

THE WORLD BANK

Research Observer

EDITOR

Emmanuel Jimenez, World Bank

CO-EDITOR

Luis Servén, World Bank

EDITORIAL BOARD

Harold Alderman, International Food

Policy Research Institute

Barry Eichengreen, University of California-Berkeley

Marianne Fay, World Bank

Jeffrey S. Hammer, Princeton University

Ravi Kanbur, Cornell University

Ana L. Revenga, World Bank

Ann E. Harrison, University of Pennsylvania

The World Bank Research Observer is intended for anyone who has a professional interest in development. *Observer* articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the *Observer* considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the Editor. Please direct all editorial correspondence to the Editor, *The World Bank Research Observer*, 1818 H Street, NW, Washington, DC 20433, USA.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

For more information, please visit the Web sites of the *Research Observer* at www.wbro.oxfordjournals.org, the World Bank at www.worldbank.org, and Oxford University Press at www.oxfordjournals.org.

Subscriptions

A subscription to *The World Bank Research Observer* (ISSN 0257-3032) comprises 2 issues. Prices include postage; for subscribers outside the Americas, issues are sent air freight.

Annual Subscription Rate (Volume 29, 2 issues, 2014)

Academic libraries

Print edition and site-wide online access: US\$220/£147/€220

Print edition only: US\$202/£134/€202

Site-wide online access only: US\$177/£117/€177

Corporate

Print edition and site-wide online access: US\$334/£221/€334

Print edition only: US\$306/£203/€306

Site-wide online access only: US\$268/£177/€268

Personal

Print edition and individual online access: US\$61/£41/€61

Please note: US\$ rate applies to US & Canada, Euros€ applies to Europe, UK£ applies to UK and Rest of World.

Readers with mailing addresses in non-OECD countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request by writing to the UK address below.

There may be other subscription rates available; for a complete listing, please visit www.wbro.oxfordjournals.org/subscriptions.

Full pre-payment in the correct currency is required for all orders. Payment should be in US dollars for orders being delivered to the USA or Canada; Euros for orders being delivered within Europe (excluding the UK); GBP sterling for orders being delivered elsewhere (i.e., not being delivered to USA, Canada, or Europe). All orders should be accompanied by full payment and sent to your nearest Oxford Journals office. Subscriptions are accepted for complete volumes only. Orders are regarded as firm, and payments are not refundable. Our prices include Standard Air as postage outside of the UK. Claims must be notified within four months of despatch/order date (whichever is later). Subscriptions in the EEC may be subject to European VAT. If registered, please supply details to avoid unnecessary charges. For subscriptions that include online versions, a proportion of the subscription price may be subject to UK VAT. Subscribers in Canada, please add GST to the prices quoted. Personal rate subscriptions are only available if payment is made by personal cheque or credit card, delivery is to a private address, and is for personal use only.

Back issues: The current year and two previous years' issues are available from Oxford University Press. Previous volumes can be obtained from the Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. E-mail: psc@periodicals.com. Tel: (518) 537-4700. Fax: (518) 537-5899.

Contact information: Journals Customer Service Department, Oxford University Press, Great Clarendon Street, Oxford OX2 6DP, UK. E-mail: jnls.cust.serv@oup.com. Tel: +44 (0)1865 353907. Fax: +44 (0)1865 353485.

In the Americas, please contact: Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513, USA. E-mail: jnlorders@oup.com. Tel: (800) 852-7323 (toll-free in USA/Canada) or (919) 677-0977. Fax: (919) 677-1714. **In Japan, please contact:** Journals Customer Service Department, Oxford University Press, 4-5-10-8F Shiba, Minato-ku, Tokyo, 108-8386, Japan. E-mail: custserv.jp@oup.com. Tel: +81 3 5444 5858. Fax: +81 3 3454 2929.

Postal information: *The World Bank Research Observer* (ISSN 0257-3032) is published twice a year, in February and August, by Oxford University Press for the International Bank for Reconstruction and Development/THE WORLD BANK. Postmaster: send address changes to *The World Bank Research Observer*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Periodicals postage paid at Cary, NC and at additional mailing offices. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Research Observer*, The World Bank, 1818 H Street, NW, Washington, D.C. 20433, USA.

Environmental and ethical policies: Oxford Journals, a division of Oxford University Press, is committed to working with the global community to bring the highest quality research to the widest possible audience. Oxford Journals will protect the environment by implementing environmentally friendly policies and practices wherever possible. Please see <http://www.oxfordjournals.org/ethicalpolicies.html> for further information on environmental and ethical policies.

Digital Object Identifiers: For information on dois and to resolve them, please visit www.doi.org.

Permissions: For information on how to request permissions to reproduce articles or information from this journal, please visit www.oxfordjournals.org/jnls/permissions.

Advertising: Advertising, inserts, and artwork enquiries should be addressed to Advertising and Special Sales, Oxford Journals, Oxford University Press, Great Clarendon Street, Oxford, OX2 6DP, UK. Tel: +44 (0)1865 354767; Fax: +44(0)1865 353774; E-mail: jnlsadvertising@oup.com.

Disclaimer: Statements of fact and opinion in the articles in *The World Bank Research Observer* are those of the respective authors and contributors and not of the International Bank for Reconstruction and Development/THE WORLD BANK or Oxford University Press. Neither Oxford University Press nor the International Bank for Reconstruction and Development/THE WORLD BANK make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability for any errors or omissions that may be made. The reader should make her or his own evaluation as to the appropriateness or otherwise of any experimental technique described.

Paper used: *The World Bank Research Observer* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

Indexing and abstracting: *The World Bank Research Observer* is indexed and/or abstracted by ABI/INFORM, CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/EconLit, PAIS International, RePEc (Research in Economic Papers), Social Services Citation Index, and Wilson Business Abstracts.

Copyright © 2014 The International Bank for Reconstruction and Development/THE WORLD BANK

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without prior written permission of the publisher or a license permitting restricted copying issued in the UK by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1P 9HE, or in the USA by the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Typeset by Techset Composition Limited, Chennai, India; Printed by The Sheridan Press, USA

Research Observer

Volume 29 • Number 2 • August 2014

- Progress on Global Health Goals: Are the Poor Being Left Behind?
Adam Wagstaff, Caryn Bredenkamp, and Leander R. Buisman 137
- Is Green Growth Good for the Poor?
Stefan Dercon 163
- Entry Regulation and the Formalization of Microenterprises in Developing Countries
Miriam Bruhn and David McKenzie 186
- Cash Transfers and Child Labor
Jacobus de Hoop and Furio C. Rosati 202
- The Promise of Performance Pay? Reasons for Caution in Policy Prescriptions in the Core Civil Service
Zahid Hasnain, Nick Manning, and Jan Henryk Pierskalla 235

Progress on Global Health Goals: Are the Poor Being Left Behind?

Adam Wagstaff, Caryn Bredenkamp, and Leander R. Buisman

We examine differential progress on health Millennium Development Goals (MDGs) between the poor and the better off within countries. Our findings are based on an original analysis of 235 DHS and MICS surveys spanning 64 developing countries over the 1990–2011 period. We track five health status indicators and seven intervention indicators from all four health MDGs. In approximately three-quarters of countries, the poorest 40 percent have made faster progress than the richest 60 percent on MDG intervention indicators. On average, relative inequality in these indicators has been falling. However, in terms of MDG outcome indicators, in nearly half of the countries, relative inequality has been growing. Moreover, in approximately one-quarter of the countries, the poorest 40 percent have been slipping backwards in absolute terms on both MDG interventions and outcomes. Despite reductions in most countries, relative inequalities in MDG health indicators are still appreciable, with the poor facing higher risks of malnutrition and death in childhood and lower odds of receiving key health interventions.

With only a few months to go until the December 31, 2015, target date for the attainment of the Millennium Development Goals (MDGs), global and country-specific progress toward the health-related MDGs is being fervently discussed by ministries of health, development partners, civil society organizations, and health advocates. Recent estimates ([Go and Quijada 2012](#); [United Nations 2012](#); [World Bank 2013](#)) show that although progress toward most non-health MDG targets is largely on track, progress toward the health MDG targets is not.

Largely absent from the discussion on the health MDGs has been the question raised by [Gwatkin \(2005\)](#) nearly 10 years ago: whether progress within countries has been pro-poor (i.e., faster among the poor and hence inequality reducing) or pro-rich.¹ Only three studies have shed light on this question, and, as we show below, these studies reached different conclusions. As a result, it is currently unclear

whether progress to date on the health MDGs has been pro-poor or pro-rich for both interventions (e.g., immunization) and outcomes (e.g., under-five mortality).

In this paper, we re-analyze this unresolved question using a larger number of MDG indicators, more countries, more survey years, more recent data, and more measures of inequality than previous studies. We find that with the exception of HIV prevalence, where progress has, on average, been markedly pro-rich, progress on the MDG health outcome (health status) indicators has, on average, been neither pro-rich nor pro-poor. Average rates of progress are similar among the poorest 40 percent and among the richest 60 percent. On average, the concentration index (the measure of relative inequality that we use) neither rose nor fell. A rosier picture emerges for MDG intervention indicators: whether we compare rates of change for the poorest 40 percent and richest 60 percent or consider changes in the concentration index, we find that progress has, on average, been pro-poor.

However, behind these broad-brush findings lie variations around the mean. Not all countries have progressed in an equally pro-poor way. In almost half of countries, (relative) inequality in child malnutrition and child mortality fell, but it also increased in almost half of countries, often quite markedly. We find some geographic concentration of pro-rich progress; in almost all countries in Asia, progress on underweight has been pro-rich, and in much of Africa, inequalities in under-five mortality have been growing. Even on the MDG intervention indicators, we find that a sizable fraction of countries have progressed in a pro-rich fashion. Progress on skilled birth attendance and antenatal care has been pro-rich in 20 percent of countries, and in as much as 25–30 percent of countries, inequalities in immunization have grown. We also find that despite mostly pro-poor progress over time (at least on the intervention indicators), relative inequalities in MDG health indicators persist. The poor face a higher risk of malnutrition and death in childhood, and their odds of receiving key health interventions are appreciably lower.

Although our analysis focuses on the MDG indicators, it has implications for the measurement and monitoring of trends in inequalities in health status and health interventions more generally. Our finding that progress on health interventions has been more pro-poor than progress on health outcomes raises questions about where to focus efforts on monitoring and how to incentivize, financially or otherwise, reaching the poor.

