Human Development Quarterly Update

Latest research findings

Shrinking classroom age variance raises student achievement—considerably

In developing countries, classes typically exhibit a high variance in student age. Using exogenous variation in the variance of student age in ability-mixing schools, L. Choon Wang [1] finds robust negative effects of classroom age variance on fourth graders' achievement. In a simulation, he demonstrates that re-grouping students by age in the sample can improve math and science test scores by roughly 0.1 standard deviations. According to past estimates for the United States, this effect size is similar to that of raising expenditures per student by 26 percent.

Mass media and public services: more accountability or larger private spending?

Using unique data from Benin, Phil Keefer and Stuti Khemani [2] find that literacy rates among school children are higher in villages exposed to signals from a larger number of community radio stations. The effect is identified based on a "natural experiment" in the northern communes of Benin where within-commune variation in village access to radio stations is exogenous to observed and unobserved village characteristics. In contrast to prior research, the authors find that this media effect does not operate through government accountability: government inputs into village schools and household knowledge of government education policies are no different in villages with greater access to community radio. Instead, households with greater access are more likely to make financial investments in the education of their children.

Do school feeding programs really work?

Despite the popularity and widespread implementation of school feeding programs, evidence on the impact of school feeding on school participation and nutritional status is mixed. Alison Buttenheim, Harold Alderman and Jed Friedman [3] evaluate school feeding programs in three northern districts of the Lao People's Democratic Republic (Lao PDR). Feeding modalities included on-site feeding, take-home rations, and a combination. District-level implementation of the intervention sites and selective take-up present considerable evaluation challenges. Addressing these problems as best they can, the authors find minimal evidence that the school feeding schemes increased enrollment or improved children's nutritional status.

The food crisis, household welfare and HIV/AIDS treatment

Using panel data from Mozambique collected in 2007 and 2008, Damien de Walque, Harounan Kazianga, Mead Over and Julia Vaillant [4] find that the food crisis led to a real deterioration of welfare in terms of income, food consumption, and nutritional status among both HIV and comparison households. However, HIV households did not suffer more from the crisis than others. Furthermore, initiation of treatment and better services in health facilities counter-balanced the effect of the crisis by improving the health of patients and their labor force participation. Finally, the authors find no effect of the change in welfare on the frequency of visits, but do find that people who experienced a negative income shock also experienced a reduction or a slower progression in treatment outcomes.

Crossing the threshold: an analysis of IBRD graduation policy

Graduation from the International Bank for Reconstruction and Development (IBRD) is not an automatic consequence of reaching a particular income level, but is supposed to be based on a determination of whether the country has reached a level of institutional development and capital-
market access that enables it to sustain its own development process without recourse to Bank funding. Using panel data for 1982 through 2008, Jac Heckelman, Steve Knack and Halsey Rogers [5] find that the observed correlates of Bank graduation are generally consistent with the stated policy. Countries that are wealthier, more creditworthy, more institutionally developed, and less vulnerable to shocks are more likely to have graduated. Predicted probabilities generated by the model correspond closely to the actual graduation and de-graduation experiences of most countries (such as Korea and Trinidad and Tobago), and suggest that Hungary and Latvia may have graduated prematurely—a prediction consistent with their subsequent return to borrowing from the Bank in the wake of the global financial crisis.

A practical comparison of the bivariate probit and linear IV estimators

Richard Chiburis, Jishnu Das and Michael Lokshin [6] present asymptotic theory and Monte-Carlo simulations comparing maximum-likelihood bivariate probit and linear instrumental variables (IV) estimators of treatment effects in models with a binary endogenous treatment and binary outcome. They clarify the relationship between the ‘average treatment effect’ obtained in the bivariate probit model and the ‘local average treatment effect’ estimated through linear IV. They compare the mean-square error and the actual size and power of tests based on these estimators across a wide range of parameter values relative to the existing literature. Finally, they assess the performance of misspecification tests for bivariate probit models. The authors recommend two changes to common practices: bootstrapped confidence intervals for both estimators, and a score test to check goodness of fit for the bivariate probit model.

