

World Bank Research *Digest*

VOLUME 11 ★ NUMBER 2 ★ WINTER 2017

Estimating International Poverty Lines

Is the World Bank's international poverty line of \$1.90 a day both robust and relevant? Recent analysis suggests that it is

The share of people living in extreme poverty, as measured by the international poverty line estimated by the World Bank, has become one of the most prominent indicators for assessing progress in global economic development. It has been a central indicator for the Millennium Development Goals and is now an important one for the Sustainable Development Goals. Since the dollar-a-day poverty line was introduced in 1990, the guiding concept for estimating the international poverty line has been to base the line on a typical value of a subsample of national poverty lines for some of the world's poorest countries. The most recent World Bank international poverty line, \$1.90 per person per day, is based on the simple average of national poverty lines from the 15 poorest countries from a sample of 74 national poverty lines dating from 1985 to 2005.

The use of 15 national poverty lines that are more than 14 years old on average creates concerns for several reasons. One worry is that they provide weak support for the international poverty line: small changes in the composition of the 15-country reference group can result in large changes in the estimated international poverty line, altering the poverty status of tens of millions of people.

The age of the lines is also a concern. To state the value of these lines

in a common year, inflation data for several years are needed to update their values. The estimated value of \$1.90 (expressed in 2011 dollars adjusted for purchasing power parity, or PPP) requires an average of 14 years of inflation data from each of the 15 countries in the reference group, some of which have had lapses in the consistency of their inflation measures. Errors in inflation data can similarly change the poverty classification of tens of millions of people.

Another concern is whether national poverty lines from past decades are too frugal or outdated for assessing societies' views of basic needs. Are old poverty lines relevant for today's international poverty line?

In a recent paper Jolliffe and Prydz construct a new set of national poverty lines that help inform many issues related to global poverty measurement, including analysis on the relevance and robustness of the \$1.90 international poverty line. They derive a set of national poverty lines that are more recent, represent more countries, and directly correspond with national poverty counts as reported by the World Bank for monitoring the Millennium Development Goals and the Sustainable Development Goals.

The authors' approach is based on estimating national poverty lines by combining national poverty headcounts with corresponding consumption and income distributions as reported in the World Bank's PovcalNet database (a public research tool for international poverty estimation).

IN THIS ISSUE

Estimating International Poverty Lines	... page 1
Global Inequality between 1988 and 2008 and the "Elephant Graph"	... page 2
Challenges in Counting the World's Hungry	... page 3
Incentive Effects of Antipoverty Programs	... page 4
Importing High Food Prices by Exporting	... page 5
The Decision to Invest in Child Quality over Quantity	... page 6
Supporting Parents to Improve Preschool Outcomes	... page 7
Can Conditional Cash Transfers Improve Child Nutrition?	... page 8
Globalization, Comparative Advantage, and Fertility Decisions	... page 9
Experimental Justice Reform	... page 10
Recent World Bank Research Publications and Papers on Poverty and Inequality	... page 12

(continued on page 11)

Global Inequality between 1988 and 2008 and the “Elephant Graph”

The global growth incidence curve (or “elephant graph”) reveals the contrast between fast growth in Asia and slow growth in the West

A recent paper by Lakner and Milanovic presents new evidence on global inequality. The paper uses a newly compiled and improved database based on national household surveys from 1988 to 2008 and with data drawn mostly from the World Bank’s PovcalNet database and the Luxembourg Income Study. Every person in the world is assigned the average income of his or her income decile. Incomes are measured on a per capita basis and adjusted for differences in purchasing power across countries.

The data show that global income inequality remains high compared with inequality within countries: between 1988 and 2008 the global Gini index declined only slightly, from around 72 to 70. But global inequality may be even higher than this measure indicates. Household surveys often fail to measure top incomes accurately—and as a result, might underestimate inequality and fail to notice changes in inequality at the very top of the

distribution. The paper is the first analysis of global inequality that tries to adjust for missing top incomes. It approximates top incomes by the gap between national accounts consumption and survey means, in combination with a Pareto-type imputation of the upper tail. This results in an estimated global Gini index that is much higher, at nearly 76 in 2008, and the downward trend almost disappears.

Differences in average income between countries remain the most important source of global inequality. While this source of inequality has declined, thanks to the strong growth in some parts of Asia, it still explains around three-quarters of global inequality as measured by the Theil-L index. On the other hand, (population-weighted) inequality within countries has increased, both in levels and as a share of global inequality.

The aspect of the paper receiving most attention has been the anonymous global growth incidence curve (GIC) shown by the solid blue line in figure 1, which has come to be known as the “elephant graph.” The figure reflects the income growth rate for different percentiles of the global distribution between 1988 and 2008. Growth has been distributed such that it was

fastest for percentile groups just above the global median (at around 3 percent a year), followed by the global top 1 percent, while the 80th percentile saw very little growth.

Because of churning in the global distribution, the country composition of the global percentile groups tracked by the anonymous GIC can change over time. For example, many Asian decile groups have

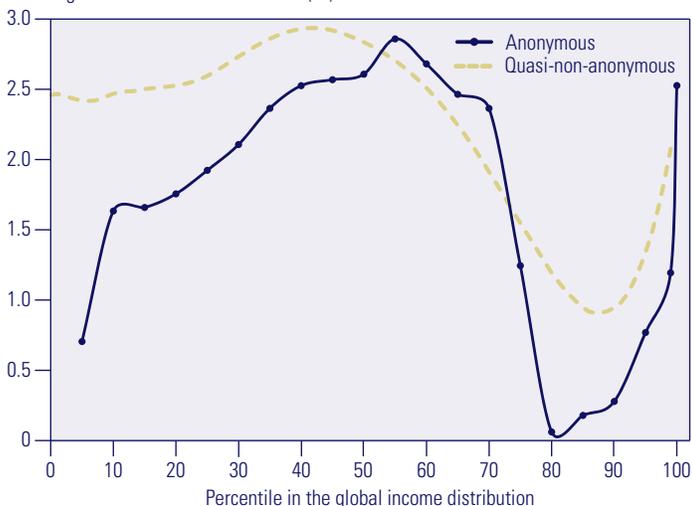
seen faster growth than the rest of the global distribution and thus occupy higher global percentiles in 2008 than in 1988. The quasi-non-anonymous GIC in figure 1 controls for such churning as well as for differences in population growth rates. Specifically, this curve keeps a country-decile fixed at its 1988 position and shows what growth it experienced over the next 20 years. Even after controlling for such churning in the global distribution, the growth rate dips around the 80–85th global percentile.