What Did Previous Studies Find about Whether Progress Toward the Health MDGs Has Been Pro-Poor?

Table 1 indicates, for each of the main four health-related MDGs, the goal, the targets, the official indicators, and the additional core intermediate indicators that

Table 1. MDG Health Goals, Targets, and Indicators

<i>MDG goal</i>	<i>Targets</i>	<i>Official indicators</i>	<i>Additional core intermediate monitoring indicators[§]</i>
Goal 1: Eradicate extreme poverty and hunger	Halve, between 1990 and 2015, the proportion of people who suffer from hunger	Prevalence of underweight children under five years of age Proportion of population below minimum level of dietary energy consumption	Percentage of children aged 6 to 59 months who received one dose of vitamin A in the past six months, proportion of infants under six months who are exclusively breastfed
Goal 4: Reduce child mortality	Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	Under-five mortality rate Infant mortality rate Measles immunization among children under one	Proportion of infants under six months who are exclusively breastfed, proportion of surviving infants who have received a dose of measles vaccine by their first birthday, proportion of children with fast or difficult breathing in the past two weeks who received an appropriate antibiotic, proportion of children with diarrhea in the past two weeks who received oral rehydration therapy (ORT), proportion of children under five who slept under an insecticide-treated net the previous night (in malarious areas), proportion of children with fever in the past two weeks who received an appropriate anti-malarial (in malarious areas)
Goal 5: Improve maternal health	Reduce by three-quarters, between 1990 and 2015, the maternal mortality ratio	Maternal mortality ratio Proportion of births attended by skilled health personnel	Contraceptive prevalence rate, percentage of women with any antenatal care, provision of emergency obstetric care, syphilis in pregnant women and proportion that are properly treated, percentage of women receiving antenatal care who receive at least two to three intermittent preventive malaria treatments during pregnancy (in malarious areas)

Continued

Table 1. *Continued*

<i>MDG goal</i>	<i>Targets</i>	<i>Official indicators</i>	<i>Additional core intermediate monitoring indicators[§]</i>
Goal 6: Combat HIV/AIDS, malaria and other diseases	Have halted by 2015 and begun to reverse the spread of HIV/AIDS	HIV prevalence among 15- to 24-year-old pregnant women Condom use rate of the contraceptive prevalence rate Number of children orphaned by HIV/AIDS	Percentage of persons using a condom at last higher-risk sex, percentage of sexually transmitted infection clients who are appropriately diagnosed and treated according to guidelines, percentage of HIV-positive women receiving anti-retroviral treatment during pregnancy to prevent mother-to-child transmission of HIV
	Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	Prevalence of and death rate associated with malaria Proportion of population in malaria risk areas using effective malaria prevention and treatment measures	Percentage of patients with uncomplicated malaria who received treatment within 24 hours of onset of symptoms, percentage of children under five sleeping under insecticide-treated nets, percentage of pregnant women sleeping under insecticide-treated nets, percentage of pregnant women who have taken chemoprophylaxis or drug treatment for malaria
		Prevalence of and death rates associated with tuberculosis Proportion of tuberculosis cases detected and cured under directly observed treatment, short-course (DOTS)	Percentage of estimated new smear-positive tuberculosis cases that were registered under the DOTS approach

Source: [Wagstaff and Claeson \(2004\)](#).

Notes: [§] Additional core intermediate monitoring indicators are those recommended by a technical consultation between the World Bank and UN technical agencies and are as reported in Table 2.1 of [Wagstaff and Claeson \(2004\)](#). That table also lists optional indicators.

Table 2. Previous Studies of Differential Progress by Wealth toward the Health MDGs

	<i>Moser et al. (2005)</i>	<i>Suzuki et al. (2012)</i>	<i>Victoria et al. (2012)</i>
Time period	1990–2001	1995–2007	1990–2007
No. countries	22	40	35
Data source	Demographic and Health Survey (DHS)	Demographic and Health Survey (DHS), Multiple Indicator Cluster Survey (MICS)	Demographic and Health Survey (DHS)
No. surveys per country	2	2	2
Official indicators	Under-five mortality	Prevalence of underweight children under five years of age Under-five mortality rate Infant mortality rate Measles immunization among children under one Proportion of births attended by skilled health personnel	Skilled birth attendant Measles immunization
Additional core intermediate monitoring indicators	None	Use of insecticide-treated bednets by under-five children Proportion of infants under six months who are exclusively breastfed Proportion of children with diarrhea in the past two weeks who received oral rehydration therapy (ORT) Proportion of children with fever in the past two weeks who received an appropriate anti-malarial (in malarious areas) Contraceptive prevalence rate	Use of insecticide-treated bednets by under-five children
Stratifying variable	Wealth index	Wealth index	Wealth index
Summary statistic	Changes in ratio of the richest 20 percent to the poorest 20 percent	Changes in ratio of the richest 20 percent to the poorest 20 percent	Changes in concentration index
Findings	An approximately equal number of countries with widening and narrowing inequalities	Pro-poor progress on underweight Pro-rich progress on under-five mortality Pro-poor progress on most intervention indicators except skilled birth attendance	Inequalities have been falling over time

Source: Authors' compilation based on listed studies.

emerged from a technical consultation between the UN technical agencies and the World Bank in 2001.

Table 2 summarizes the three previous studies (Moser et al. 2005; Suzuki et al. 2012; Victora et al. 2012) that shed light on the issue of whether progress toward the health MDGs has been pro-poor at the country level. The studies vary in the time period under analysis and in the number of countries studied. Two use only the Demographic and Health Survey (DHS), whereas one uses both the DHS and the Multiple Indicator Cluster Survey (MICS). One study focuses on only one MDG and one official indicator, another looks at two MDGs and one official indicator of each, and a third looks at three MDGs and six official indicators.

In the absence of consumption data, the preferred measure of household living standards in developing countries (Deaton 1997), all three studies distinguish between poor and better-off households using the ‘wealth index’ developed by Filmer and Pritchett (2001). Two of the previous three studies address the inequality issue by looking at changes in the ratio of the outcome among the poorest quintile to the outcome among the richest quintile. The other uses the concentration index (Kakwani et al. 1997), analogous to the Gini coefficient except that when considering inequality in, say, underweight, individuals are ranked by the household’s score on the wealth index rather than by underweight. Both approaches speak to the issue of relative – rather than absolute – inequality and shed light on whether progress has been pro-poor in a relative sense (in which case relative inequality is said to have fallen).

The findings of the three previous studies are mixed. Examining under-five mortality (U5MR) between 1990 and 2001, Moser et al. (2005) find no clear trend, whereas Suzuki et al. (2012) find that the progress between 1995 and 2007 in the MDG 1, 4, and 5 outcomes was generally pro-rich. On interventions, both Suzuki et al. (2012) and Victora et al. (2012) mostly find pro-poor progress, except for skilled birth attendance (SBA), where Suzuki et al. (2012), but not Victora et al. (2012), find that progress has been pro-rich.

Data and Indicators

Data

Data are drawn from the DHS and the MICS surveys. These global household survey programs are the only regular source of comparable household-level data from which the health MDG monitoring indicators can be constructed. We analyzed the raw data from 164 DHS surveys and 71 MICS surveys covering a total of 91 countries between 1990 and 2011. After eliminating countries with only one period of data and one dataset with some implausible values, we were left with 64 countries

Table 3. Details of the Datasets Used in the Current Study

MDG	Indicator	No. countries	No. surveys	Surveys per country		
				Average	Minimum	Maximum
1	Stunting	53	156	2.9	2	5
1	Underweight	54	158	2.9	2	5
4	IMR	41	126	3.1	2	5
4	U5MR	41	125	3.0	2	5
4	Full immunization	60	178	3.0	2	6
4	Measles immunization	63	188	3.0	2	6
5	ANC4+	41	127	3.1	2	5
5	SBA	41	127	3.1	2	5
5	Contraceptive prevalence	38	87	2.3	2	4
6	Condom use in risky intercourse	15	32	2.1	2	3
6	HIV prevalence	9	18	2.0	2	2
6	Malaria nets (children)	23	54	2.3	2	3
	Average	40	115	2.7	2	4.5

Source: Authors' calculations from original DHS and MICS datasets.

Notes: IMR: Infant mortality rates; U5MR: Under-five mortality rates; ANC4+: Antenatal care visits (4+); SBA: Skilled birth attendance.

with at least two periods of data. Depending on the MDG indicator (see below for selection criteria and definitions), we have data for between 9 and 63 countries, with an average of 40 countries per indicator (table 3). There are some indicators with up to six surveys per country, but the average is 2.7.

Thus, we cover more countries than previous studies: 64 countries compared with the 22 analyzed by Moser et al. (2005), the 40 analyzed by Suzuki et al. (2012), and the 35 analyzed by Victora et al. (2012). Unlike previous studies whose trends have been estimated from only two surveys per country per indicator, ours are estimated from several surveys per country per indicator, for an average of 2.7.