New articles and books

Making schools work: New evidence on accountability reforms

A new World Bank book by Barbara Bruns, Deon Filmer and Harry Patrinos [7] explores threats to education quality in the developing world that cannot be explained by lack of resources. It reviews the observed phenomenon of service delivery failures in public education: cases where programs and policies increase the inputs to education but do not produce effective services where it counts—in schools and classrooms. It documents what we know about the extent and costs of such failures across low and middle-income countries. And it further develops the conceptual model posited in the World Development Report 2004: that a root cause of low-quality and inequitable public services—not only in education—is the weak accountability of providers to both their supervisors and clients.

No small matter: The impact of poverty, shocks, and human capital investments in early childhood development

Education is often seen as a fundamental means to improve economic prospects for individuals from low income settings. However, even with increased emphasis on basic education for all, many individuals fail to achieve basic skills to succeed in life. This new book, edited by Harold Alderman [8], presents evidence that one core reason is that by the time a child is old enough to attend school there is already a wide disparity in cognitive skills and in emotional and behavioral development among children from households of different socioeconomic backgrounds. Low levels of cognitive development in early childhood strongly correlate with low socio-economic status (as measured by wealth and parental education) as well as malnutrition. These disadvantages are often exacerbated by economic crises. As this book documents, however, there are programs that have proven effective in promoting a child’s development through caregiver-child interaction and stimulation, and with well designed preschool programs. Focused preschool programs can serve as a key investment in a strategy to reduce the transmission of poverty from poor parents to their children.

Migration and the overweight and underweight status of children in rural China

China’s rapid economic growth has been accompanied by a large-scale rural-to-urban migration, but over time more children are being left behind rural areas. This paper by Alan de Brauw and Ren Mu [9] studies how the overweight and underweight status of rural children is associated with the out-migration of others in their household. They find that migration is related to different nutritional outcomes for the left-behind children. Older children (aged 7–12) are more likely to be underweight; younger children (aged 2–6) are less likely to be overweight if left behind without the care of a grandparent. They also find evidence that the remaining adult household members spent less time preparing meals, whereas older children take up more household chores.
**Variations in condom use by type of partner in 13 sub-Saharan African countries**

Using nationally representative data from 13 sub-Saharan African countries, [Damien de Walque and Rachel Kline][10] reinforce and expand upon previous findings that men report using condoms more frequently than women do, and that unmarried respondents report that they use condoms with casual partners more frequently than married individuals report using them with their spouses. Based on descriptive, bivariate, and multivariate analyses, they also demonstrate to a degree not previously shown that married men from most countries report using condoms with extramarital partners about as frequently as unmarried men report using them with casual partners. Married women from most of the countries included in the study reported using condoms with extramarital partners less frequently than unmarried women reported using them with casual partners. This result is especially troubling because marriage usually ensures regular sexual intercourse, thereby providing more opportunities for a person to pass HIV infection from an extramarital partner to his or her spouse.

**Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting**

Using a randomized controlled trial design, [Markus Goldstein, Damien de Walque and 10 coauthors][11] set out to find whether short message service (SMS) reminders about adherence to ART among patients attending a rural clinic in Kenya could promote adherence to antiretroviral therapy (ART). 431 adult patients who had initiated ART within 3 months were enrolled and randomly assigned to a control group or one of the four intervention groups. Participants in the intervention groups received SMS reminders that were either short or long and sent at a daily or weekly frequency. Adherence was measured using the medication event monitoring system. 53% of participants receiving weekly SMS reminders achieved adherence of at least 90% during the 48 weeks of the study, compared with 40% of participants in the control group (P = 0.03). Participants in groups receiving weekly reminders were also significantly less likely to experience treatment interruptions exceeding 48 h during the 48-week follow-up period than participants in the control group (81 vs. 90%, P = 0.03).

**The cost of complying with human rights treaties: the convention on the rights of the child and basic immunization**

The determinants of compliance with human rights treaties likely vary according to the right in question, yet heterogeneity in the pathways through which ratification affects various human rights outcomes has received limited attention. [Varun Gauri][12] first develops an account of treaty compliance that incorporates the intrinsic benefits to the state of compliance, regime costs associated with certain rights, the political costs that NGOs, judges, and others are able to impose for non-compliance, and the fiscal and economic costs of compliance. He argues that for child survival rights, fiscal and economic costs are likely to be dispositive, and that as a result richer countries are more likely to comply. He then uses an instrumental variable approach to investigate whether ratification of the Convention of the Rights of the Child (CRC) was associated with stronger effort at the country level on child survival rights. He finds that ratification of the CRC was correlated with a subsequent increase in immunization rates, but only in upper middle and high income countries.