There were some 207 million people around the 80–85th percentile in 1988. About 172 million (more than 80 percent) were from today’s OECD countries. Even if the formerly Communist economies are dropped, there are still 145 million people (or 70 percent) from the “old rich” economies.

While figure 1 shows relatively large gains for the part of the distribution around the median, these gains are measured in relative (percentage) terms. Precisely because global incomes are distributed very unequally, the absolute gains are much greater for higher percentiles. While average per capita income increased by PPP\$25,000 for the top 1 percent between 1988 and 2008, it increased by only PPP\$400 at the global median. Thus 44 percent of the increase in global income between 1988 and 2008 went to the top 5 percent of the world population.

Figure 1. Growth Incidence, 1988–2008

Annual growth rate of real PPP income (%)



Note: The anonymous GIC shows the growth rate for each ventile (as well as the top 1 percent) of the global distribution, ignoring changes in country composition. For 63 countries observed in 1988 and 2008, the quasi-non-anonymous GIC plots growth for a country-decile against its position in the 1988 global distribution.

Christoph Lakner and Branko Milanovic. 2016. “Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession.” World Bank Economic Review 30 (2): 203–32. For recent discussion of the elephant graph, see <http://blogs.worldbank.org/developmenttalk/rebuttal-elephant-graph-discussion>.

Challenges in Counting the World's Hungry

Variations in the way household surveys measure food consumption lead to big differences in estimates of hunger

Global numbers on the prevalence of hunger are widely followed, not least because ending hunger is the second of the 17 Sustainable Development Goals. But both the standard approach to estimating these numbers and an alternative that has been increasingly promoted face measurement challenges. These raise questions about the cross-country validity of both approaches as currently practiced, according to findings in a recent paper by De Weerd, Beegle, Friedman, and Gibson.

The use of household survey data is fundamental in both approaches. The standard approach combines mean per capita food consumption derived from national food balance sheets (FBS) with estimates of the distribution of caloric intake (caloric variance) across households that are based on household survey data. The alternative relies solely on household consumption and expenditure surveys (HCES), measuring hunger directly as a function of the observed total household food consumption (in relation to estimates of household caloric need).

For both the FBS and HCES methods, one challenge is the wide

variation in the design of household consumption surveys around the world. The variation occurs in many dimensions, including these:

- The method of data capture, which is typically either diary or a recall questionnaire.
- The reference period over which consumption is measured, which can vary from one day to one week to one month.
- The degree of commodity detail, which can range from a handful of aggregate commodities to more than 400 relatively detailed commodities.

Using a consumption survey experiment in Tanzania, the authors study the implications of survey design for the measurement of hunger. It turns out that hunger estimates are even more sensitive to survey design than total consumption estimates are, because most of the design differences in global practice are due to differences in how food consumption is measured. For example, a review of consumption surveys from more than 100 countries finds that while the most common recall period for food consumption is seven days, the seven-day recall is used in only 31 percent of the surveys.

The study explores the net effect of reporting error arising from seven different, very common designs of consumption survey, including one that the authors consider to be the survey “gold standard”—an intensively

supervised personal diary with detailed cross-checks across household members. Because the study is a randomized within-village experiment that has good covariate balance and extends over an entire calendar year, the results are not subject to seasonal or community-level confounders. The only difference in population hunger numbers should be due to the particular survey design administered. And what a difference it is.

The average estimated daily per capita (kilo)calorie intake ranges from 1,794 to 2,677, depending on the survey module. As a consequence, the estimated prevalence of hunger, based solely on HCES data, ranges from 19 to 68 percent (table 1). This wide range in prevalence translates into a difference of more than 23 million people in Tanzania, a country with a population of 45 million according to the 2012 census.

And it's not just prevalence estimates that are influenced by survey design. The targeting of the hungry will also be affected. The hunger profiles suggested by each module differ because the hunger indicator varies significantly with household size, wealth, education, and other measures. For example, relative to the gold standard, recall surveys underestimate hunger as the household grows richer, but overestimate it as the household increases in size. Therefore, simple mean correction factors for each module would not be sufficient when comparisons need to be made across population groups with different characteristics. This study's analysis could potentially be used to make corrections, though it would be difficult to measure and control for all relevant characteristics. Perhaps the bigger challenge is to determine the extent to which any correction factor estimated from this study can apply to different contexts and countries (for example, settings with different income levels or staple foods).

Table 1. Survey Design and Estimates of Calorie Availability

Type of HCES design	Mean kilocalories per capita	Hunger prevalence (%)	
		Direct estimate from HCES data	Hybrid estimate from FBS and HCES data
Long list of 59 food items; 14-day recall	1,794	68	25
Long list of 59 food items; 7-day recall	2,129	48	23
Short list of 17 food items; 7-day recall ^a	2,066	48	23
Long list of 59 food items; usual 12-month recall	1,909	59	26
14-day household diaries with frequent visits	2,412	27	20
14-day household diaries with infrequent visits	2,517	23	21
14-day individual diaries with frequent visits	2,677	19	21

a. Because the 17 foods account for 77 percent of the food budget, calorie availability is scaled up by 1/0.77.

(continued on page 11)

Incentive Effects of Antipoverty Programs

Does China's targeted antipoverty program trap participants in poverty by creating disincentives for them to earn additional income?

Targeted social policies aimed at reducing poverty are widespread today, both among rich countries and, increasingly, among emerging middle-income economies. Concerns about their incentive effects have long been prominent. Do they create disincentives for participants to earn more income, leading to a poverty trap?

This long-standing debate is relevant to a major new antipoverty program in China. Introduced in the 1990s, the Minimum Livelihood Guarantee program—popularly known as the Dibao program—covered about 21 million people in urban areas and 54 million in rural areas by 2013.

The Dibao program is aimed at providing locally registered urban households whose per capita income falls below a predetermined local poverty line with a transfer payment sufficient to bring their income up to that line. The program's design implies that participants who have a small increase in nonprogram income would face an equal reduction in program receipts—a 100 percent benefit withdrawal rate (or marginal tax rate). Incentives to escape poverty would be weak or absent. But there are many reasons why the actual benefit withdrawal rates of an antipoverty program might differ from the nominal rate. In the Dibao program one possible reason is the scope for local discretion and innovation.