Indicators

Table 4 shows the MDG indicators and the definitions used. Our list includes five health status indicators (stunting, underweight, infant mortality, under-five mortality, and HIV prevalence); the rest are intervention indicators. Our aim was to include as many of the official indicators as possible for each MDG.² We have included in our analysis five additional indicators that are almost always used in MDG monitoring exercises, drawn from the additional core intermediate indicators listed in table 1. Specifically, we include the prevalence of stunted children (MDG1), the percentage of children aged one who have been fully immunized (MDG4),³ the percentage of

Table 4. MDG Indicators Used in the Current Study

<i>MDG goal</i>	<i>Indicator</i>	<i>Definition</i>
Goal 1: Eradicate extreme poverty and hunger	Underweight	Percentage of children with a weight-for-age z-score < -2 standard deviations from the reference median [§]
	Stunting	Percentage of children with a height-for-age z-score < -2 standard deviations from the reference median [§]
Goal 4: Reduce child mortality	Infant mortality rate (IMR)	Number of deaths among children under 12 months of age per 1,000 live births [§]
	Under-five mortality rate (U5MR)	Number of deaths among children under 5 years of age per 1,000 live births [§]
	Measles immunization	Percentage of children aged 12 to 23 months who received measles either verified by card or by recall of respondent
	Full immunization	Percentage of children aged 12 to 23 months who received BCG, measles, and three doses of polio and DPT, either verified by card or by recall of respondent
	Malaria nets (children)	Percentage of children under 5 sleeping under an (ever) insecticide-treated bednet the previous night
Goal 5: Improve maternal health	Skilled birth attendant (SBA)	Percentage of births to mothers aged 15 to 49 that were attended by skilled health attendant (SBA)
	Antenatal care visits – 4 or more (ANC4+)	Percentage of mothers aged 15 to 49 who received at least four antenatal care (ANC4+) visits from any skilled personnel (as defined in the country's DHS or MICS)
	Contraceptive prevalence	Percentage of women aged 15 to 49 who currently use a modern method of contraception
Goal 6: Combat HIV/AIDS, malaria and other diseases	HIV prevalence	Percentage of adults aged 15 to 49 testing positive for HIV 1 or 2
	Condom use in risky intercourse	Percentage of women aged 15 to 49 who had more than one partner in the past year and used a condom during last sexual intercourse
	Malaria nets (children)	Percentage of children under 5 sleeping under an (ever) insecticide-treated bednet the previous night

Notes: [§] WHO 2006 child growth standards used to calculate z-score.[§] Mortality rate calculated using the true cohort life table approach. The DHS reports use the synthetic cohort life table approach.

children in malaria-endemic countries who sleep under an insecticide-treated bednet (MDG4 and MDG6), antenatal visit coverage (MDG5), and whether a condom was used during sexual intercourse with a high-risk partner (MDG6).

Our study thus extends the range of indicators well beyond those previously studied. Unlike Moser et al. (2005), who focused on MDG outcomes (in fact, only one outcome: under-five mortality) and Victora et al. (2012), who focused on MDG interventions, but similar to Suzuki et al. (2012), we analyze data on both MDG outcomes and MDG interventions. Although Suzuki et al. (2012) analyze more indicators than we do, we analyze more MDG indicators.⁴

Stratification by 'Wealth'

As in previous studies, we stratify households by Filmer and Pritchett's 'wealth' index.⁵ Where available, we use the wealth index variable that is included in the public release of most DHS or MICS datasets. For a handful of MICS surveys, we had to construct the wealth index ourselves. In each case, when using wealth quintiles, we sorted *households* into quintiles. Individuals then acquire the quintile of their household, so that in an analysis of, for example, child mortality, the 'poorest quintile' may account for more than 20 percent of children in the sample.

Results

We ask first how well countries *as a whole* are progressing on the MDG indicators and display our results in such a way as to show the full variation across countries and across indicators in rates of progress. We then examine progress among the poorer wealth quintiles,⁶ both in absolute terms and relative to the rest of the population. Finally, we explore more systematically how wealth-related inequalities in the MDG indicators have changed over time. Table 5 integrates the key summary statistics that emerge from our analysis. Figures 1–6 provide much more detail and highlight the heterogeneity across developing countries.

How Fast Is the Developing World Progressing toward the Health MDGs?

Figure 1 shows the average annual growth rates for all countries (ranked in ascending order of their growth rate) on each of our MDG indicators; see also columns 1 and 2 of table 5.⁷ The growth curves allow us to see at a glance the median annual growth rate for each indicator *and* the fraction of countries experiencing negative and positive growth. For the health status (intervention) indicators, progress entails a country having a negative (positive) value of average annual growth in table 5 and being below (above) the zero line on the y-axis in figure 1. Thus, for the health status (intervention) indicators, the further below (above) the horizontal axis the growth curve is in figure 1, the swifter progress has been.

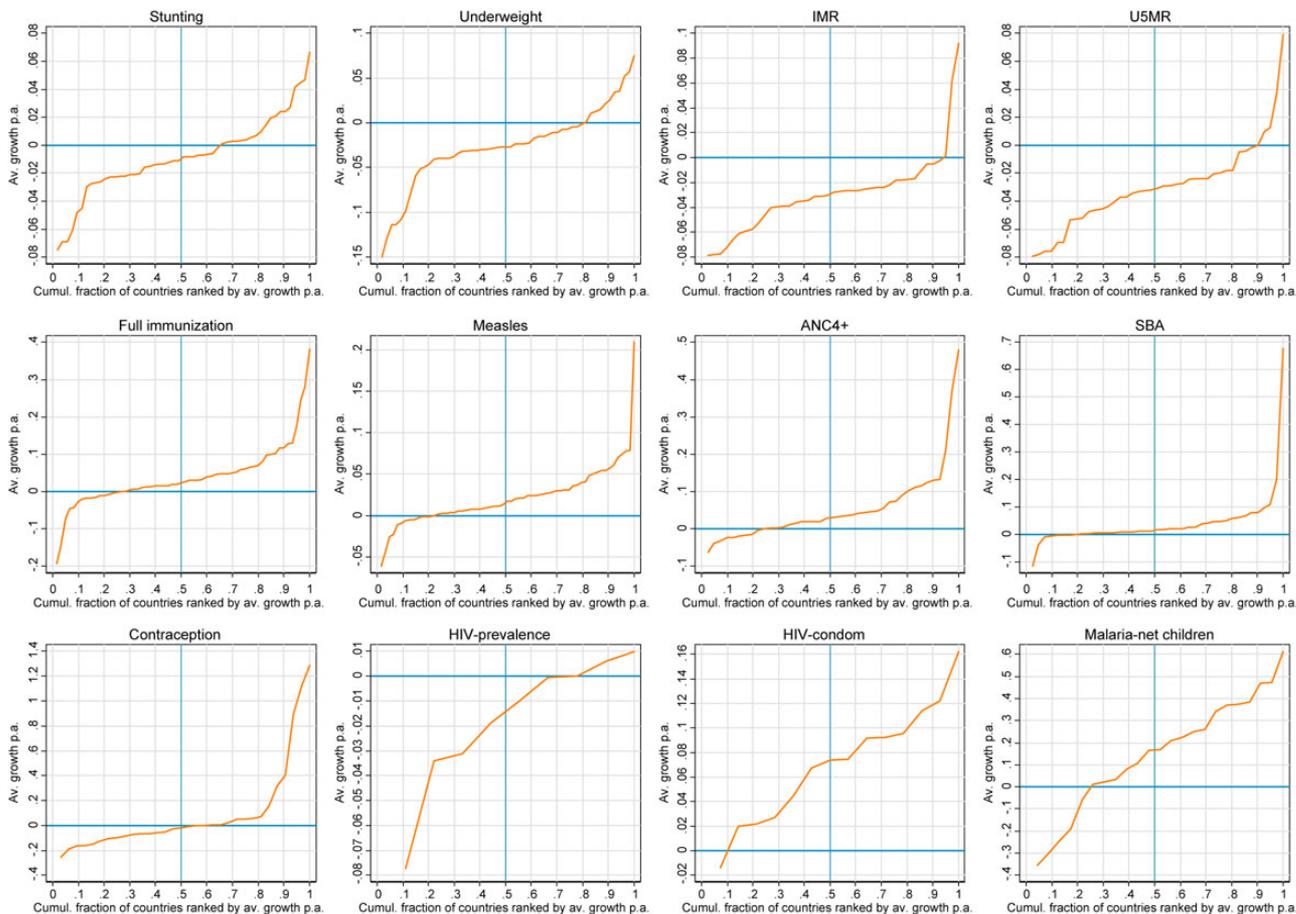
Table 5. Key Summary Statistics on Differential Progress by Wealth toward the Health MDGs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	<i>Annual growth</i>					<i>Relative inequality</i>				<i>Absolute inequality</i>			
<i>Indicator</i>	<i>Median growth rate pop</i>	<i>% countries where pop mean improved</i>	<i>Median growth rate poorest 40%</i>	<i>% countries where mean for poorest 40% improved</i>	<i>Median excess growth of poorest 40%</i>	<i>Median initial CF[§]</i>	<i>% countries with fall in CF[§]</i>	<i>Median change in CF[§] p.a.</i>	<i>as % of median initial CF[§] (absolute value)</i>	<i>Median initial absolute CF[§]</i>	<i>% countries with fall in absolute CF[§]</i>	<i>Median change in absolute CF[§] p.a.</i>	<i>as % of median initial absolute CF[§] (absolute value)</i>
Stunting	-0.0084	64	-0.0087	64	0.0023	-0.1035	58	-0.0009	0.8	-0.0406	45	0.0001	0.3
Underweight	-0.0269	80	-0.0239	74	0.0064	-0.1345	56	-0.0012	0.9	-0.0246	43	0.0004	1.8
IMR	-0.0281	93	-0.0283	93	-0.0004	-0.1039	49	0.0010	0.9	-0.0080	34	0.0003	3.4
U5MR	-0.0307	88	-0.0328	90	0.0019	-0.1209	61	-0.0009	0.7	-0.0143	34	0.0003	2.1
Full immunization	0.0242	73	0.0295	72	0.0105	0.0934	68	-0.0040	4.2	0.0419	58	-0.0008	1.9
Measles	0.0169	78	0.0217	78	0.0091	0.0655	75	-0.0025	3.8	0.0386	73	-0.0012	3.1
ANC4+	0.0310	78	0.0445	78	0.0164	0.2001	78	-0.0049	2.4	0.0571	44	0.0003	0.6
SBA	0.0183	83	0.0277	80	0.0130	0.2807	73	-0.0043	1.5	0.1080	49	0.0004	0.3
Contraception	-0.0143	41	0.0065	50	-0.0009	0.0691	53	-0.0015	2.2	0.0097	53	-0.0005	5.6
HIV prevalence	-0.0101	67	-0.0026	56	0.0198	0.1526	67	-0.0035	2.3	0.0066	89	-0.0004	5.4
HIV condom	0.0740	93	0.1095	100	0.0412	0.2446	93	-0.0113	4.6	0.0625	57	-0.0002	0.4
Malaria-net children	0.1691	78	0.3353	74	0.0480	0.2950	65	-0.0304	10.3	0.0171	52	-0.0001	0.8
Av. health status	-0.0208	78	-0.0193	75	0.0060	-0.0620	58	-0.0011	1.1	-0.0162	49	0.0002	2.6
Av. interventions	0.0456	75	0.0821	76	0.0196	0.1783	72	-0.0084	4.2	0.0478	55	-0.0003	1.8

Source: Authors' calculations from original DHS and MICS datasets.

