

ONLINE APPENDIX: NOT FOR PUBLICATION

When Measure Matters: Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries

M. Shahe Emran
IPD, Columbia University
William Greene
New York University
Forhad Shilpi
DECRG, World Bank

Truncation Bias in Estimated IGRC and IGC: Estimates from Alternative Age Ranges

The age range used in Tables 2-4 of the paper is 13-60 years. This is motivated by the fact that the average schooling attainments in rural Bangladesh and India remain low in the survey years, so that a 13 years lower threshold may not be binding for most of the rural children. In Bangladesh data, the average years of schooling is only 4.43 years; for sons it is 5.5 years and for daughters 3.4 years. The average schooling in India is 5 years, and for sons it is 7 years and for daughters 3.7 years. To explore the sensitivity of the conclusions with respect to the age range of children, we estimate the IGRC and IGC across a number of different age ranges. For the sake of brevity, we report estimates from the following age ranges: (i) 13-50 years, (ii) 16-60, and (iii) 20-69 years.

Since many children start first grade at age 6, a 13 years age cut-off implies 7 years of potential schooling as the minimum threshold in our sample (primary schooling is 5 years). The observed schooling attainment, however, may vary across 13 year old children for a variety of reasons. For example, children from poor households may start schooling later than usual, and they may also have to interrupt schooling because of negative economic shocks.¹ The variations in schooling attainment at age 13 (or even younger) can thus provide us useful evidence on the

¹ According to one estimate for India, 53 percent of students drop out before completing primary (5 years). Among every 100 girls enrolled, only 40 progress to 4th grade, 18 reaches 8th grade, and only 1 is lucky enough to go up to 12th grade (India Education Report, 2005, pp. 6-7).

role played by family background. However, one might worry that some children at age 13 have not yet completed schooling, and it is important to check if the results hold when the lower threshold for children's schooling is raised. We thus estimate the IGRC and IGC in a series of samples, starting with 14 years and raising the lower threshold incrementally by one year at a time up to 20 years. The evidence from this exercise is very reassuring; while the magnitudes differ across the samples, the main conclusions reached on the basis of 13-60 years age range remain intact. For the sake of brevity we report estimates for 16 and 20 years as the lower age threshold for children. A 16 year age cut-off implies potentially 10 years of schooling which coincides with one of the most important public examination in both Bangladesh and India (called Secondary School Certificate (S.S.C) or 'Matriculation' examination). After 12 years of schooling (18 years of age cut-off), the students sit for a second important public examination, called Higher Secondary Certificate (H.S.C) or Intermediate examination. In our Bangladesh data, about 10 percent of 20 years of age or older has 10 years or more schooling, and 5 percent has 12 years or more schooling.

For each age range, we present the bias estimates and omit the underlying estimates of IGRC and IGC for full and coresident samples. All of the underlying estimates are available from the authors upon request. Note that we do not discuss statistical significance of the estimates, all of the estimates are significant at the 5 percent or lower level. But as noted in the main text, the estimated standard errors in IGRC and IGC are very small reflecting the large sample size. Thus statistical significance is not very informative, and the focus should be on the magnitude of the bias.

The estimates for three different age ranges for the all children sample including both sons and daughters are presented in Table T.1. The estimates, both for India and Bangladesh, lead to the same set of conclusions derived from the 13-60 years age range in section (4) of the main text of the paper. The IGRC estimates, in general, suffer from substantial downward bias because of truncation due to coresidency. In Bangladesh, the bias in the coresident sample estimate of IGRC is more than 10 percent in seven out of nine cases with an average bias of 15.56 percent. The extent of bias in the case of India is smaller, the average is 14.7 percent, but still the bias is more than 10 percent in seven out of nine cases.

In contrast, the coresidency bias in the IGC estimates are again much smaller, in only one out of nine cases the bias is more than 10 percent in Bangladesh (10.9 percent when father's age is the indicator of parental education and the age range for children is 1350 years). The average bias in the IGC estimate for Bangladesh is only 4.6 percent, less than one-third of that in the IGRC estimates (15.56 percent). The estimates for India are similar; in three out of nine cases the bias is more than 10 percent, the highest being 12 percent, and the average is 9 percent, significantly smaller than the corresponding average for IGRC (14.7 percent).²