In a recent paper Ravallion and Chen study the Dibao program with the goal of assessing whether it has created a poverty trap—whether it operates in practice in the ways that its formal rules suggest. The paper focuses on the problem of estimating the mean benefit withdrawal rate, given by the average rate at which transfer receipts respond to differences in household income.

Most methods of calculating the benefit withdrawal rate found in

practice have either calculated the transfers or taxes implied by the formal rules or calculated conditional means of actual transfers or taxes at each level of net income. It is well recognized that behavioral responses can invalidate either method. The authors study the bias in statistical estimates of the benefit withdrawal rate induced by latent incentive effects and income measurement errors. They also identify a third source of bias (not previously discussed in the literature): correlated incidence heterogeneity. For example, on moral grounds program administrators may resist reducing the benefits of the poorest family when its income rises slightly. The extent to which this happens will naturally vary with the amount of local discretion.

The authors offer a new approach to estimating average benefit incidence that can be implemented with essentially the same data as prevailing methods of nonbehavioral benefit incidence analysis, but without ignoring incentive effects and measurement errors. In a departure from past work they focus directly on a key policy parameter, the benefit withdrawal rate. The key assumption is that incentive effects and classical measurement errors affect only certain lagged income components but that these still have predictive power for isolating exogenous variation in total income net of transfers or taxes. This justifies an instrumental variables estimator for the mean benefit withdrawal rate.

To implement this approach, the authors use a specially designed and commissioned survey to study the Dibao program. The results suggest a sizable bias in the benefit incidence picture that is implied by either the formal rules or the usual statistical practice of calculating conditional means at different net incomes. The estimated mean benefit withdrawal rate is much lower than the program's formal rules suggest, yet about twice that implied by the standard statistical approach.

By focusing on the key parameter for policy design, the authors also offer

some insights for policy reform in the light of the literature on optimal taxation. In theory, the Dibao program's design is ideal for protection from poverty but bad for promotion out of poverty because it imposes a 100 percent marginal tax rate on poor participants—a poverty trap. But the authors find no sign of this in the data. Indeed, the benefit withdrawal rate appears to be quite low when compared with the range of values suggested by optimal tax studies. Incentives for promotion appear to be strong, but performance in protection is weak.

The authors argue that the reason may be found in local implementation practices. Local agents implicitly put a far higher weight on promotion than implied by the program's design. But the data indicate heterogeneity in this, with richer cities tending to put higher weight on protection and poorer cities appearing to have a greater problem with participant capture. One possible explanation is that richer cities have more professional local administrators, less prone to participant capture.

The key policy conclusion is that the Dibao program is unlikely to provide a strong disincentive for participants to earn extra income. Incentive effects appear to be more serious than presumed by standard nonbehavioral benefit incidence analysis but still much less severe than basic incentive theory would suggest. Indeed, the findings suggest that reforms to the program should strive for a higher benefit withdrawal rate in local implementation, alongside expanded coverage.

Martin Ravallion and Shaohua Chen. 2015. "Benefit Incidence with Incentive Effects, Measurement Errors and Latent Heterogeneity: A Case Study for China." *Journal of Public Economics* 128: 124–32.

Importing High Food Prices by Exporting

In Lao PDR, which uses export controls to keep rice prices stable, prices are more likely to rise after a good harvest than a bad one

Rapidly rising staple food prices are a challenge in many developing countries: they directly affect poor households, which often spend more than half their income on food, and they increase the risk of food riots and political unrest. Rice trade policy is therefore often not guided by economic principles alone, and many decisions tend to be highly politicized.

In the Lao People's Democratic Republic, one of the least developed countries in Southeast Asia, glutinous (sticky) rice consumption provides almost 70 percent of calorie and protein intake. In 2010 the price of rice in the country increased by more than 50 percent, leading the government to impose a ban on the export of rice. The ban lasted only until February 2011, yet it was widely criticized by economists and multilateral organizations. Evaluating whether this criticism was justified or not requires a complete picture of how staple food prices are determined.

Lao PDR has an extensive set of controls on international rice trade, at both the national and provincial level. The central government has the responsibility for maintaining stable rice prices and can both impose trade bans and release stocks to prevent price increases. In general, however, decisions about international trade are made at the provincial level, though this is not allowed by the World Trade Organization. Because prices in neighboring countries are higher than in Lao PDR, a temporary removal of an export ban in one province is likely to generate both increased exports and interprovincial trade.

In a recent paper Durevall and van der Weide evaluate the drivers of glutinous rice prices and the reason for recent price spikes in Lao PDR, using data on prices, exports, and harvests. The focus is on the role of exports

to Thailand and Vietnam, the country's main trading partners. The study provides a number of insights. First, Lao PDR imports rice price changes in the long run by exporting its rice to Thailand and Vietnam. Second, rice prices in Lao PDR are considerably lower than in Thailand and Vietnam, despite the evidence of market integration. Third, and most interestingly, good harvests and large exports result in higher domestic prices later in the year. This is consistent with export controls being relaxed after good harvests, presumably to prevent sharp price falls, triggering a surge in exports early in the season that leaves stocks depleted ahead of the next harvest.

The use of export controls with the aim of maintaining stable and affordable prices is common in developing countries with a potentially exportable staple food. In Lao PDR this policy creates two regimes: a no-trade regime in which prices are determined primarily by domestic factors (there is also illegal trade) and a trade regime in which they are close to the (higher) international prices. When there is a good harvest and an opportunity to export, millers (the main traders) wish to sell as much as possible. Stocks therefore tend to decline fast in a surplus market because of the easing of controls, which produces a larger than normal price increase at the end of the season.

Why do millers not behave as if they expect a price spike at the end of the season and thus store rice during the year? One conjecture is that the scope for storage is limited by uncertainty resulting from a combination of segmented markets, the irregularity of the export opportunities, and the potential risk of government intervention. Millers get the opportunity to sell large quantities of rice abroad only occasionally, and because they do not know how long it will last, they sell as much as they can.

Lifting the trade restrictions permanently would then seem to be a natural policy option. The expectation is that this would raise domestic rice prices by reducing the price gap with

Thailand and Vietnam, and thus significantly increase incomes from agriculture and stimulate modernization of the sector. Indeed, there is emerging evidence that many poor households also become better off in the medium term as the benefits from the agricultural expansion eventually trickle down, likely in the form of greater demand for farm workers and higher rural wages. But the benefits of this indirect effect are of course uncertain. Some intermediate solution that protects the poor may therefore be advisable. Cash or food transfers are one policy option for protecting vulnerable households against rising food prices.