Notes: [§] CI is the concentration index – see text for definition.

Figure 1. Progress toward the Health MDGs for the Entire Population



Source: Authors' calculations from original DHS and MICS datasets.

Notes: The annual growth is as a fraction, so, e.g., 0.02 means 2% annual growth.

On all indicators except contraception, the majority of countries have made progress. For example, reading across the zero line on the y-axis in the stunting chart, we see that in approximately 65 percent of countries the prevalence of stunting has fallen; for underweight, the figure is 80 percent. All but 22 percent of countries have seen increases in antenatal care coverage. Reading along the median line on the x-axis in the infant mortality rate (IMR) chart, we see that the median annual rate of reduction of IMR is approximately 0.03, or three percent; the median annual rate of increase in measles immunization is approximately two percent.

Several results shown in figure 1 are worth underscoring. First, despite progress for the developing world as a whole, not all countries have made progress toward the MDGs. For some MDGs, the fraction that has not made progress is very high. More than half of countries have seen no progress in contraceptive prevalence, one-quarter of countries have not recorded increases in immunization coverage, and one-fifth of countries have not made progress in reducing underweight, with the figure for stunting as high as one-third. Moreover, it is not only a question of some countries not making progress; there are countries that are actually going backwards. In some cases, such as contraception, the countries that are going backwards are doing so fairly slowly; in others, such as infant- and under-five mortality, immunization, and malaria nets, the rates of negative progress are quite high.

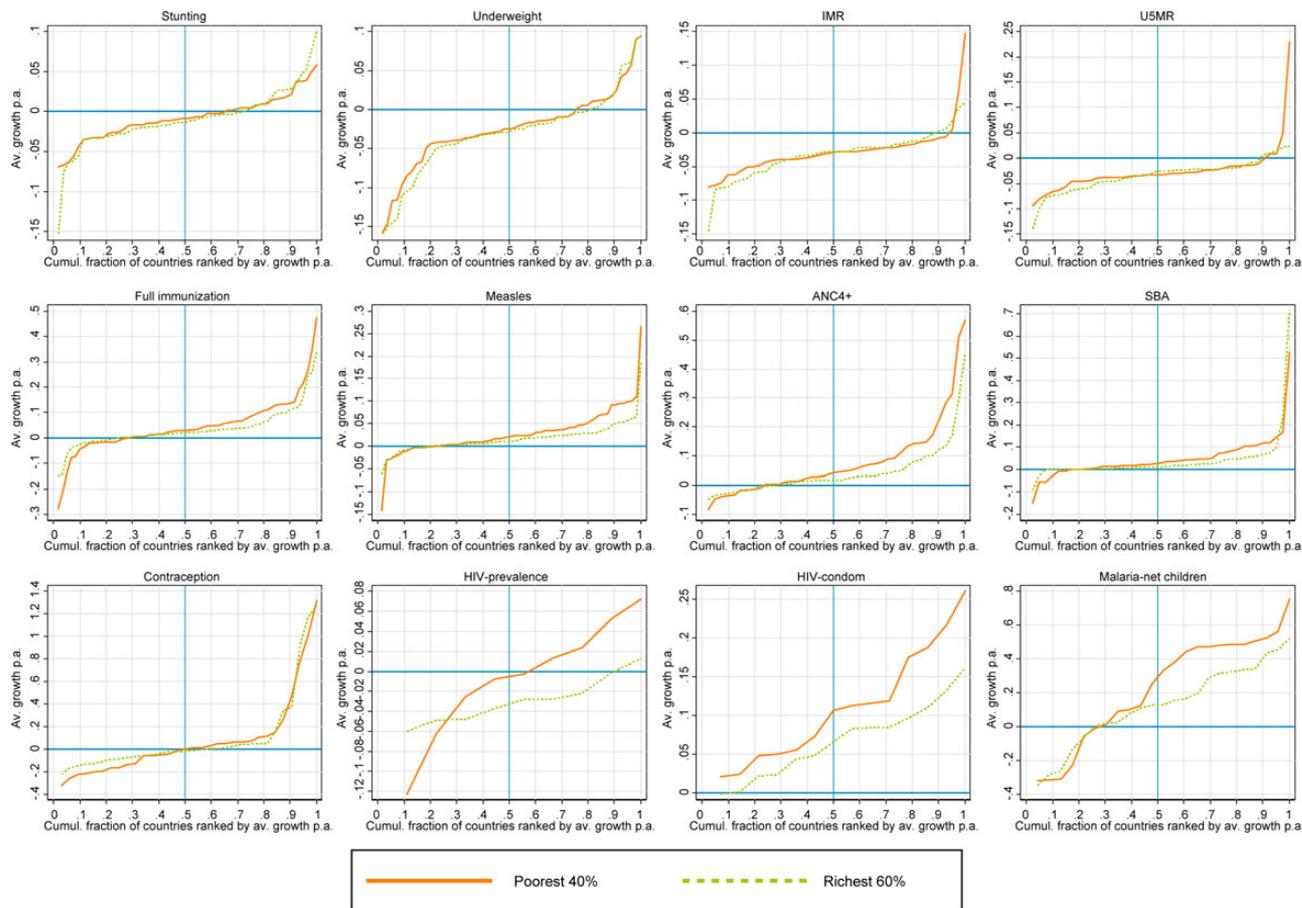
A second noteworthy finding illustrated in figure 1 is the variation *across indicators* in the amount of progress made. The conclusion we draw from this variation, however, depends on whether we look at median rates of growth or the fraction of countries making progress. For example, the median country recorded roughly similar rates of reduction for underweight, IMR and U5MR. However, 93 percent of countries saw progress on IMR, whereas only 80 percent saw such progress on underweight. Overall, progress has been greater for child mortality (Goal 4) than for malnutrition (Goal 1) and maternal health (Goal 5).

A final point to note is that the evenness with which progress is spread *across countries* varies by indicator. The growth in measles immunization is very flat; very few countries record an exceptionally high growth rate, and none achieved a growth rate of more than 20 percent. The skilled birth attendance (SBA) curve is also fairly flat, but it has a steeper and more extended right tail; approximately five percent of countries averaged growth in excess of 10 percent, some recording growth rates substantially higher than 10 percent. The malaria net growth curve shows much more variation in the spread of growth rates across countries.

How Fast Are the Poor Progressing toward the Health MDGs?

Figure 2 shows the growth curves for the poorest 40 percent and the richest 60 percent within each country; see also columns 3 and 4 of table 5. The first conclusion is that for all MDG indicators except contraception, the poorest 40 percent

Figure 2. Progress toward the Health MDGs among the Poorest 40 Percent and Richest 60 Percent



Source: Authors' calculations from original DHS and MICS datasets.

Notes: The annual growth is as a fraction, so, e.g., 0.02 means 2% annual growth.

have made progress, whether we measure this in terms of the median growth rate or the fraction of countries in which progress among the poorest 40 percent has occurred. We also see similar broad patterns in the data of the poorest 40 percent and the data for the population as a whole: negative progress among some countries, variations across indicators in the speed of progress, and small pockets of rapid progress on some indicators.

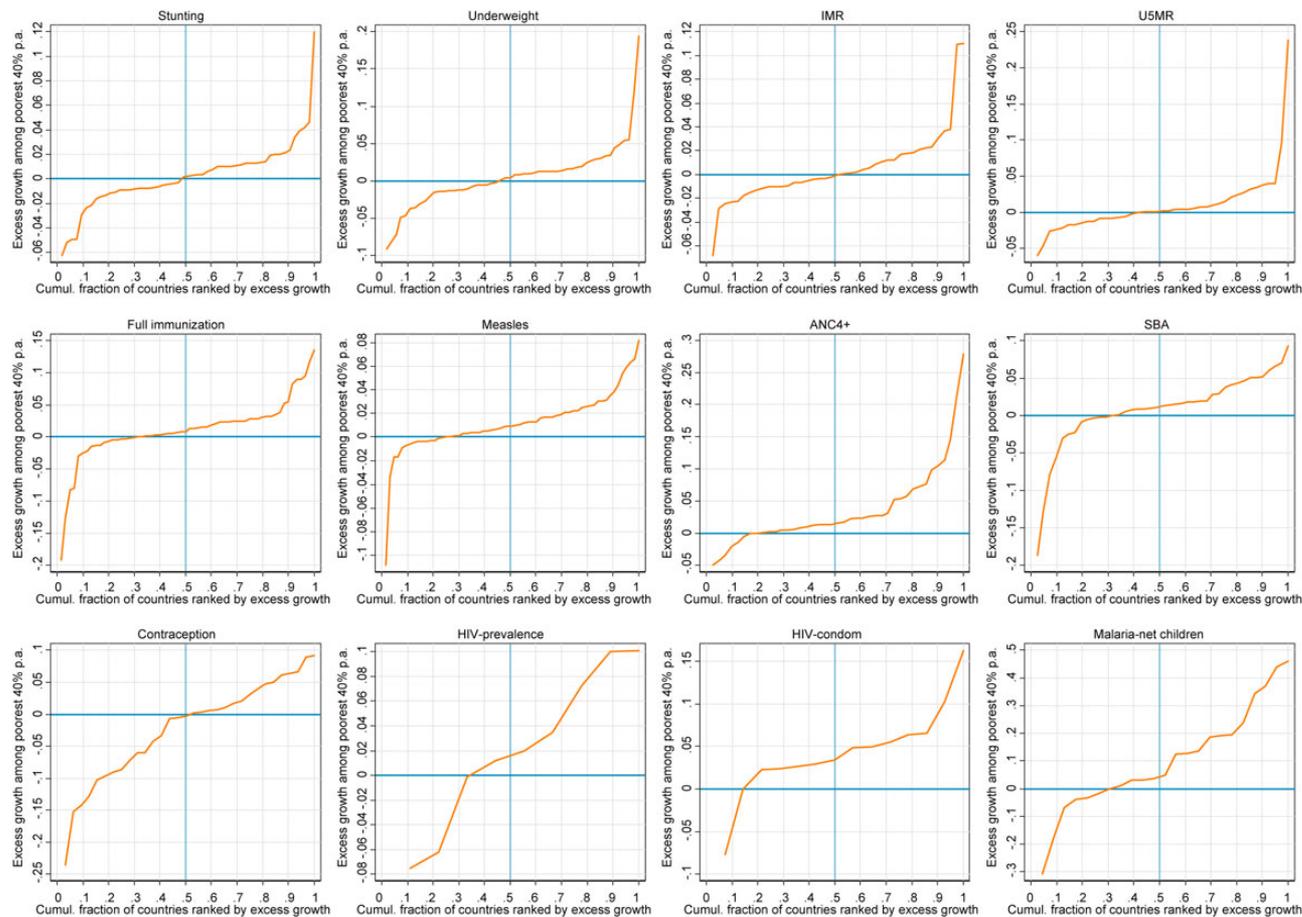
Although the poorest 40 percent have, broadly speaking, shared in the progress toward the health MDGs of the population as a whole, figure 2 reveals some differences in the speed of progress between the poorest 40 percent and richest 60 percent. The degree to which progress has been pro-poor or pro-rich is also shown in figure 3 (and column 5 of table 5), which shows how countries vary in excess growth among the poorest 40 percent (i.e., the excess of the growth rate among the poorest 40 percent over the growth rate among the richest 60 percent). For an intervention indicator, such as immunization, positive excess growth among the poorest 40 percent means faster progress among the poor (i.e., pro-poor progress). For health status indicators, such as underweight and mortality, the opposite is the case; positive excess growth among the poorest 40 percent means pro-rich progress.

