Comment

Julian Cristia: Latin America fares poorly in international learning assessments even when compared with countries of similar economic development (see Berlinski and others 2011, figure 1). There is significant evidence showing that academic achievement affects important adult outcomes such as employment and wages and some suggestive evidence on its aggregate effects on GDP growth (Dougherty 2003; Hanushek and Woessmann, 2008). This scenario provides a powerful motivation to identify programs that can increase students' learning.

This paper contributes to the quest for effective programs by evaluating the impacts of the Brazilian Public School Math Olympics (OBMEP) on students' learning. It uses linked data on students' test scores, participation in the competition, and sociodemographic characteristics. The authors implement a difference-in-differences design augmented with a statistical technique aimed at increasing the similarity between schools participating in the program and a comparison group (propensity score reweighting). Results suggest that the program produced a positive and significant effect on mathematics test scores for participating students in the ninth grade. These effects seem to be concentrated among high-achieving students and in schools that have participated in the OBMEP multiple times.

The evaluated program is quite different from other interventions that have been studied in education. Traditional interventions have aimed at providing certain inputs in the educational process (for example, books, materials, furniture, upgraded infrastructure, and teacher training). The evidence on limited impacts of this type of input-based intervention coupled with the substantive effects on school attendance produced by conditional cash transfer programs has fueled a shift toward studying the effects of interventions aimed at increasing student effort. In recent years there has been a wave of studies evaluating the effects of providing monetary incentives to students (Jackson 2010; Fryer 2011). The OBMEP program shares the same focus on

incentives as those studies, but it differs in providing nonmonetary rewards to top performers.

Because of this unique program feature (the provision of nonmonetary rewards to top performers), it can be expected that the intervention may produce behavioral changes mainly among high-achieving students. Consequently, as opposed to many other interventions in education, this program could increase average performance but at the cost of increased inequality, given that the expected effects could be larger on top students. This expectation seems to have materialized in practice because program take-up is heavily concentrated among large, urban, high-performing schools and because, even within this set of schools, the effects seem to be concentrated in top students. However, if the program is highly cost effective, then it could free up resources to fund compensating actions and produce learning gains across the whole distribution. For example, it could be implemented in tandem with increases in class size in top-performing schools and reductions in low-performing ones.

Because the program is heavily tilted toward high-achieving schools, it creates a substantial challenge for generating unbiased estimates of its effects. The authors make a significant effort to document that the estimated differences in test scores between participating schools and a comparison group represent causal effects. But the resulting evidence does not rule out the possibility of biases in the estimated results. First, the authors document that there are no significant differences between students in fifth grade in participating schools and the comparison group. Because fifth-graders did not participate in the competition, the absence of statistically significant differences may suggest no selection. However, the estimated coefficient is positive and quite large (about 60 percent of the baseline effect), and it is not statistically significantly different from the main effect. Second, results indicate a positive and significant effect in reading. Though the authors suggest that this is evidence of spill-over effects, another plausible explanation is that the estimated effects are biased. That is, "effects" on reading (a nontargeted outcome) may indicate that, in the absence of the program, participating students would have performed better than those in the comparison group.

Data on intermediate channels would have helped to assess the credibility of effects on final outcomes. For example, data on study time at home and time-on-task at school in mathematics and language would have suggested whether participating in the program indeed affected behavior and, in particular, whether the increases in study time were present both in math and other subjects. Unfortunately, the authors cannot explore potential mechanisms because there are no administrative data on them.

To sum up, my view is that the reviewed paper has made significant inroads in assessing an important policy question but has met certain hurdles in terms of the existing variation in program (and it consequences on potential biases) and on the available data. Overall, the results provide some suggestive evidence on positive program effects concentrated among top-performing students, at low cost. This study could be an excellent starting point to design and implement a large-scale, randomized evaluation in collaboration with the agencies running the program. This follow-up study could exploit again the administrative data to obtain large sample sizes (and hence precise estimates) but include primary data collection on a subsample of treatment and control schools to shed light on intermediate channels. Under plausible assumptions, the costs associated with producing such an evaluation would be a small fraction of the value of the expected benefits in terms of the better use of public resources that this new study could induce.

References

Berlinski, S., and others. 2011. "Computers in Schools: Why Governments Should Do Their Homework." In *Development Connections: Unveiling the Impact of New Information Technologies,* edited by A. Chong. Washington, D.C.: Palgrave Macmillan.

Curi, Andréa Zaitune, and Naércio Aquino Menezes Filho. 2007. "A Relação entre o Desempenho Escolar e os Salários No Brasil." Insper Working Paper 058.

Deci, E. L., R. Koestner, and R. M. Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin* 125 (6): 627–68.

Dehejia, R., and S. Wahba. 1999. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94: 1053–62.

Dougherty, C. 2003. "Numeracy, Literacy, and Earnings: Evidence from the National Longitudinal Survey of Youth." *Economics of Education Review* 22: 511–21.

Fryer, R. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." *Quarterly Journal of Economics* 126: 1755–98.

Glewwe, P., and M. Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." In *Handbook of Economics of Education,* vol. 2, edited by E. Hanushek and F. Welch. Elsevier.

Hanushek, E., and L. Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46: 607–68.

———. 2009. "Schooling, Cognitive Skills, and the Latin American Growth Puzzle." Working Paper 15066. Cambridge, Mass.: National Bureau of Economic Research.

IBGE. 2000. Censo Demográfico.

Imbens, G. M., and J. M. Wooldridge. 2008. "Recent Developments in the Econometrics of Program Evaluation." Working Paper 14251. Cambridge, Mass.: National Bureau of Economic Research.

INEP/MEC. 2010. "Microdados da *Prova Brasil* 2005 e 2007 e do Censo Escolar 2006."

Jackson, K. 2010. "A Little Now for a Lot Later: An Evaluation of a Texas Advance Placement Incentive Program." *Journal of Human Resources* 45: 591–639.

Lepper, Mark R., Jennifer Henderlong Corpus, and Sheena S. Iyengar. 1995. "Intrinsic and Extrinsic Motivational Orientations in the Classroom." *Journal of Educational Psychology* 97 (2): 184–96.

López-Calva, L., and N. Lustig. 2010. "Declining Inequality in Latin America: A Decade of Progress?" Brookings Institution Press and the United Nations Development Programme.

Murnane, R. J., and others. 2000. "How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?" *Journal of Policy Analysis and Management* 19 (4): 547–68.

OBMEP. 2011. "OBMEP em Números" (www.obmep.org.br/obmep_em_numeros.html).

PISA. 2010. "PISA 2009 Results" (www.oecd.org/edu/pisa/2009).

Robins, J. M., and A. Rotnitzky. 1995. "Semiparametric Efficiency in Multivariate Regression Models with Missing Data." *Journal of the American Statistical Association* 90 (429): 122–29.

Robins, J. M., A. Rotnitzky, and D. Scharfstein. 1999. "Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models." In *Statistical Models in Epidemiology: The Environment and Clinical Trials,* IMA vol. 116, edited by M. E. Halloran and D. Berry, pp. 1–92. New York: Springer-Verlag.

Rosenbaum, Paul, and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.