However, trade liberalization may not be politically feasible because of its distributional consequences, as evidenced by the number of governments that use trade policies to influence food prices. Adopting a variable export tax with transparent rules may be a viable second-best policy option. Ideally, once economic growth has reduced the dependence of most poor households on glutinous rice, export restrictions can be fully dismantled.

Dick Durevall and Roy van der Weide. 2016. "Importing High Food Prices by Exporting: Rice Prices in Lao PDR." *Journal of Agricultural Economics*. Published electronically August 4. doi:10.1111/1477-9552.12179.

The Decision to Invest in Child Quality over Quantity

Family size matters for household investment in education in Vietnam: parents with fewer children spend more per child on tutoring

Over the past four decades there has been much study of the relationship between household choices on the quantity and quality of children. The hypothesis driving the literature is that parents make tradeoffs between the number of children they bear and the “quality” of those children—shorthand for the amount of investment that parents make in their children’s human capital. If this hypothesis is true, it has implications for policies aimed at increasing economic growth and reducing poverty. For example, this could motivate policy makers to work on policies aimed at helping couples avoid unwanted births or subsidizing birth control.

A recent paper by Dang and Rogers investigates a new measure of households’ investment in their children: private tutoring (or extra classes) in mainstream subjects at school in which children are tested. Private tutoring is now widespread in many countries and may be an especially good measure of a household’s decision to invest voluntarily in children’s human capital—compared with enrollment, for example, which may also reflect exogenous factors such as compulsory schooling laws. Evidence indicates that private tutoring improves students’ academic performance in a number of countries.

Given the rapid expansion of educational attainment around the developing world, the tradeoffs that households make between the quantity and quality of children may increasingly manifest themselves outside the formal education system. Indeed, this is likely in developing and richer countries alike. In response to the widening inequality in the United States, for example, there have been calls for more summer and extracurricular programs boosting the skills of low-income students, to help level the playing field

between these students and their richer peers.

Moreover, in countries where the private-school sector is almost non-existent (at least at the pre-tertiary level), such as Vietnam, private tutoring represents a type of flexible household education investment that is most likely to be the equivalent of household investment in private education in other contexts. Vietnam also offers a particularly interesting case study because it has undergone rapid change in fertility and educational attainment, with fertility falling sharply at the same time that educational attainment has been rising rapidly. The government has paid much attention to family planning and has promulgated policies over the past 50 years encouraging (and, for government employees, requiring) families to restrict their number of children to one or two, but very few if any studies rigorously investigate the quantity-quality tradeoff for this country.

To identify the causal impacts of family size on household investment in private tutoring in rural Vietnam, the authors use as an instrument the distance from the household’s commune to the nearest family planning center. In contrast to those used in most previous studies, this instrumental variable allows them to study the effects of family size for families with one child or more. The authors also provide the most comprehensive empirical investigation to date of different aspects of household investment in private tutoring for each child.

On the conceptual front, the authors explicitly investigate the nexus between private tutoring and regular school—particularly the absolute and relative differences between household investments in regular school and those in private tutoring—which has not been examined before. They also slightly extend the standard Becker-Lewis quantity-quality tradeoff framework to provide further insights that can then guide the empirical analysis, which uses data from three rounds of the Vietnam Household

Living Standards Survey (2002, 2006, and 2008).

Estimation results indicate that each additional sibling reduces a rural household’s investments in a child’s schooling as measured through a variety of indicators: it reduces education expenditure by 0.4 standard deviations and tutoring expenditure by 0.5 standard deviations; it decreases the child’s probability of being enrolled in tutoring by 32 percentage points; and it cuts the average time spent on tutoring by 74 hours a year and the number of years of tutoring by 1.4. With respect to the differences between investments in tutoring and regular school, estimation results show that one more sibling reduces by 31 percentage points the probability of attending tutoring (unconditional on whether the child is enrolled in school or not); reduces by 243,000 dong (about \$15) the amount spent on education net of tutoring expenditure; and reduces by 8 percentage points the share of tutoring expenditure in education expenditure and by 20 percentage points the years attending tutoring as a share of completed years of schooling.

These results provide considerable support for the quantity-quality tradeoff in the Vietnamese context. The analysis also suggests that compared with traditional indicators such as enrollment, data on tutoring may be a more illuminating indicator of parents’ willingness to invest in the quality of education of their children and can offer a promising avenue of research in other contexts.

Hai-Anh Dang and Halsey Rogers. 2016. “The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam.” *World Bank Economic Review* 30 (1): 104–42.

Supporting Parents to Improve Preschool Outcomes

For informal preschools in resource-poor settings, integrating support of parenting can be more effective than simply improving classroom quality

Lack of adequate preparation is among the key risk factors for poor performance in primary school. In recognition of this, the Sustainable Development Goals call for all children, by 2030, to have access to quality early childhood development, care, and preprimary education so that they are ready for primary education.

A common approach to trying to improve children's outcomes centers on increasing enrollment in preschool programs or trying to improve the quality of existing programs. Research has shown that being assigned to higher-quality classrooms in kindergarten can modestly increase test scores among children. But an evaluation of Chile's *Un Buen Comienzo* (A Good Start), a large-scale, randomized study of an effort to improve the quality of preschool education, showed that while the program led to significant improvements in many classroom characteristics and teacher behaviors, it had no effects on children's language or literacy skills.

Another approach revolves around support of parents. A study in Mexico found positive effects on child development from group-based parenting support, as did a similar study in Colombia that used a home-visiting approach.

In a recent study using a cluster-randomized controlled trial, Özler, Fernald, Kariger, McConnell, Neuman, and Fraga tested the effectiveness of teacher training at informal schools in a resource-poor setting. The aim was to assess whether such school-based interventions are more effective in improving early childhood development and primary school readiness when combined with group-based parenting training.

The study focused on community-based childcare centers in Malawi,

which are estimated to serve more than half a million children in around 5,000 communities across the country. These centers generally have poor-quality facilities, lack basic play and learning materials, and operate for a few hours each weekday morning—and they are run by teachers who are typically untrained and unpaid and have a low level of education. To support early childhood development and learning, the government decided to improve the supply of play and learning materials in these centers and provide teachers with additional training and mentoring.