On the MDG outcomes, the picture varies somewhat by indicator. For HIV prevalence, progress has been markedly pro-rich. The median rate of reduction for the richest 60 percent was almost three percentage points higher than that for the poorest 40 percent. Although 90 percent of countries saw a reduction in HIV prevalence among the richest 60 percent, only approximately 55 percent of countries saw a reduction in HIV prevalence among the poorest 40 percent. Figure 3 shows that in the case of HIV prevalence, excess (negative) growth among the poorest 40 percent was recorded in only approximately 35 percent of countries.

By contrast, no striking differences in progress between the poor and the better off emerge for malnutrition and child mortality. The growth curves for the poorest 40 percent and the richest 60 percent are close to one another and actually intersect several times, and the excess growth curves in figure 3 all intersect the median line on the x-axis around zero on the y-axis. For both malnutrition indicators, the median rate of reduction among the richest 60 percent is very slightly larger than the median rate among the poorest 40 percent, and the fraction of countries where the richest 60 percent saw reductions in malnutrition is very slightly higher than the fraction of countries where the poorest 40 percent saw reductions. Figure 3 shows that for malnutrition, the fraction of countries that saw negative excess (i.e., pro-poor) growth among the poorest 40 percent is only slightly less than 50 percent, and the median excess growth rate is almost zero. On the excess growth measure, progress on child mortality emerges as neither pro-poor nor pro-rich.

On MDG intervention indicators, a distinctly pro-poor picture emerges. In figure 2, for all indicators except contraception, the growth curve for the poorest 40 percent lies above that for the richest 60 percent. Whether we consider the difference in median

Figure 3. Incidence of Excess Growth among the Poorest 40 Percent on Health MDG Indicators



Source: Authors' calculations from original DHS and MICS datasets.

Notes: The annual growth is as a fraction, so, e.g., 0.02 means 2% annual growth.

growth rates or the difference in the fraction of countries recording progress, we see a pro-poor result. We obtain the same pro-poor picture for MDG intervention indicators in figure 3, where all curves except that for contraception intersect the zero line on the y-axis to the left of the median line on the x-axis. Thus, whether we look at the median rate of excess growth among the poorest 40 percent or at the fraction of countries where there is positive excess growth among the poorest 40 percent, we see the same pro-poor pattern emerges. Using the latter indicator of pro-pooriness, figure 3 suggests that progress has been especially pro-poor for antenatal care and condom use, whereas progress in the use of contraception has been neither pro-poor nor pro-rich.

Figure 3 points to large differences across countries in the degree to which progress toward the health MDGs has been pro-poor. For example, while on average progress has been broadly similar for the poorest 40 percent and the richest 60 percent for malnutrition and child mortality, there are some countries (to the left of each graph in figure 3) where progress has been dramatically pro-poor (i.e. growth rates have been more negative for the poorest 40 percent than for the richest 60 percent) but also some countries (to the right of each graph in figure 3) where progress has been markedly pro-rich (i.e. growth rates have been less negative for the poorest 40 percent). Likewise, while, on average, progress on intervention indicators has been pro-poor, figure 3 shows that in plenty of countries, progress has been slower for the poorest 40 percent than for the richest 60 percent.

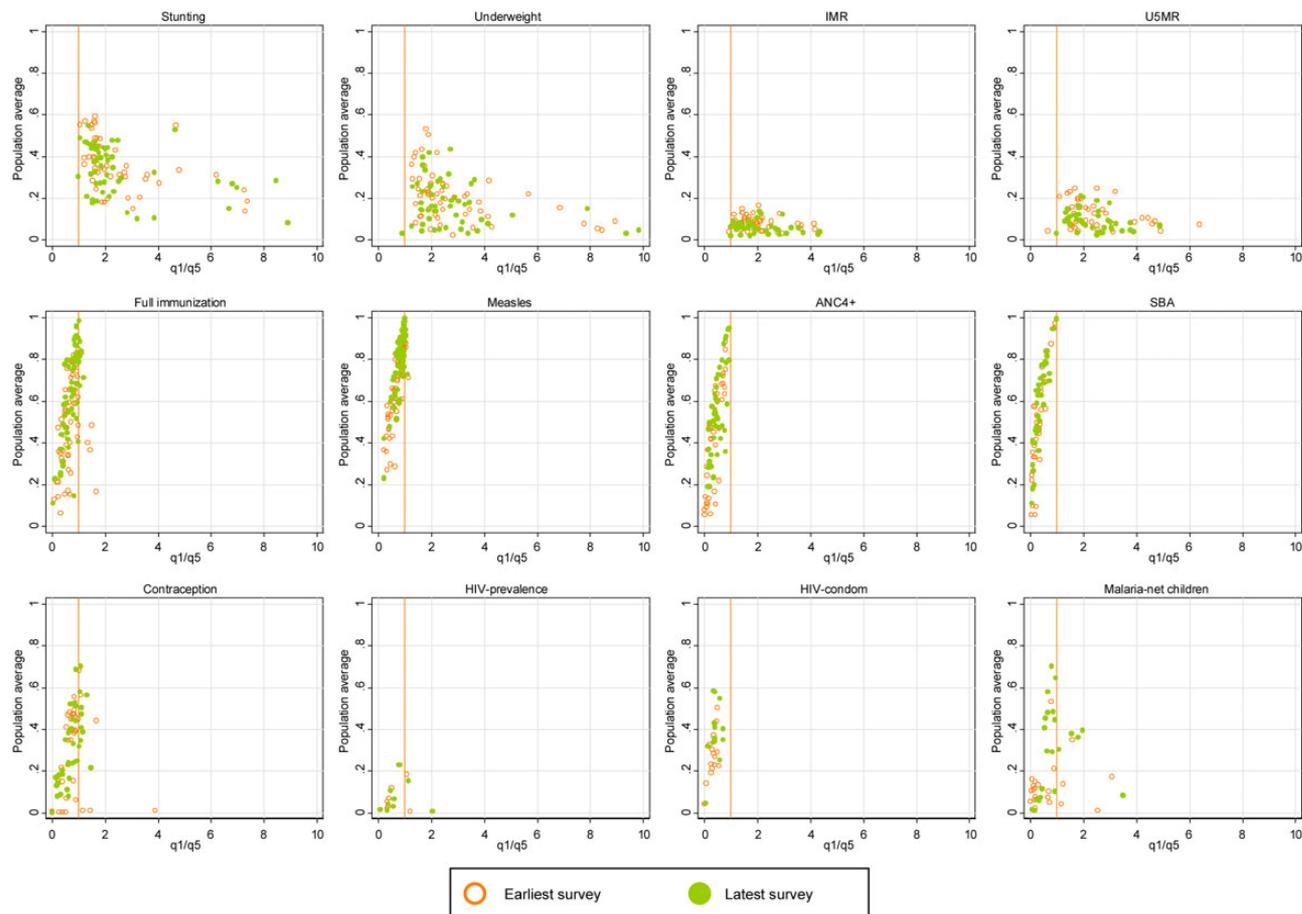
When we examine geographic patterns, we find that for most indicators, most continents have a mixed picture: some countries have achieved pro-poor progress, but not all. This is less true of Asia and the Americas; for antenatal care and skilled birth attendance, for example, almost all countries in these regions have achieved pro-poor progress. By contrast, in almost all countries in Asia, progress on underweight has been pro-rich. More detailed information on geographic patterns was presented by Wagstaff et al. (2014, Figure 4).

To what extent is overall pro-poor progress attributable to some countries improving systematically across all MDG indicators? Treating each country-indicator combination as an observation and defining the outcome as whether the country recorded excess growth among the poorest 40 percent for that indicator, we performed a two-way analysis of variance. The 12 indicator and 64 country codes were entered as independent variables. We find that two-thirds of the explained variance was attributable to variation across countries, holding indicators constant. In other words, two-thirds of the explained variation can be explained by some countries systematically doing better than others in achieving pro-poor progress *across all indicators*.

How are Wealth-Related Inequalities in the Health MDGs Changing?

Faster rates of improvement among the poorer wealth groups imply a reduction in wealth-related inequalities in the MDG health indicators. Figures 2 and 3 are thus

Figure 4. Trends in the Ratio of the Poorest Quintile's MDG Indicator Value to the Richest Quintile's Value



Source: Authors' calculations from original DHS and MICS datasets.