**Fungibility and the impact of development assistance: Evidence from Vietnam’s health sector**

The apparent fungibility of aid is a challenge to the evaluation of donor-funded development projects, requiring a comparison of the observed outcomes with the outcomes that would have occurred if the project had not gone ahead. Where projects are targeted on specific geographic areas, counterfactual outcomes in each can differ from observed outcomes because the amount of government spending (gross of aid) differs, the productivity of government spending differs, or both. [Adam Wagstaff][13] estimates the benefits of two concurrent World Bank health projects in Vietnam targeted on specific provinces. He derives estimates from a model linking outcomes (under-five mortality) to government spending before and after the project and in project and nonproject provinces, and are presented for different assumptions regarding fungibility of funds (zero and full fungibility) and the impacts of the project on the productivity of government spending (the project modifies productivity in both sectors equally and in neither sector). The estimated mortality reductions are highly insensitive to the assumed degree of fungibility, but highly sensitive to the assumed productivity effects (the estimates range from 1 to 25%). The wide range reflects the uncertainty due to the lack of a genuine control group of provinces.

**Research in the news**

Jishnu Das, Shantayay Devarajan, and Jeffrey Hammer, in a piece called **“Lost in Translation”** in a special series in the Boston Review on behavioral economics called
“Small Changes, Big Results”, remind readers that randomized control trials at best tell us what happens if a specific intervention occurs, not whether the government ought to finance let alone deliver the intervention. They argue that “because the experimental approach is based on what is feasible to randomize, it may lead us in the wrong direction. Experimental interventions must be completely under the control of the experimenter...” This narrows] the scope of possible activity to assessment of private goods, such as mosquito nets, chlorine pills, and school uniforms. Yet, those interventions easiest to control and measure may be the ones least in need of government effort.”

The work by Tahir Andrabi and Jishnu Das on “winning hearts and minds” in the aftermath of the Pakistan earthquake is picked up Tim Harford in his FT Undercover Economist column. Harford notes that while striking “it will be awkward news in some quarters”. He reports on a Foreign Policy article by Christopher Stokes, general director of Medecin sans Frontieres, who decries the notion of “politically useful” aid, arguing that “winning the trust of all parties in a conflict and gaining access to the affected population depends on being understood as purely humanitarian”.

And on the blogs

In a comment-provoking post on the Bank’s “Africa Can... End Poverty” blog, Damien de Walque asks: Is male promiscuity the main route of HIV/AIDS transmission in Africa? He concludes not, and that “the proportion of discordant female couples is difficult to explain unless women are also sexually active outside the marriage (or cohabiting union).”

Never at a loss for an eye-catching question, Damien de Walque, writing with Harounan Kazianga and Mead Over on the Bank’s “Let’s Talk Development” blog, asks: Could easier access to AIDS treatment increase risky sexual behaviors? The post cites new evidence from Mozambique where the authors find that increased access to therapy led to men engaging in more risky behaviors because they believe, wrongly, that AIDS can be cured, while women increase their risky behavior because they believe, correctly, that antiretroviral therapy can treat AIDS but cannot cure it.

In a post on the Bank’s “Governance for Development” blog, Stuti Khemani asks: Do informed citizens hold governments accountable? She concludes that it depends on the political and economic context within which media operates. In Benin, where she found it had no effect via accountability, the context neither encourages media to provide “accountability” information nor facilitates citizen mobilization in response to that information.

Writing on the Let’s Talk Development blog, Adam Wagstaff asks on the basis of Google Trends data: Is it time to rethink the development agency’s flagship report? Interest in all the agencies’ reports seems to have been waning, as has interest in international development. Meanwhile searches for “BMW” and “Mercedes” have stayed flat, as have searches for “Stiglitz”, “Eminem” and “Bono”. On the Global Dashboard blog, Alex Evans picks up the blog in a post entitled “International development: yesterday’s news?”