Table T.2 reports estimates for the father-son and mother-daughter persistence in schooling attainment for different age ranges of children. As to be expected, the magnitudes of the estimates vary across different age ranges, but the main conclusions of the paper remain valid. The coresidency bias in the IGRC estimates is very high in the estimates for Bangladesh; the lowest bias is 15 percent and the highest 46 percent, with an average of 27 percent, a very high bias by any standard. The bias estimates for India are smaller in magnitude consistent with its higher coresidency rates. However, the average bias in IGRC is still more than 10 percent (10.6 percent). *More important for the research on intergenerational mobility in developing countries constrained by the coresident samples is the clear evidence that in all 12 cases reported in Table T.2, the bias in the IGC estimate is much smaller than that in the corresponding IGRC estimate. The average bias in the IGC estimates is only 7.5 percent in Bangladesh (27 percent for IGRC), and 4.5 percent in India (10.6 percent for IGRC).*

² Note that there are two negative estimates of bias out of a total of 36 estimates in Table T.1, implying that the estimate from full sample is smaller than that from the coresident sample. It is, however, important to underscore that, in both of these cases, the numerical magnitudes of estimates from full and coresident samples are extremely close; the IGRC estimates are 0.83 (full sample) and 0.84 (coresident sample), and the IGC estimates are 0.53 (full sample) and 0.54 (coresident sample).

**Table T.2: Truncation Bias in IGRC and IGC: Robustness Checks for Different Age Ranges
Father-Son, and Mother-Daughter**

Biases	Bangladesh		India	
	<u>Father-Son</u>	<u>Mother-Daughter</u>	<u>Father-Son</u>	<u>Mother-Daughter</u>
16-60 Year Age group				
Intergenerational Regression Coeff. (IGRC)	22%	32%	6%	16%
Intergenerational Correlations(IGC)	9%	4%	2%	7%
20-69 Year Age group				
Intergenerational Regression Coeff. (IGRC)	15%	18%	6%	7%
Intergenerational Correlations(IGC)	6%	6%	2%	4%
13-50 Year Age group				
Intergenerational Regression Coeff. (IGRC)	30%	46%	7%	22%
Intergenerational Correlations(IGC)	9%	11%	2%	10%

Notes: (1) Bias is defined as [(Estimate from Full Sample – Estimate from Coresident Sample) *100] / Estimate from Coresident Sample. (2) All of the Bias Estimates are Significant at the 5 Percent or Lower Level. (3) Data Sources: India: Rural Economic and Demographic Survey (REDS) 1999; Bangladesh: Matlab Health and Socioeconomic Survey 1996.

Figure F2: Intergenerational persistence in education: non-parametric graphs for Bangladesh

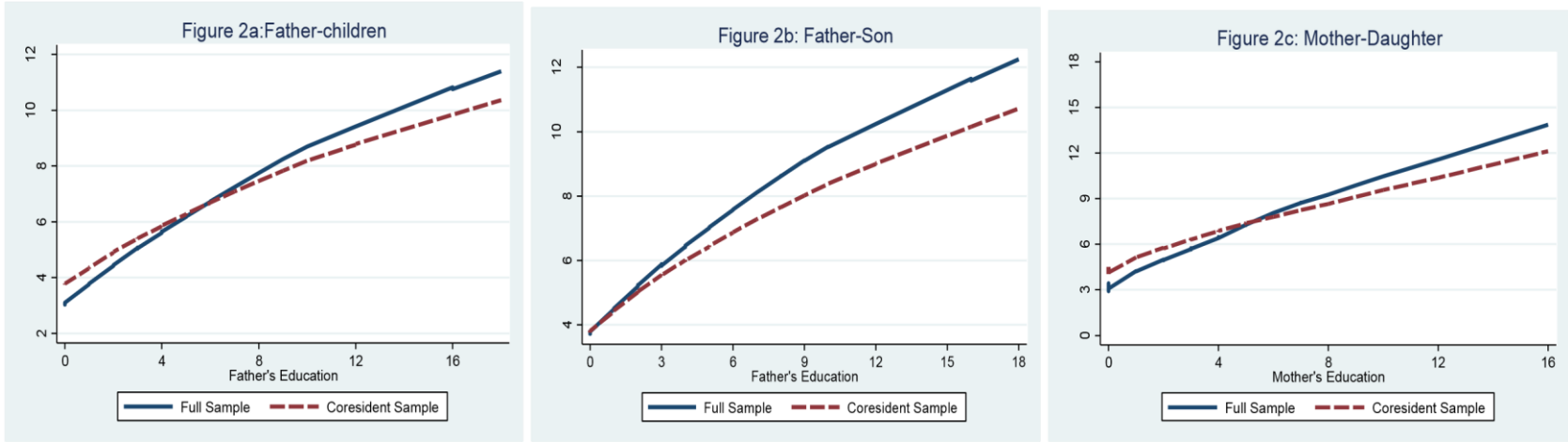


Figure F3: Intergenerational persistence in education: non-parametric graphs for India