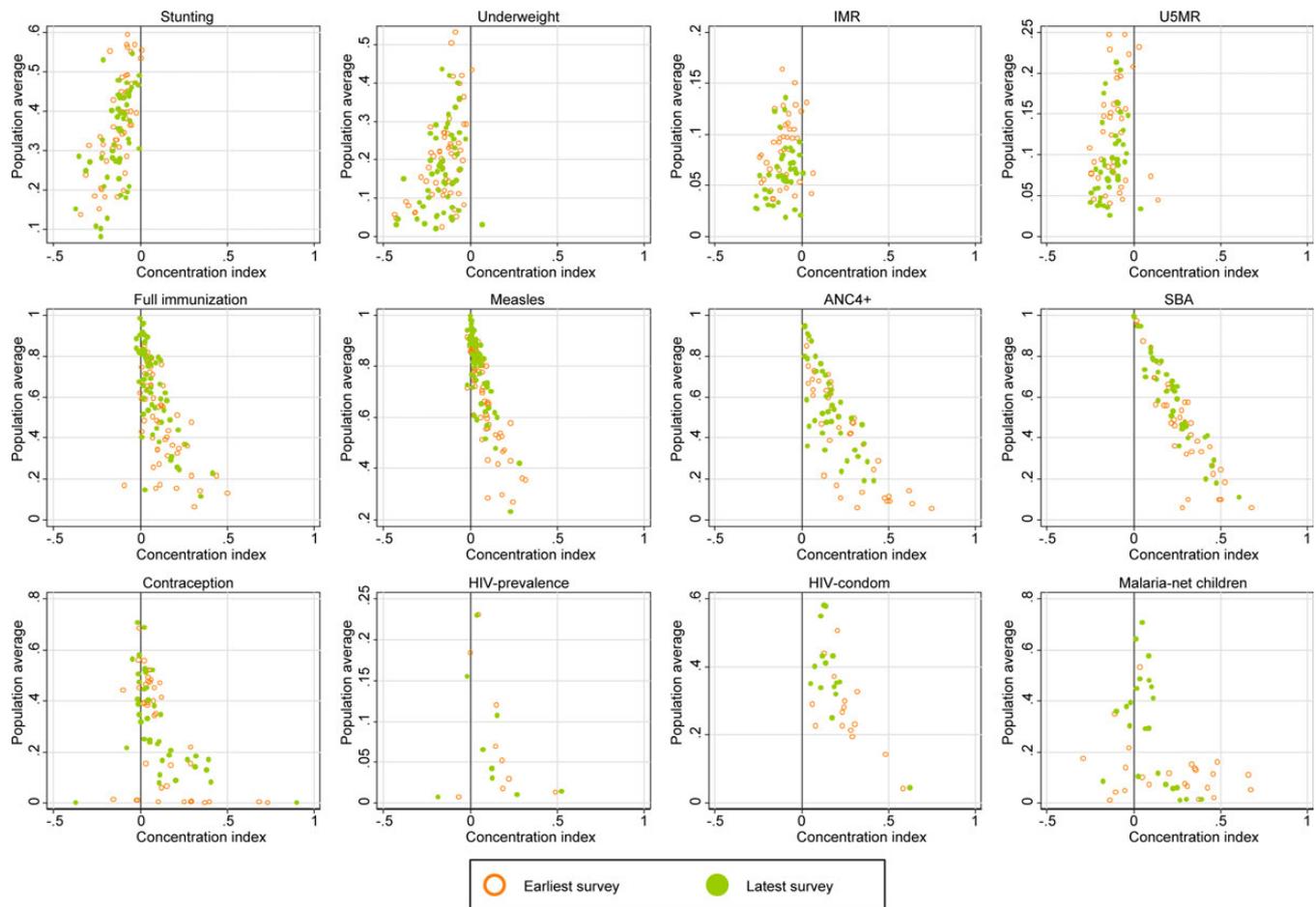
Notes: The population rate is as a fraction, so, e.g., 0.2 means a population average of 20%. This is true also for IMR and U5MR, which are typically expressed per 1,000.

suggestive of inequality reduction for the intervention indicators and of little change in inequality for the outcome indicators. However, the charts do not provide a definitive answer on inequality trends. It could be argued that splitting the population into two at the 40 percent mark is a somewhat arbitrary way of tracking inequality. Moreover, the charts do not indicate by *how much* inequality has been reduced or the amount of inequality remaining at the time of the latest survey. Finally, pro-poor progress – in the sense of faster growth among the poor – implies a reduction in *relative* inequality but not necessarily in *absolute* inequality. If the poor start behind the better off, as they do on MDG intervention indicators, they will see a larger *absolute* increase only if their increase is considerably larger in percentage terms.

A simple – and very incomplete – relative inequality measure is the ratio of the poorest quintile's value to that of the richest quintile. Figure 4 shows this measure for the earliest survey in each country and the latest survey. A more complete measure of wealth-related health status and health intervention inequality is the concentration index (Kakwani et al. 1997). A positive value indicates that the better off, on average, have higher values of the outcome variable (good in the case of an intervention, bad in the case of a negative health status indicator such as malnutrition or mortality), whereas a negative value indicates that they have, on average, lower values (bad in the case of an intervention, good in the case of an indicator such as malnutrition or mortality). With a binary outcome, such as the variables used in this paper, the bounds of the concentration index are $\mu-1$ and $1-\mu$, where μ is the mean of the outcome variable (Wagstaff 2005). Thus, as the mean increases, the range of possible values shrinks, tending to zero as the mean goes to one. Figure 5 shows the population mean and the concentration index for the earliest and latest surveys for all countries for all MDG indicators; see also column 6 of table 5.⁸

Four points are worth noting in figures 4 and 5. First, reinforcing the findings of earlier studies, we find pro-rich inequalities on all but one indicator: the poor, on average, have higher rates than the better off on all MDG health status indicators except HIV and lower rates on the MDG intervention indicators. Second, for the intervention indicators, we see the ratios in figure 4 moving closer to one and the concentration indices in figure 5 approaching zero as the population average moves closer to one. All of the charts in figure 5 have a somewhat half-arrowhead appearance, reflecting the fact that the bounds shrink as the population mean rises toward one and that inequality is pro-rich almost everywhere (so that we see only the left or right of the arrowhead). Third, in the case of the intervention indicators, the markers in both charts seem to be rising over time (indicating that population averages are improving) and closing in on the vertical line on the x-axis (indicating that inequalities are falling). For the malnutrition and child mortality indicators, we see the markers falling over time but no clear pattern in terms of inequality. We will return to this in a moment. Fourth, inequalities favoring the better off are still very much in evidence even in the latest round of surveys.

Figure 5. Trends in the Concentration Index of MDG Indicators



Source: Authors' calculations from original DHS and MICS datasets.

Notes: The population rate is as a fraction, so, e.g., 0.2 means a population average of 20%. This is true also for IMR and U5MR, which are typically expressed per 1,000.

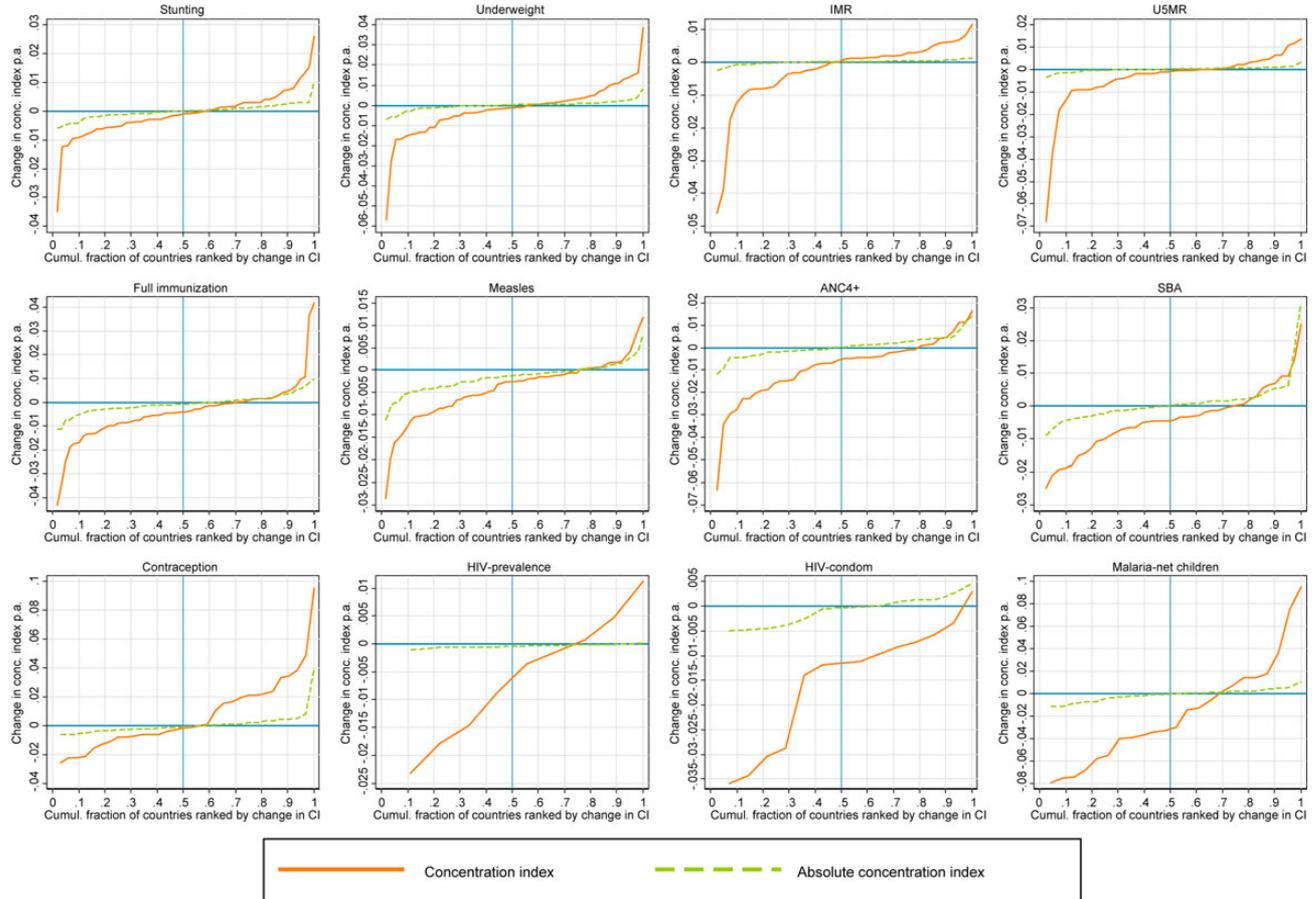
The analogue of the concentration index for the absolute inequality measurement is the absolute concentration index, equal to the mean multiplied by the standard concentration index.⁹ A positive value indicates higher values, on average, among the better off, and a negative value indicates higher values, on average, among the less well off. In the case of a binary outcome, the bounds of the absolute concentration index start at zero when the population mean is zero, grow as the mean rises above zero reaching minimum and maximum values of -0.25 and +0.25 as the mean reaches one half. As the population mean rises further, the bounds shrink, reaching zero again as the mean reaches one.¹⁰

When we look at absolute inequality (see column 10 of table 5 and figure 7 in Wagstaff et al. (2014)), there are five findings that are worth highlighting. First, absolute inequality also favors the better off. Second, as the population mean rises, absolute inequality does not always decrease; instead, there is a sign – especially visible for the ANC and SBA indicators – of inequalities growing as the population mean increases before diminishing. This is consistent with the bounds of the absolute concentration index having a rotated ellipse shape with minima and maxima of -0.25 and +0.25. Third, the smaller bounds of the absolute concentration index explain why the values for the absolute concentration index are smaller than those for the regular concentration index. Fourth, as with relative inequality, it appears that the upward drift in the intervention markers over time has been associated with less absolute inequality, whereas for the malnutrition and mortality indicators, the pattern is less clear. Fifth, as with relative inequality, we still see evidence of inequalities even in the latest data.