The study designed an experiment in which the control group received only a standard kit of supplies from UNICEF, while a second arm also received teacher training and mentoring. In a third arm the trained teachers were assigned to receive a small monthly stipend during the first school year following the intervention, intended to increase retention and motivation among these otherwise unpaid workers. In the fourth and final arm of the trial the school-based teacher training program was complemented by a 12-module, group-based, parenting education program for the primary caregivers of the children enrolled at the center. This model is less costly than stand-alone home-visiting programs, and because the newly trained teachers and mentors in the community deliver the group-based parenting training, it can be easily and inexpensively scaled up.

Primary child outcomes improved at the 18-month follow-up, but only in the treatment group receiving the integrated intervention—with teacher training and parenting education. Children in this group had significantly higher scores in an assessment of language skills, and they exhibited more pro-social behaviors than both the control group and the teacher-training-only group. The gains at the child level from the added parenting education were accompanied by substantial improvements in family care indicators, such as how many times a day primary

caregivers read to children or played with them.

Teacher training alone (or with monthly stipends for retention) did not improve children's outcomes, despite significant improvements in the classroom environment and teacher behaviors. Moreover, a rich battery of child assessments, conducted 36 months after baseline, showed no treatment effects among the six- to eight-year-old children in any treatment arm, indicating a substantial fade-out of program impacts in the integrated intervention.

The findings echo those of Chile's *Un Buen Comienzo*, which also found that improvements in classroom quality did not translate into improvements in child-level outcomes at the end of the two-year teacher training intervention. Both the study in Chile and the one in Malawi highlight the difficulty of converting program-induced improvements in classroom quality into better child outcomes.

The trial in Malawi combined the two main existing approaches to early childhood investments—preschool quality improvements and parenting support—which had not previously been tested together. It found promising evidence that this approach can improve children's outcomes over and above teacher training—at least in the short run.

In addition, the trial design allowed the study to identify the causal effect of exogenously improving classroom and parenting quality on child development outcomes. The analysis suggests that in this context, improving classroom quality had only a negligible impact while improving parenting quality had significant and large effects. The results suggest that group-based parenting support can have significant benefits for child development in an informal preschool setting, but that the early benefits faded over time.

Berk Özler, Lia C. H. Fernald, Patricia Kariger, Christin McConnell, Michelle Neuman, and Eduardo Fraga. 2016. "Combining Preschool Teacher Training with Parenting Education: A Cluster-Randomized Controlled Trial." *Policy Research Working Paper 7817*, World Bank, Washington, DC.

Can Conditional Cash Transfers Improve Child Nutrition?

Evidence from the Philippines suggests that conditional cash transfer programs can reduce severe malnutrition among children

Conditional cash transfer (CCT) programs are a central pillar of social protection policy in developing countries. These programs typically make the receipt of benefits conditional on such behaviors as use of health services or school attendance—with the aim of fostering income growth through investments in health and schooling while also addressing existing poverty.

In the two decades since Mexico's PROGRESA program introduced cash transfers linked to participation in health services for low-income households, the impact of such targeted social protection programs has been assessed on a range of outcomes. Virtually all those that have included incentives to use health services have led to greater participation in preventive health and nutrition activities. On average, however, these programs have not led to similar improvements in nutritional outcomes for children. In particular, they have generally failed to reduce stunting, a measure of chronic child malnutrition: an analysis of 17 programs that combined conditional and unconditional cash transfers shows an effect size that is neither statistically significant nor biologically meaningful.

There are many possible reasons for this disconnect between the success of CCT programs in improving health care utilization and their limited impact on nutritional outcomes for young children. One reason is that many evaluations have focused on programs in Latin American countries with a relatively low prevalence of stunting or underweight. Another is that the evaluations often are conducted for a relatively short period and therefore cannot capture cumulative effects. In addition, most studies include all children under age five and thus include both those in the period of greatest risk

of growth faltering—that is, children under age two—and those who may be less responsive to interventions at health care facilities. Including older children, and therefore those whose height is less responsive to nutritional interventions, may mask the program impact when the period of study does not include the first months of the child's life. But the issue is not only one of research design; the quality of health services may also limit the extent to which the use of services translates into a nutritional impact.

In a recent paper Kandpal, Alderman, Friedman, Filmer, Onishi, and Avalos use data from an experimental evaluation of the Philippine CCT program to explore the effects of CCTs on child nutrition. They focus on children in the vulnerable age group of 0–3 years in a high-malnutrition setting. The Philippine program provides cash transfers to poor households with children ages 0–14 conditional on investments in child health and education as well as use of maternal health services. Eligible households that satisfy compliance conditions receive a combination of health grants and education grants every two months. The grants range from \$11 to \$32, depending on the number of eligible children in the household. The maximum monthly transfer represents 23 percent of beneficiaries' household incomes.

The authors estimate treatment effects in 65 treated and 65 control villages, using data collected after 31 months of implementation. To capture the range of potential impacts, they analyze a number of outcomes: standardized measures of child height-for-age and weight-for-age; stunting and severe stunting, which reflect deficits in height-for-age; and underweight and severely underweight status, which reflect low weight-for-age. They also assess the impact on perinatal care, institutional delivery, presence of skilled birth attendants, breastfeeding practices, immunization, growth monitoring and deworming, care-seeking, and children's intake of foods that are rich in protein.

Results show that the CCT program led to a 15 percent improvement in height-for-age and a 40 percent reduction in severe stunting in children. In considering some behavioral changes that may explain these gains in nutritional outcomes, the authors document improvements in the utilization of perinatal care, though no changes in the rate of institutional deliveries, presence of skilled birth attendants, or breastfeeding practices. In addition, poor children ages 0–3 in treated villages were more likely to have received age-appropriate health services, including growth monitoring and immunizations, in the previous six months. Infant and young child feeding practices also appear to have been affected, with young children in treated areas being more likely to consume eggs and dairy than children in control areas.

The Philippine CCT program may have been effective at improving child nutrition for several reasons. Increased income from the cash transfer may have facilitated an improvement in diet diversity, while the program conditions, by promoting timely perinatal care and regular growth monitoring and health service use, may have led to increased diet diversity and clinical counseling. Finally, the information provided through the enforcement of program conditions may have combined with the additional cash to result in greater investments in children's welfare and improved health and sanitation.

Eeshani Kandpal, Harold Alderman, Jed Friedman, Deon Filmer, Junko Onishi, and Jorge Avalos. 2016. "A Conditional Cash Transfer Program in the Philippines Reduces Severe Stunting." *Journal of Nutrition* 146 (9): 1793–800.

