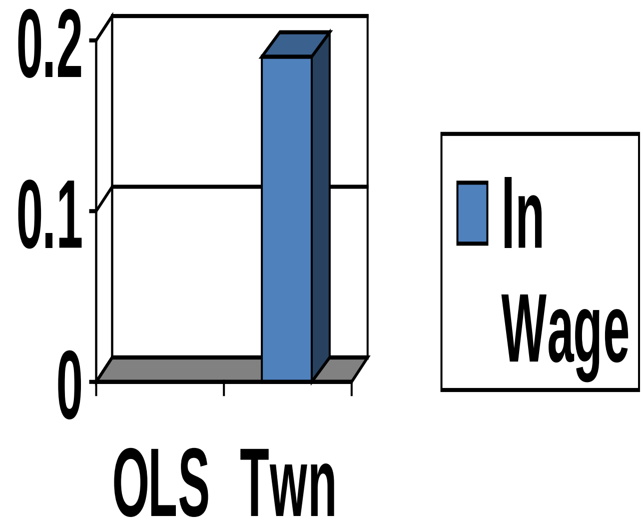
Probable bias because BW correlated with unobserved endowments (parallel relation for BW).

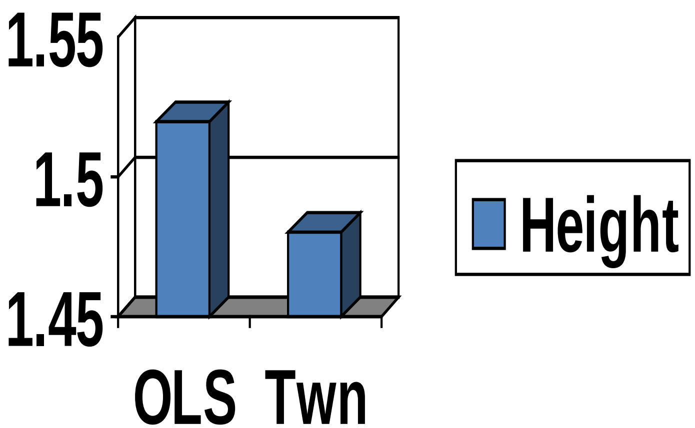
Adult sisters identical twins estimates.

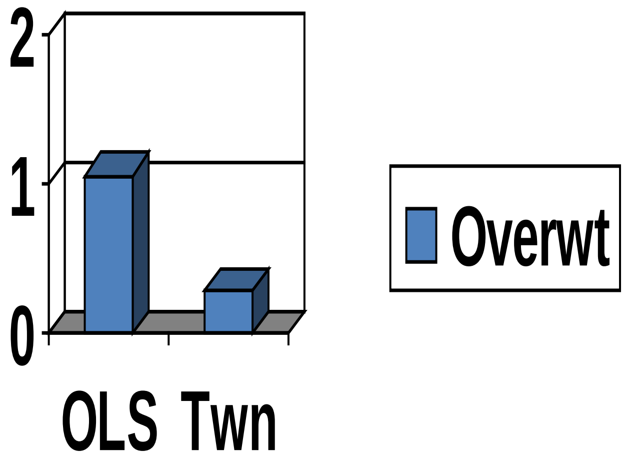
Behrman, J. R. and M. R. Rosenzweig (2004). "Returns to Birthweight." The Review of Economics and Statistics **86**(2): 586-601

Impact of Higher Birth Weight with (Twn) and without (OLS) Control for Endowments (Twn)









$$(1''') \; R = a_0 + a_1 S + a_2 \, Stunt + a_3 \, X + a_4 \, U^f + a_5 \, U^v + u,$$