Figure 6 shows the global distribution, for each indicator, of annual average changes in the standard and absolute concentration indices; see also columns 7, 8, 9, 11, 12 and 13 of table 5. Two points are worth highlighting.

First, reductions in relative inequality are more evident for the intervention indicators than for the health status indicators. For IMR, roughly half of countries saw relative inequality increase, and half saw it fall. For the other four health status indicators, only a slight majority of countries saw a reduction in relative inequality, and the (absolute) percentage change per year in relative inequality was small (typically less than one percent). The health status indicator that bucks this trend is HIV prevalence, where 67 percent of countries saw relative inequality fall and the regular concentration index fell by 2.3 percent per year. Keep in mind, however, that in the initial round of surveys, HIV was more prevalent among the better off, not the poor. Thus, although a reduction in inequality occurred, it was not a pro-poor change. On the intervention indicators, reductions in inequality are more pronounced. On average, nearly three-quarters of countries record reductions in relative inequality on the intervention indicators, and the annual reduction in the concentration index averages approximately four percent.

Figure 6. Variations across Countries in Trends in Relative and Absolute Wealth-Related Inequality in MDG Indicators



Source: Authors' calculations from original DHS and MICS datasets.

Notes: The vertical axis measures the annual average change in the index, not the percentage change.

Second, fewer countries have experienced reductions in absolute inequality than relative inequality and the median reductions in absolute inequality are smaller. On the malnutrition and child mortality indicators, we actually see *increases* in absolute inequality in the majority of countries, with changes in the absolute concentration index for these indicators averaging three percent. On the intervention indicators, we see reductions in absolute inequality, but the percentage reductions in absolute inequality are, on average, half as large as the percentage reductions in relative inequality. The different behaviors of the relative and absolute inequality measures reflect the fact that an improvement for someone in the richest half of the population will cause absolute inequality to rise, but it need not cause relative inequality to rise. Depending on the initial concentration index value, improvements for people well above the mid-point of the wealth distribution could still be compatible with a reduction in relative inequality.¹¹

Discussion

The developing world as a whole is making progress toward all MDG indicators. Unsurprisingly, perhaps, given the different starting points, progress has been slower on some indicators (for example, immunization, which had higher initial rates of coverage) than on others (for example, antenatal care coverage, which had lower initial rates). More surprising is the fact that for every indicator, there are several countries that have gone backwards. For some indicators, the fraction of such countries is quite high: one-quarter for immunization and one-fifth and one-third for underweight and stunting.

The main focus of this paper, however, has been on differential progress between the poor and the better off. There are three key messages of this paper. First, in a majority of countries, the poorest 40 percent are making faster progress than the richest 60 percent, and, on average, relative inequality in the MDG indicators has been falling. Second, in an appreciable fraction of countries, the poorest 40 percent have not been progressing in absolute terms. Moreover, in a sizable fraction of countries, the poorest 40 percent have progressed less quickly than the richest 60 percent, and relative inequality has been growing. This is especially true of the child health status indicators (for child malnutrition and mortality, the fraction in this latter category is 40 to 50 percent), but it is also true of some intervention indicators (for immunization, the figure is almost 40 percent). Third, although, on average, relative inequality has been falling, appreciable inequalities remain. In the latest round of surveys used in this study (of which all but two are from the 2000s) and based on a comparison of odds ratios, the richest 60 percent of children in the developing world are 1.2 times as likely to sleep under a bednet, 1.5 times as likely

to be immunized, and 3.7 times as likely to have been delivered by a skilled birth attendant as the poorest 40 percent of children.

On the subset of issues that we address and that previous authors have addressed, we see some similarities in results as well as some differences. Our results are broadly consistent with those of [Victora et al. \(2012\)](#) and [Suzuki et al. \(2012\)](#) who found pro-poor progress on most of the MDG intervention indicators that were also examined here; the exception is SBA, which [Suzuki et al. \(2012\)](#) concluded had been pro-rich. On inequalities in health status indicators, our results on under-five mortality are consistent with those of [Moser et al. \(2005\)](#), who found that progress had been neither pro-poor nor pro-rich on balance. However, our results on under-weight are less encouraging than those of [Suzuki et al. \(2012\)](#), who found pro-poor progress in more than 60 percent of countries. Differences in findings may be due to variations in the methodological approaches adopted, such as differences in the number of countries included, time period examined, the definition of poor and non-poor and the measure of inequality.

Our first finding – that in most countries the poor have not been left behind by the health MDGs – will be reassuring to donors, international development and technical agencies, national governments, NGOs, program implementers, and health care professionals. However, our other two findings – that relative inequalities have grown in a sizable minority of countries, especially on health status indicators, and that despite reductions in most countries, inequalities are still appreciable – will be a cause for concern. Governments, donors and NGOs ought to be able to exert a fairly direct influence over the coverage of immunization, skilled birth attendants, and bednets, among others. Our findings suggest that [Gwatkin \(2005\)](#) was right to warn that programs that improve population averages may not necessarily disproportionately benefit the poor. In addition, they imply that it is indeed important to monitor progress not only on population averages but also on inequalities – a point made by the UN’s 2013 high-level panel report on the post-2015 development agenda ([United Nations 2013](#)). Our results should prompt further discussion of appropriate policies to reduce inequalities in intervention coverage and health status. For example, are universalist policies the answer, or is there a greater need for targeting efforts toward the poor? Do ‘new’ initiatives such as conditional cash transfers and pay-for-performance stand a better chance of reducing inequalities than ‘traditional’ approaches such as pure input-financing?

Our finding that progress on interventions has been considerably more pro-poor than progress on health status is a puzzle that merits further research. One hypothesis is that the quality of health care is worse for lower socioeconomic groups; though the poorest 40 percent may have experienced a larger percentage increase in, for example, antenatal visits, they have not observed the same improvement in the survival prospects of their babies. If true, this finding would point to the need for a monitoring framework that captures not only the quantity of care (as is

currently the case) but also its quality. Another hypothesis is that although inequalities in the official monitoring indicators have been narrowing, inequalities in at least some of the interventions that are not monitored may have been widening. Obvious examples are health interventions delivered privately (i.e., in the home) by caregivers, which are often complementary to monitored interventions, such as breastfeeding, hand-washing, and antibiotics for a child suffering from dysentery or pneumonia. These are indicators that a UN technical consultation suggested should be tracked but that did not obtain official indicator status. As a result, the interventions in question may have been under-promoted. If this is the case, it may suggest the need for a broadening of the monitoring framework to capture interventions delivered by caregivers as well as a policy shift in the health sector toward preventive and curative interventions delivered in the home and in the community.

Notes

Adam Wagstaff: Development Research Group, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA; Caryn Bredenkamp: Health, Nutrition and Population Global Practice, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA; Leander R. Buisman: Institute of Health Policy and Management, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Corresponding author: Adam Wagstaff. Development Research Group, The World Bank, 1818 H Street, NW, Washington, DC 20433, USA. Tel: +1 202 473 0566, awagstaff@worldbank.org. The research reported in this paper was supported by the World Bank's Rapid Social Response Multi-Donor Trust Fund (RSR-MDTF). We are grateful to Nicole Klingen, Gayle Martin, Owen Smith, Aparnaa Somanathan, Owen O'Donnell, Ellen Van de Poel, Jos Verbeek, Xiaoqing Yu, the WBRO editor, and three anonymous referees for advice and useful comments on earlier versions of the paper. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments of the countries they represent.

1. Throughout this paper, "pro-poor" indicates that changes in MDG health indicators are such that there is a reduction in inequality (in the indicator) between the poor and better-off. This is in the spirit of the definition of pro-poor growth proposed by [Kakwani and Pernia \(2000\)](#) (i.e. growth that reduces income inequality) and the definition of pro-poorness used in the context of the redistributive effects of taxes and transfers (i.e. taxes and transfers that reduce income inequality) ([Lambert 2001](#)).

2. Some official indicators, such as the number of tuberculosis (TB) cases detected and treated under DOTS, cannot be constructed from the DHS or MICS. Deaths from malaria and TB are also not captured by regular household surveys such as the DHS and MICS, partly because of sample size issues but also because of the difficulty of ascertaining after the fact the cause of death. Maternal mortality is often computed using variables from the DHS, but maternal deaths are so rare that quintile-level analysis would not be meaningful.

3. The official indicator captures only measles immunization.

4. We include two additional MDG6 indicators: condom use and HIV prevalence.

5. The wealth index is created by using principal components analysis to construct a weighted average of a variety of indicators capturing the ownership of assets (e.g., car and television) and the characteristics of the family's home (e.g., type of floor and roof). Subsequent to the work of Gwatkin et al. ([2000](#); [2007](#)), who published tabulations by 'wealth quintile' of a large battery of health,

nutrition and population indicators for all DHS surveys, the organization responsible for the DHS (then Macro International) decided to include the Filmer-Pritchett wealth index in each public-release DHS dataset. UNICEF, which is responsible for the MICS, subsequently decided to do the same with the MICS.

6. The poorest 40 percent is the target group in the World Bank Group's (WBG) new goal of 'promoting shared prosperity'; see, e.g., <http://www.worldbank.org/content/dam/Worldbank/document/WB-goals2013.pdf>, where the WBG states, "We will monitor progress in shared prosperity using the income growth of the bottom 40 percent of a nation's population. This implies a direct focus on the income of the less well-off, as opposed to the common practice of focusing only on growth of GDP per capita and implicitly relying on the "trickle down" impact of growth on the bottom of the distribution. The measure captures the two elements central to the notion of shared prosperity: a growing economy and a fundamental concern for equity."