Globalization, Comparative Advantage, and Fertility Decisions

Countries that have a comparative advantage in sectors with a greater demand for female labor also have lower fertility

Attempts to understand population growth and the determinants of fertility date as far back as the late 1700s, when Thomas Malthus wrote *An Essay on the Principle of Population*. In more recent times, postulating that fertility decisions are influenced by women's opportunity cost of time, economists have incorporated choice over fertility into growth models in order to understand the joint behavior of population and economic development throughout history. Most analyses examine individual countries in a closed-economy setting. But in an era of ever-increasing integration of world markets, the role of globalization in determining fertility can no longer be ignored.

In a recent study Do, Levchenko, and Raddatz examine both theoretically and empirically the impact of comparative advantage in international trade on fertility outcomes. The authors' conceptual framework is based on several observations. First, women bear a disproportionate share of the burden in raising children. While there has been a debate on whether this empirical regularity is due to nature or nurture, the outcome remains that having a child reduces a woman's labor supply more than a man's. Second, industries vary a great deal in the gender mix of their labor force: some industries employ primarily women, others primarily men. For example, female workers account for as little as 8–9 percent of employment in industries such as heavy machinery, and as much as 60–70 percent in some types of textiles and apparel. Third, differences in technology and resource endowments will lead to some countries having a comparative advantage in industries that happen to hire a lot of women.

All this being so, countries with a comparative advantage in female-intensive goods will exhibit lower

fertility. The authors' main thesis thus combines the hypothesis that fertility is affected by women's opportunity cost of time with the insight that this opportunity cost is higher in countries with a comparative advantage in industries that have a higher demand for female labor.

The authors provide empirical evidence for this phenomenon using industry-level export data for 61 manufacturing sectors in 145 countries over five decades. They combine the classification of industries according to their intensity of female labor use with data on countries' export shares to construct, for each country and time period, a measure of its *female labor needs of exports*. This index captures the degree to which a country's comparative advantage is in female-intensive sectors. The measure is used to show empirically that fertility is lower in countries with a comparative advantage in female-intensive sectors.

One way to provide a simple illustration of the results is to compare countries with the largest positive and negative changes in female labor needs of exports between the 1960s and today. For the countries with the largest observed increases in female labor needs of exports, the common pattern is that they change their specialization from agriculture-based sectors to wearing apparel. In Cambodia, for example, 80 percent of exports in the 1960s were in food products sectors. By the 2000s, 85 percent of Cambodian exports were in wearing apparel. Since food products sectors are right in the middle of the distribution of female intensity, and wearing apparel is the most female-intensive sector, this type of change in specialization will lead to large increases in female labor needs of exports.

The largest observed decreases in female labor needs of exports are driven by the discovery of natural resources. For example, Niger was an agricultural exporter in the 1960s, with nearly 80 percent of exports in food products. By the 2000s more than 60 percent of Niger's exports were in

refined petroleum products and nuclear fuel. The natural-resource-based sectors are among the least female intensive, with a share of female workers in total employment of 11–13 percent. This explains why countries with major shifts toward natural resources show a reduction in their female labor needs of exports.

This pattern is more general. Analysis controlling for income, overall trade openness, and regional location shows a pronounced negative relationship between fertility and female labor needs of exports.

Fertility is an economic decision, and like all economic decisions it has long been considered an appropriate—and important—subject of analysis by economists. It is now clear that international trade—or, more precisely, comparative advantage—matters for fertility decisions. The authors' results emphasize the heterogeneity of the effects of trade on countries' industrial structures and gender outcomes. From a policy perspective, the results suggest that it will be more difficult for countries with technology-based comparative advantage in male-intensive goods to undertake policy measures to reduce the gender gap in labor market conditions, potentially leading to a slower pace of women's empowerment. In an increasingly integrated global market, the road to female empowerment is paradoxically very specific to each country's productive structure and exposure to international trade.

Experimental Justice Reform

A new, experimental approach to justice reform seeks to foster inclusive decision-making and pragmatic, locally based solutions

Adam Smith famously declared in 1755, two decades before the publication of *The Wealth of Nations*, that little more was required “to carry a state to the highest degree of opulence from the lowest barbarism but peace, easy taxes and a tolerable administration of justice...” More recently a remarkably broad set of actors, across the political spectrum, have endorsed building and maintaining “the rule of law” as a primary issue facing every country. Indeed, in 2010 then World Bank president Robert Zoellick argued at a major conference that “the rule of law must be at the center of the development agenda.”

One might imagine that a high-priority development goal ostensibly uniting an otherwise unlikely assortment of constituencies would be grounded in a robust foundation of theory, method, and evidence—or at least a steadily expanding record of practical achievement. If anything, however, the law and development field has been beset from the outset by trenchant critique of its dominant approach, “legal transplanting,” in which the forms of institutions, codes, and procedures from one (usually high-income) country are introduced into another (usually low-income) country. Faced with the challenge of building or enhancing a legal system in the aftermath of a civil conflict, for example, orthodoxy as funded by international donors may focus on constitutional reform, redrafting legal codes, overseeing elections, building courthouses and jails, training police officers and prosecutors, and upgrading administrative systems (for example, providing the latest software for tracking cases).

All these tasks, importantly, can be funded, procured, and implemented in ways entirely consistent with bureaucratic imperatives back in the capital cities of donor countries. They will

reliably generate a seemingly impressive array of discrete “deliverables” that can be readily photographed, counted, tracked, aggregated, and compared. From both an accounting and an accountability perspective, such tasks raise few red flags and, after a time, enable senior managers to present a coherent, even compelling, narrative to skeptical politicians and voters of how public resources were spent in sensible ways in response to a clear and present development challenge.

According to an assessment by leading analysts of the field, however, the actual achievements of legal transplanting are fragile at best. Echoing misgivings dating back to at least the 1970s, a high-profile review in 2008 argued that thinking of law as a readily transferable technology is a serious error—because law consists of institutions deeply embedded in particular political, economic, and social contexts.

Since roughly 2000 an array of different organizations have sought to heed calls to explore a completely different approach to justice reform—an approach that Desai and Woolcock, in a recent article, term “new experimentalism.” For these organizations the aspirational goal has been to explore the experience of justice-seeking from diverse perspectives, incorporating in particular the experiences and aspirations of users (for example, those seeking to access a prevailing justice system), mid-level public authority, sovereign administrative authority, and transnational private interests; to help forge and protect a political space in which locally generated evidence can inform (more) equitable contests between those representing these multiple voices; and to engage in iterative problem-solving (as opposed to solution-selling) in response to concerns nominated and prioritized by these users (as opposed to foreign experts).