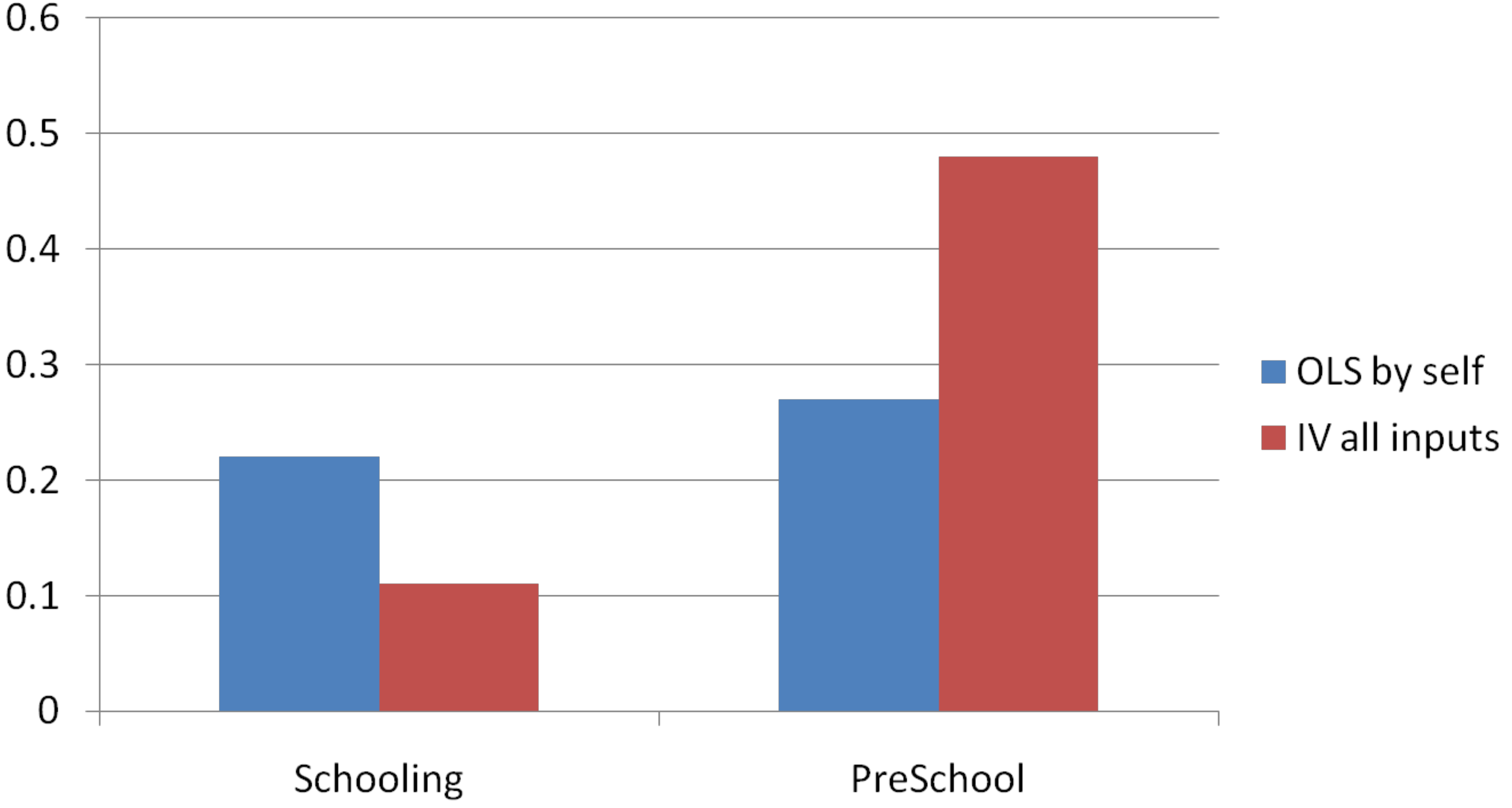
Where R is adult reading comprehension, S is completed grades of schooling, Stunt is stunted when of preschool age, U^f is family endowments for child, U^v is individual-specific endowments for child.

Probable bias because S and Stunt correlated with unobserved endowments (parallel relation for both).

Instrumental variables estimates for Guatemala.

Behrman, JR, J Hoddinott, JA Maluccio, E Soler-Hampejsek, EL Behrman, R Martorell, A Quisumbing, M Ramirez and AD Stein"What Determines Adult Skills? Impacts of Pre-School, School-Years and Post-School Experiences in Guatemala", Washington, DC: IFPRI Discussion Paper No. 826, 2008.