In another post on the Let’s Talk Development blog, Adam Wagstaff looks at the democratization of development economics. “The good news is that the developing world’s total and share of economics articles have risen over the last quarter century. The bad news is threefold: the increase is due almost entirely to growth in just three of the six Bank regions; the developing world has not seen any increase in “development and planning” articles; and one large region—South Asia—has seen a reduction in the number of economics and “development and planning” articles.”

Also on the Let’s Talk Development blog, Adam Wagstaff offers four cheers for the “results agenda”. It gets us to work back from outcomes to think about multiple ways of achieving better outcomes, including fixing the demand side. It invites us to focus on ways to incentivize results, whether we’re talking about rewarding clinics with better pay or rewarding countries with larger loans. It forces us to get serious about evidence, using impact evaluation but other methods too. And it invites us to think more seriously about the role of knowledge in the Bank’s work.

Jishnu Das in a post on the Bank’s Education for Global development blog, takes on the issue of corporal punishment and child abuse in schools. Das cites a survey by UNICEF in Pakistan a decade ago that “found that 88 percent of kids surveyed had been physically punished in the week before the survey – not one child reported never having been beaten. These beatings can be quite severe, often leading to serious injuries, and in some cases, death.”

DECRG’s HD work continues to be picked up by bloggers inside and outside the Bank. Steve Knack’s work on aid and the quality of tax administration is picked up by Amitai Etzioni in a piece entitled “Corruption reduction” in the online Harvard International Review. Varun Gauri’s work is picked up in a post “For the sake of fairness: Justice in development” on the Bank’s Governance for Development blog. Damien de
Walque’s work is picked up on a post on the JOURNALMEX blog. Tahir Andrabi and Jishnu Das’s work is picked up in a post on the Rug Pundits blog entitled “Doubt and development”. The new book by Barbara Bruns, Deon Filmer and Harry Patrinos is picked up by Ariel Fiszbein in a post on the Education for Global Development blog.

Linking research & operations: anti-corruption and regulatory reform in ECA

Researchers often assist regional staff in Economic and Sector Work, including not only country-specific products such as Poverty Assessments, but also regional “flagship” reports. Below, Steve Knack reflects on his productive collaboration over a 5-year period with ECA staff on anti-corruption issues.

In 1990, legal frameworks and other institutions necessary for market economies to function efficiently and equitably were missing or underdeveloped in the “transition” countries of ECA region, most of which experienced large declines in output in the 1990s. Privatization and adoption of new regulatory systems provided enormous opportunities for bribe-seeking and improper use of connections and influence. By the late 1990s, corruption had emerged as one of the most serious obstacles to growth in the region. Beginning in 2000, a series of reports on Anticorruption in Transition (ACT) has monitored progress on this problem, based on the Business Environment and Enterprise Performance Surveys (BEEPS) conducted every three years by the EBRD and ECA.

The most notable finding in the ACT series is a trend toward lower administrative corruption in the region overall. This headline message has often been obscured by seemingly conflicting evidence, however, and one of the main challenges in recent reports has been to sort through the proliferation of data sources and indicators to provide a clear and consistent story to policymakers. I played a supportive role on the third ACT report, based on the 2005 BEEPS, and was a lead author on the most recent report, based on the 2008 BEEPS.

Anticorruption in Transition 3

The first report (ACT-1) [14] was based only on the 1999 BEEPS, and set baselines. A favorable trend on administrative corruption was first noted in ACT-2 [15], based on comparisons from the 1999 and 2002 BEEPS. Some observers pointed out that this trend was not apparent in other well-known data sources on corruption, including the aggregate indexes produced by Transparency International (TI) and the Bank’s own Worldwide Governance Indicators (WGI) project. When ECA staff later discovered another favorable trend in their BEEPS 2002 and 2005 comparisons, they decided to pre-empt the skeptics by requesting me to write a background paper [16] on corruption indicators that might clear up any apparent inconsistencies. Results from the background paper, later published in the Journal of Public Policy [17], were summarized in a Box and cited elsewhere in the ACT-3 report [18]. It generally validated the favorable trend in administrative corruption found in the BEEPS, in making the following points:

- Different corruption indicators often measure different aspects of corruption, so less-than-perfect agreement among them is to be expected.
- This heterogeneity is reflected in the BEEPS itself, and its favorable trend in administrative corruption is not exhibited for “state capture”; moreover, administrative corruption in some areas (e.g. tax) improved more than in others (e.g. judicial systems).
- The TI and WGI indicators are aggregated from numerous sources of “expert” perceptions, and some sources do not provide an explicit definition of corruption provided. It is therefore unclear to what extent they “should” behave as measures of administrative corruption, “state capture,” or other types.
- Changes in experts’ perceptions over time are more subject to measurement error than changes in firms’ reported experiences. Often, changes in perceptions indicators represent corrections in response to newly-available information, rather than a belief that corruption has actually changed.
- Evidence from the World Economic Forum’s “Executive Opinion Surveys” was remarkably consistent with evidence from the BEEPS, in showing improvement in administrative corruption but not in state capture, and in showing more improvement in certain areas (licenses and permits, tax and customs, utilities) than in others (public procurement, judicial systems).

Trends in corruption and “regulatory burden,” 2005-2008

The sample design and questionnaire for the 2008 BEEPS were modified for conformity with the Bank’s other Enterprise Surveys. This change facilitated comparisons with other regions, but made it more difficult to track progress over time within ECA leading up to 2008.
With some input from other Research Department colleagues, I provided advice to ECA on several products based on the 2008 BEEPS, including:

- A dataset for 2005 and 2008 that maximizes comparability over time, mostly by dropping firms with certain characteristics (based on size or sectors) from one year or the other.
- A BEEPS at a Glance report and country profiles based on this dataset, posted on the ECA web site. The country profiles are particularly useful in showing which countries deviate from the overall regional trends.

More substantially, I was a lead author on a report [19] similar to those in the ACT series, assessing progress on corruption and regulatory burden, and providing some evidence on what reforms are more likely to be effective. A major goal in the report was to prevent misleading interpretations of data caused by changes in the way questions were asked. Two illustrations follow:

- The BEEPS appeared at first glance to indicate a doubling – from 6% in 2005 to 12% in 2008 - in the average time spent by managers in the region in dealing with public officials and government regulations in the region. However, this increase turned out to be an artifact of a change in question wording. The “time tax” question in 2008, but not in 2005, included a follow-up statement prompting respondents with examples that, in effect, invited them to interpret the question more broadly. A similar question in the WEF shows no adverse trend over time.
- The three ACT reports made extensive use of a question on the “bribe tax,” i.e. the share of firm revenues paid as bribes to public officials. In the 2005 and earlier BEEPS, used in those reports, respondents were all asked to state their answer as a percentage. In the 2008 survey, respondents were given a choice of (1) stating their answer as a percentage, or instead (2) providing an estimate of total bribes paid to officials in local currency units. For firms responding in this way, the bribe tax is computed using a separate survey question on firm sales in local currency units. The average bribe tax for firms using Method 1 is roughly 50 times as large as for firms using Method 2. Allowing the Method 2 option beginning in 2008 therefore likely creates a bias toward finding reductions in the bribe tax between 2005 and 2008. To correct for this bias, by estimating the average effect on responses of the change in question wording, it would be necessary to conduct a controlled experiment, e.g. in which half of the firms were randomly assigned to use Method 1 and the other half Method 2. In the 2008 survey, however, firms could choose for themselves which method to use, and their decisions were not random: firms reporting more frequent bribe paying (on a separate question) were more likely than other firms to answer the bribe tax question in terms of percentages – the method generating higher values. This selection bias – whereby firms facing more severe corruption problems chose Method 1 - implies that one could not merely drop the firms reporting by Method 2, and compute country aggregates from respondents reporting by Method 1, for comparisons with 2005.

For the next round of BEEPS, I have supported ECA staff in their attempt to add some questions on “state capture” back into the survey. These questions were all dropped for comparability with the other regions’ surveys, but the 2005 BEEPS and more recent WEF data indicate it is a continuing problem in ECA. I have also recommended conducting randomized experiments in countries with larger samples (e.g. Russia), to better understand the implications of asking questions (such as in the two examples above) in different ways.

References


Previous issues

HD Quarterly Previous Issues