New experimentalist approaches to justice reform thus diverge from orthodoxy in two key ways: they steward an inclusive process of decision-making

around inherently contentious issues (such as land and natural resources), and they seek to ensure its equitable functioning through the use of regular feedback grounded in locally generated data. New experimentalism incorporates a pragmatic orientation when describing the nature of its participants but also positions the development practitioner as relatively nonideological (with some commitment to liberal process norms); consistently self-critical; and able to produce different kinds of information and translate it into politically salient action for different kinds of audiences. Such practitioners are embedded “organizational entrepreneurs” rather than distant technical specialists.

A leading instantiation of this alternative approach is the World Bank’s Justice for the Poor program (J4P). Now 14 years old, J4P can claim significant achievements in terms of its global impact on the theory and practice of justice reform, seeking more to influence at scale rather than operate at scale (such as through its contribution to numerous *World Development Reports* and the work of other advocacy groups such as the World Justice Project). It remains to be seen whether these initial gains can be more formally routinized into everyday practice within and beyond the World Bank, but doing so will entail remaining true to its own founding principles of regular self-critique.

Deval Desai and Michael Woolcock. 2015. “Experimental Justice Reform: Lessons from the World Bank and Beyond.” *Annual Review of Law and Social Science* 11: 155–74.

(continued from page 1)

Estimating International Poverty Lines

Because PovcalNet harmonizes its measures of well-being in per capita terms, the resulting national poverty lines are also in per capita terms. Matching the poverty lines with reported poverty headcounts and harmonizing the lines in per capita terms helps to overcome comparability problems with previous databases of national poverty lines.

The resulting database provides 864 national poverty lines covering 129 countries, more than a tenfold increase over the database of national poverty lines used to estimate the \$1.90 line. Most of these 864 national poverty lines, and most of the countries, are bunched together at relatively low values. Of all the poverty lines, 37 percent are less than \$3 a day and 52 percent are less than \$5 a day. Of those that are less than \$5 a day, there is a noticeable mass near \$1.90.

To assess concerns about inflation adjustments and the continued

relevance of \$1.90 as an estimate of basic needs, the authors subsample the set of national poverty lines closest in time to 2011 (the PPP reference year). This results in a sample of national poverty lines from 115 countries, which require on average only one year of inflation data to be expressed in 2011 currency. Following the same guiding principle of basing the international poverty line on a typical value of a subsample of the lowest of national poverty lines, the authors examine two subsamples of national poverty lines—one from countries classified as low income and the other from countries in the bottom quartile (based on a measure of national consumption). The median national poverty line from the bottom quartile is \$1.86, and the median from the low-income subsample is \$1.91.

Despite the use of different approaches for selecting the reference subsample, and of a much more

complete and recent set of national poverty lines, both estimates result in an international poverty line that directly corresponds to the World Bank definition of extreme poverty. This finding suggests that the current international poverty line of \$1.90 is robust to measurement concerns and continues to be a relevant marker of extreme poverty in the poorest of countries.

Dean Jolliffe and Espen Beer Prydz. 2016. "Estimating International Poverty Lines from Comparable National Thresholds." *Journal of Economic Inequality* 14 (2): 185–98. doi:10.1007/s10888-016-9327-5.

(continued from page 3)

Challenges in Counting the World's Hungry

The focus so far has been on the HCES method, but the global hunger estimates derived from the FBS method are also subject to the vagaries of cross-country survey design. Recall that in this method, while the mean is taken from the national food balance sheets, the variance of calorie consumption is measured through household consumption and expenditure surveys. There are numerous grounds for questioning the accuracy of the balance sheets, including the degree to which they capture national postharvest stores and losses and the accuracy of their root crop yields. But it is clear that variation in survey

design (as seen in table 1) will also affect the FBS-derived estimates: estimates of the prevalence of hunger range from 20 to 26 percent on the basis of survey-design-driven differences in the estimated variance of the calorie distribution.

HCES data are ubiquitous and hold the potential for being a useful complement to existing measures of hunger. Their appropriateness for measuring hunger could be improved through more thoughtful harmonization of survey design. But until more is done to undertake these design changes, and to understand the various sources of error and how they differ between

survey methods and, more broadly, between the HCES method and the hybrid FBS-HCES method, caution is warranted in drawing inferences from comparisons of survey-based hunger estimates over time and space.

Joachim De Weerd, Kathleen Beegle, Jed Friedman, and John Gibson. 2016. "The Challenge of Measuring Hunger through Survey." *Economic Development and Cultural Change* 64 (4): 727–58.

Recent World Bank Research Publications and Papers on Poverty and Inequality

Chapters in Books

Andrews, Matt, Lant Pritchett, Salimah Samji, and Michael Woolcock. 2015. "Building Capability by Delivering Results: Putting Problem-Driven Iterative Adaption (PDIA) Principles into Practice." In *A Governance Practitioner's Notebook: Alternative Ideas and Approaches*, edited by Alan Whaites, Eduardo Gonzalez, Sara Fyson, and Graham Teskey, 123–33. Paris: Organisation for Economic Co-operation and Development.

D'Souza, Anna, and Dean Jolliffe. 2016. "Coping with Food Price Shocks in Afghanistan." In *Food Price Volatility and Implications for Food Security and Policy*, edited by Matthias Kalkuhl, Joachim von Braun, and Maximo Torero, 543–81. Springer.

Ferreira, Francisco, Sergio Firpo, and Julian Messina. 2016. "Understanding Recent Dynamics of Earnings Inequality in Brazil." In *New Order and Progress: Development and Democracy in Brazil*, edited by Ben R. Schneider. Oxford: Oxford University Press.

Ferreira, Francisco, and Vito Peragine. 2016. "Individual Responsibility and Equality of Opportunity." In *Oxford Handbook of Well-Being and Public Policy*, edited by Matthew Adler and Marc Fleurbaey. Oxford: Oxford University Press.

Journal Articles

Doudich, Mohamed, Abdeljaouad Ezrari, Roy van der Weide, and Paolo Verme. 2015. "Estimating Quarterly Poverty Rates Using Labor Force Surveys: A Primer." *World Bank Economic Review* 30 (3): 475–500.

Ferreira, Francisco, Shaohua Chen, Andrew Dabalen, Yuri Dikhanov, Nada Hamadeh, Dean Jolliffe, Ambar Narayan, Espen Prydz, Ana Revenga, Prem Sangraula, Umar Serajuddin, and Nobuo Yoshida. 2016. "A Global Count of the Extreme Poor in 2012: Data Issues, Methodology and Initial Results." *Journal of Economic Inequality* 14 (2): 141–72.