SDs in Reading Comprehensive Score per Grade of Schooling or Pre-School Child Not Stunted



III. Related Other Thoughts

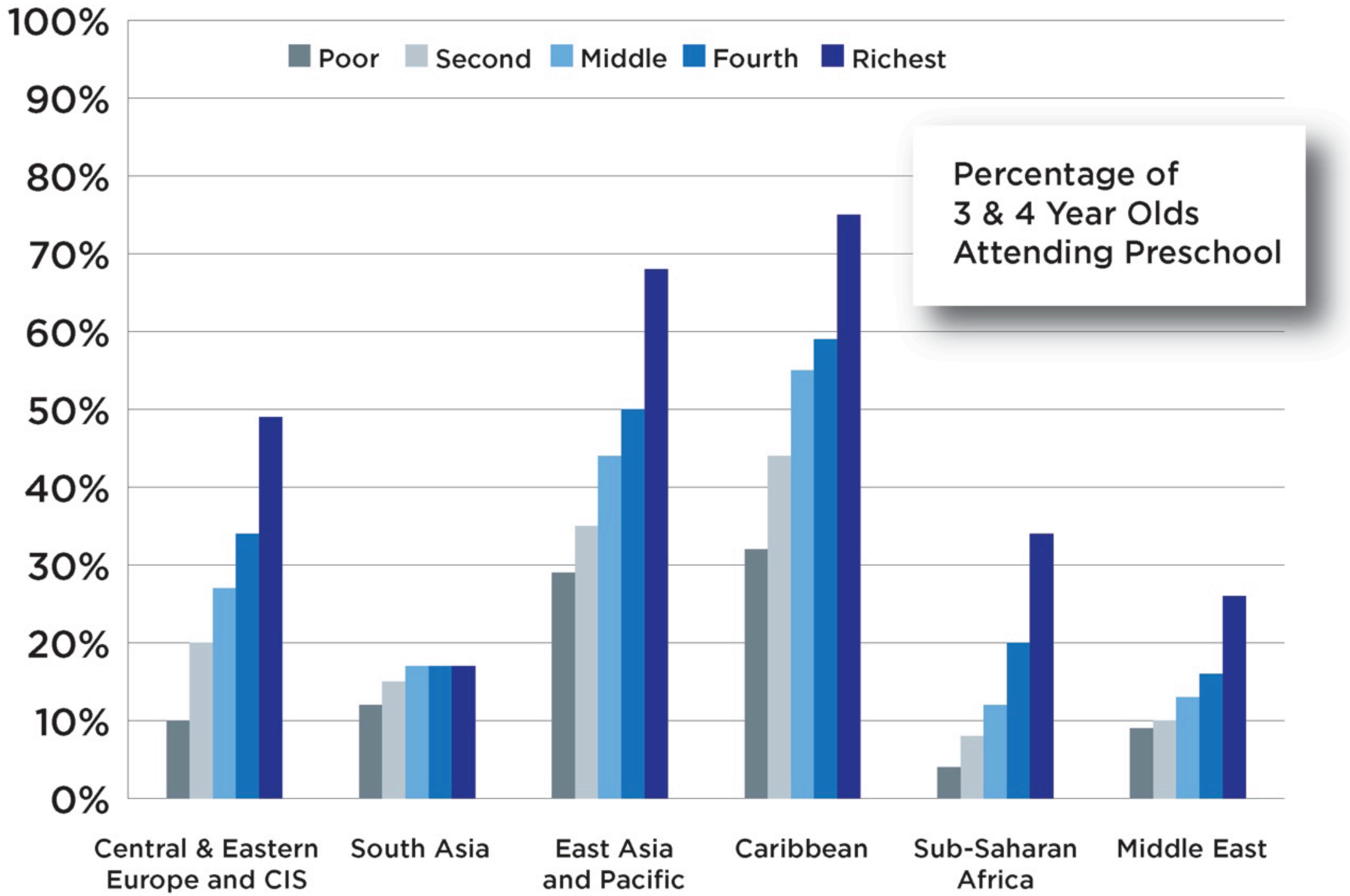
III-A. Treatment effects (T) added:

(1) $C_t = a_0 + a_1 H_t + a_2 X + a_3 U^f + a_4 U^v{}_t + a_5 T_t + u_t$

Related issue if endogenous choice of treatment (i.e., behavioral choices of T). Therefore: (i) RCTs, (ii) Natural experiments, (iii) FE (iv), IV (v) Regression discontinuity, (vi) Propensity score matching.