Ferreira, Francisco, Nora Lustig, and Daniel Teles. 2015. "Appraising Cross-National Income Inequality Databases: An Introduction." *Journal of Economic Inequality* 13 (4): 497–526.

Working Papers

Akresh, Richard, Damien de Walque, and Harounan Kazianka. 2016. "Evidence from a Randomized Evaluation of the Household Welfare Impacts of Conditional and Unconditional Cash Transfers Given to Mothers or Fathers." Policy Research Working Paper 7730, World Bank, Washington, DC.

Ali, Daniel Ayalew, Klaus Deininger, and Marguerite Duponchel. 2016. "Using Administrative Data to Assess the Impact and Sustainability of Rwanda's Land Tenure Regularization." Policy Research Working Paper 7705, World Bank, Washington, DC.

Brück, Tilman, Patricia Justino, Philip Verwimp, and Andrew Tedesco. 2016. "Measuring Violent Conflict in Micro-Level Surveys: Current Practices and Methodological Challenges." Policy Research Working Paper 7585, World Bank, Washington, DC.

Brunori, Paolo, Flaviana Palmisano, and Vito Peragine. 2016. "Inequality of Opportunity in Sub-Saharan Africa." Policy Research Working Paper 7782, World Bank, Washington, DC.

Deininger, Klaus, Hari K. Nagarajan, and Sudhir K. Singh. 2016. "Short-Term Effects of India's Employment Guarantee Program on Labor Markets and Agricultural Productivity." Policy Research Working Paper 7665, World Bank, Washington, DC.

Ferreira, Francisco, Sergio Firpo, and Antonio F. Galvão. 2017. "Estimation and Inference for Actual and Counterfactual Growth Incidence Curves." Policy Research Working Paper 7933, World Bank, Washington, DC.

Friedman, Jed, Kathleen Beegle, Joachim De Weerd, and John Gibson. 2016. "Decomposing Response Errors in Food Consumption Measurement: Implications for Survey Design from a Survey Experiment in Tanzania." Policy Research Working Paper 7646, World Bank, Washington, DC.

Fujii, Tomoki, and Roy van der Weide. 2016. "Is Predicted Data a Viable Alternative to Real Data?" Policy Research Working Paper 7841, World Bank, Washington, DC.

Gaddis, Isis. 2016. "Prices for Poverty Analysis in Africa." Policy Research Working Paper 7652, World Bank, Washington, DC.

Gill, Indermit S., Ana L. Revenga, and Christian Zeballos. 2016. "Grow, Invest, Insure: A Game Plan to End Extreme Poverty by 2030." Policy Research Working Paper 7892, World Bank, Washington, DC.

Goraus, Karolina, and Gabriela Inchauste. 2016. "The Distributional Impact of Taxes and Transfers in Poland." Policy Research Working Paper 7787, World Bank, Washington, DC.

Hallegatte, Stephane, Mook Bangalore, and Adrienne Camille Vogt-Schilb. 2016. "Assessing Socioeconomic Resilience to Floods in 90 Countries." Policy Research Working Paper 7663, World Bank, Washington, DC.

Jirasavetakul, La-Bhus Fah, and Christoph Lakner. 2016. "The Distribution of Consumption Expenditure in Sub-Saharan Africa: The Inequality among All Africans." Policy Research Working Paper 7557, World Bank, Washington, DC.

Lakner, Christoph. 2016. "Global Inequality: The Implications of Thomas Piketty's *Capital in the 21st Century*." Policy Research Working Paper 7776, World Bank, Washington, DC.

Marrero, Gustavo A., Juan Gabriel Rodriguez, and Roy van der Weide. 2016. "Unequal Opportunity, Unequal Growth." Policy Research Working Paper 7853, World Bank, Washington, DC.

Narayan, Ambar, and Rinku Murgai. 2016. "Looking Back on Two Decades of Poverty and Well-Being in India." Policy Research Working Paper 7626, World Bank, Washington, DC.

Narloch, Ulf Gerrit, and Mook Bangalore. 2016. "Environmental Risks and Poverty: Analyzing Geo-Spatial and Household Data from Vietnam." Policy Research Working Paper 7763, World Bank, Washington, DC.

Newhouse, David, Pablo Suarez-Becerra, and Martin C. Evans. 2016. "New Estimates of Extreme Poverty for Children." Policy Research Working Paper 7845, World Bank, Washington, DC.

Pave Sohnesen, Thomas, and Niels Stender. 2016. "Is Random Forest a Superior Methodology for Predicting Poverty? An Empirical Assessment." Policy Research Working Paper 7612, World Bank, Washington, DC.

Skoufias, Emmanuel. 2016. "Synergies in Child Nutrition: Interactions of Food Security, Health and Environment, and Child Care." Policy Research Working Paper 7794, World Bank, Washington, DC.

van der Weide, Roy, Christoph Lakner, and Elena Ianchovichina. 2016. "Is Inequality Underestimated in Egypt? Evidence from House Prices." Policy Research Working Paper 7727, World Bank, Washington, DC.

Verme, Paolo. 2016. "Subsidy Reforms in the Middle East and North Africa Region: A Review." Policy Research Working Paper 7754, World Bank, Washington, DC.

Wydick, Bruce, Paul Glewwe, and Laine Rutledge. 2016. "Does Child Sponsorship Pay Off in Adulthood? An International Study of Impacts on Income and Wealth." Policy Research Working Paper 7563, World Bank, Washington, DC.

Working Papers can be downloaded at <http://econ.worldbank.org>
To download the *World Bank Research E-Newsletter*, go to http://econ.worldbank.org/research_newsletter

The World Bank Research Digest is a quarterly publication disseminating findings of World Bank research. The views and interpretations in the articles are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

The Research Digest is financed by the Bank's Research Committee and managed by DECDP, the research support unit of the Development Economics Senior Vice Presidency (DEC). The Research Digest is not copyrighted and may be reproduced with appropriate source attribution.

Editorial Committee: Chorching Goh (managing editor), Aslı Demirgüç-Kunt, and Shiva S. Makkı. Editor: Alison Strong; production: Roula Yazigi. For information or free subscriptions, send email to research@worldbank.org or visit http://econ.worldbank.org/research_digest.



The World Bank
1818 H Street, NW
Washington, DC 20433, USA



Printed on Recycled Paper