III-B. Selected samples (e.g., those attending preschool programs) (high value of u_t).

Preschool enrolment by region and income –less than 20% for poorer income quintiles



Proportion of young children attending preschool in 58 low-income and middle-income countries by income quintile within country summed across sample countries by region. Data are from UNICEF's 2005 Multiple Indicator Cluster Survey 3 for children aged 3 and 4 years.

III-C. Life-cycle Implications for impact evaluation: Present Discounted Values in U.S. $ of Seven Major Benefits of Moving an Infant Out of Low Birth Weight Status in a Poor Developing Country

(Alderman, H. and J. R. Behrman (2006). "Reducing the Incidence of Low Birth Weight In Low-Income Countries has Substantial Economic Benefits." World Bank Research Observer **21**(1): 25-48.)

	Annual discount rate		
Benefit	3 percent	**5 percent**	10 percent
1. Reduced infant mortality	$95	**$99**	$89
2. Reduced neonatal care	$42	**$42**	$42
3. Reduced costs of infant and child illness	$36	**$35**	$34
4. Productivity gain from reduced stunting	$152	**$85**	$25
5. Productivity gain from increased cognitive ability	$367	**$205**	$60
6. Reduced costs of chronic diseases	$49	**$15**	$1
7. Intergenerational benefits	$92	**$35**	$6
Total	$832	**$510**	$257
Share of total at 5 percent discount rate (percent)	163%	**100%**	50%

Note: The 5 percent discount rate, shown in bold in the table, is the base case estimate.

III-D. Resource Costs – Private and Public -- as Important as Benefits:

Benefit-Cost Ratios for Reducing Low Birth Weight

	Benefits/Costs
Treatment for women with asymptotic bacterial infections	0.6-4.9
Treatment for women with presumptive sexually transmitted diseases	1.3-10.7
Drugs for pregnant women with poor obstretic history	4.1-35.2

Source: Behrman, JR, H Alderman and J Hoddinott (2004), "Hunger and Malnutrition" in Bjorn Lomborg, ed., *Global Crisis, Global Solutions* Cambridge, UK, Cambridge University Press, 363-420.

III-E Context Matters – And Risks of Facile Claims of "Best Practice" (e.g. J-PAL Website)

Extra Years of Education per $100 Spent

Evaluations	Cost-Effectiveness	Impact: 90% Confidence Interval		
		Lower Bound	Point Estimate*	Upper Bound
Programs				
1 **Information in Madagascar** Giving parents information on the higher wage returns to education caused higher student attendance.	19.8 years	0.19	3.517	6.84
2 **Deworming in Kenya** Deworming children at school decreased absenteeism by 25% and was extremely cost-effective.	6.1 years	4.59	11.25	17.91
3 **Iron & Deworming in India** Children were given iron supplements and deworming pills to fight anemia, enabling them to attend school more often.	2.9 years	0.21	5.8	11.39
4 **Merit Scholarships in Kenya** Merit scholarships for high-performing girls induced all students to attend more regularly.	0.27 years	0.24	3.20	6.16
5 **Free School Uniforms in Kenya** Subsidizing uniforms, a large part of school expenses, increased student attendance.	0.72 years	2.95	6.40	9.85
6 **CCT for Primary Enrollment in Mexico** Families were given cash transfers conditional upon their children attending primary school.	0.032 years	0.77	1.23	1.68

*Point estimate of the percentage point increase in enrollment/attendance

Guess for which country interventions to increase primary schooling in the last decade are likely to be least effective?

Country	Net Primary School Enrollments in 2000
Madagascar	68%
Kenya	65%
India	79%
Mexico	97%

III-F. Structural Models

- Estimation of underlying preference relations and production functions and other constraints.
- Learn about, for example, dynamic complementarities (e.g., Cunha and Heckman).
- Can perform counterfactual experiments (e.g., Todd, P. E. and K. I. Wolpin (2006). "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." American Economics Review **96**(5): 1384-1417.)