7. Growth rates are computed using information from all available survey years. We assume a compound growth process and estimate the average annual percentage change by regressing – for each country and each indicator, separately for the population and the poorest 40 percent – the natural logarithm of the indicator on the date of the survey. The average annual rate of growth is equal to the exponential of the (regression) coefficient on the year variable minus one (see, e.g., <http://data.worldbank.org/about/data-overview/methodologies>).

8. Concentration indices were computed from the individual-level data, not the quintile data, using the convenient covariance method (see, e.g., Jenkins 1988), except in the case of the mortality variables, whose concentration indices were computed from quintile data (see, e.g., Kakwani et al. 1997).

9. The absolute concentration index is the same as the generalized concentration index. The former is derived from the absolute concentration curve and the latter from the generalized concentration curve. On absolute inequality, see Jenkins and Jäntti (2005 p.14), Moyes (1987 p.205) and Lambert (2001 p.56). In the health context, see Wagstaff et al. (1991) and Wagstaff (2009).

10. This follows from the fact that the absolute concentration index is equal to the mean times the standard concentration index.

11. Wagstaff (2009) proves both results. An increase in individual i 's outcome (holding everyone else's constant) will raise absolute inequality if person i is in the upper half of the wealth distribution. The same increase will raise relative inequality only if the person's rank in the wealth distribution exceeds $1/2(CI+1)$, where CI is the value of the (initial) concentration index. Therefore, if CI is 0.25, person i 's rank in the wealth distribution would have to be above 0.625 for relative inequality to increase.

References

- Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, MD: Johns Hopkins University Press.
- Filmer, D., and L. Pritchett. 2001. "Estimating wealth effects without expenditure data or tears: An application to educational enrollments in states of India." *Demography* 38 (1) : 115–32.
- Go, D. S., and J. A. Quijada. 2012. "The Odds of Achieving the MDGs." *The World Bank Research Observer* 27 (2) : 143–84.
- Gwatkin, D., S. Rutstein, K. Johnson, R. Pande, and A. Wagstaff. 2000. *Socioeconomic Differences in Health, Nutrition and Population*. Washington, DC: The World Bank.
- Gwatkin, D., S. Rutstein, K. Johnson, E. Suliman, A. Wagstaff, and A. Amouzou. 2007. *Socio-Economic Differences in Health, Nutrition and Population Within Developing Countries: An Overview*. Washington, DC: The World Bank.

- Gwatkin, D. R. 2005. "How much would poor people gain from faster progress towards the Millennium Development Goals for health?" *The Lancet* 365 (9461) : 813–17.
- Jenkins, S. 1988. "Calculating income distribution indices from microdata." *National Tax Journal* 61 : 139–42.
- Jenkins, S. P., and M. Jäntti. 2005. "Methods for summarizing and comparing wealth distributions." Working Paper, Colchester, University of Essex, Institute for Social and Economic Research.
- Kakwani, N., and E. M. Pernia. 2000. "What Is Pro-poor Growth?" *Asian Development Review* 18 (1) : 1–16.
- Kakwani, N., A. Wagstaff, and E. van Doorslaer. 1997. "Socioeconomic Inequalities in Health: Measurement, Computation, and Statistical Inference." *Journal of Econometrics* 77 (1) : 87–103.
- Lambert, P. 2001. *The distribution and redistribution of income: A mathematical analysis*. Manchester: Manchester University Press.
- Moser, K. A., D. A. Leon, and D. R. Gwatkin. 2005. "How does progress towards the child mortality millennium development goal affect inequalities between the poorest and least poor? Analysis of Demographic and Health Survey data." *BMJ* 331 (7526) : 1180–2.
- Moyes, P. 1987. "A new concept of Lorenz domination." *Economics Letters* 23 : 203–7.
- Suzuki, E., M. Sharan, and E. Bos. 2012. *Poverty and Health Monitoring Report*. Washington, DC: World Bank.
- United Nations. 2012. *Millennium Development Goals Report, 2012*. New York: United Nations.
- . 2013. *A new global partnership: eradicate poverty and transform economies through sustainable development. The report of the high-level panel of eminent persons on the post-2015 development agenda*. New York: United Nations.
- Victora, C. G., A. J. D. Barros, H. Axelson, Z. A. Bhutta, M. Chopra, G. V. A. França, and K. Kerberet al. 2012. "How changes in coverage affect equity in maternal and child health interventions in 35 Countdown to 2015 countries: an analysis of national surveys." *The Lancet* 380 (9848) : 1149–56.
- Wagstaff, A. 2005. "The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality." *Health Economics* 14 (4) : 429–32.
- . 2009. "Correcting the concentration index: A comment." *Journal of Health Economics* 28 (2) : 516–20.
- Wagstaff, A., C. Bredenkamp, and L. R. Buisman. 2014. "Progress Toward the Health MDGs: Are the Poor Being Left Behind?" Policy Research Working Paper 6894, World Bank, Washington, DC.
- Wagstaff, A., and M. Claeson. 2004. *The Millennium Development Goals For Health: Rising to the Challenges*. Washington, DC: World Bank.
- Wagstaff, A., P. Paci, and E. van Doorslaer. 1991. "On the measurement of inequalities in health." *Social Science & Medicine* 33 (5) : 545–57.
- World Bank. 2013. *Global Monitoring Report 2013: Rural-Urban Dynamics and the Millennium Development Goals*. Washington, DC: World Bank.

Is Green Growth Good for the Poor?

Stefan Dercon

The developing world is experiencing substantial environmental change, and climate change is likely to accelerate these processes in the coming decades. Due to their initial poverty and their relatively high dependence on environmental capital for their livelihoods, the poor are likely to suffer most due to their low resources for mitigation and investment in adaptation. Economic growth is essential for any large-scale poverty reduction. Green growth, a growth process that is sensitive to environmental and climate change concerns, can be particularly helpful in this respect. We focus on the possible trade-offs between the greening of growth and poverty reduction, and we highlight the sectoral and spatial processes behind effective poverty reduction. High labor intensity, declining shares of agriculture in GDP and employment, migration, and urbanization are essential features of poverty-reducing growth. We contrast some common and stylized green-sensitive growth ideas related to agriculture, trade, technology, infrastructure, and urban development with the requirements of poverty-sensitive growth. We find that these ideas may cause a slowdown in the effectiveness of growth to reduce poverty. The main lesson is that trade-offs are bound to exist; they increase the social costs of green growth and should be explicitly addressed. If they are not addressed, green growth may not be good for the poor, and the poor should not be asked to pay the price for sustaining growth while greening the planet.

green growth, poverty. JEL codes: O44, Q01, I31

Environmental degradation is occurring in many parts of the developing world. Nationally and locally, it is the result of deforestation, soil degradation, the depletion of water resources, and environmental pressures linked to urbanization and economic change. Globally, the overall process of climate change is expected to gradually but sharply increase in the coming decades with mean temperature rises, sea-level rises, and spatial changes in rainfall and other climatic conditions as well as the increased frequency or severity of extreme weather events.

The World Bank Research Observer

© The Author 2014. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
doi:10.1093/wbro/lku007 Advance Access publication July 16, 2014 29:163–185

Environmental change is rarely equity neutral. The poor are generally considered the main losers from both climate change and the burdens of local environmental damage and natural resource degradation. They are typically more dependent on environmental capital and climate for their economic activities because most of the poor still live in rural areas that are dependent on agriculture. Those in urban areas face the consequences of environmental hazards linked to overcrowding, pollution, and inadequate water and sanitation provision. The poor are also more vulnerable to extreme events affecting economic productivity, health, and security of livelihood with limited insurance or social protection. Furthermore, informal insurance mechanisms are not suited to address covariate risks such as climate risks or other risks affecting entire communities. The poor may also find it more difficult to adapt their livelihoods to changing environmental conditions because they lack the resources to invest in appropriate and profitable economic activities.¹

Development and poverty reduction investments are powerful instruments to mitigate these environmental impacts on the well-being of the poor and to offer them the resources to develop their resilience to further environmental pressures. Economic growth and development in the poorest economies is essential to build this resilience to adapt to and cope with the new reality (World Bank 2010a). It has been the key element in large-scale poverty reduction, most notably in Asia, although there are considerable geographical, sectoral, and structural differences in the speed with which poverty reduction is delivered in the context of growth (Ravallion 2000). Growth has been found to be important specifically for increased climate change adaptive capacity. For example, examining time series data across countries over the last 50 years, Dell et al. (2008, 2009) found that higher temperatures significantly reduced economic growth rates in poor countries but not in rich countries. Raddatz (2009) showed large declines in GDP per capita from climate-related disasters in low-income countries; in percentage terms, these declines were four times the size of the declines in rich economies. Noy (2009) showed that higher GDP per capita and better institutional and human development indicators reduced losses from climate-related disasters. Raddatz (2009) also found that the larger impacts in low-income countries are much more than could be explained by the relatively high share of agriculture in these countries, so this issue is not simply solved by diversification away from agriculture. Although they are difficult to identify statistically, Fomby et al. (2012) also found losses in both agricultural and non-agricultural growth, although the type of shocks appears to affect generalizations on this issue.

Unfortunately, engineering these growth and poverty reduction processes in a world of environmental change is problematic. It is generally acknowledged that the pressures of climate and other environmental changes will have serious implications for the growth prospects of some of the poorest countries in the developing world over the coming decades (IPCC 2007; World Bank 2010a). Changes in

climatic conditions will change the mean return to agriculture and increase its variance. Furthermore, the increased impacts of extreme weather events will affect the accumulation of productive assets, not least in low-lying areas such as coastal cities, where the potential returns to economic activity are especially high. These effects are also likely to accelerate the depletion of many forms of environmental capital and increase threats to human health from disease and water scarcity. The impacts of climate change on agriculture and other sectors of the economy are bound to affect the scope for their transformation from economies that are largely dependent on agriculture to a more diversified, higher-return economy.

Climate change is only one source of environmental pressure that has serious impacts in many developing countries. There are also difficult choices to be made with respect to curbing *current* environmentally damaging activities. The depletion of soil, forests, coastal fisheries, and more accessible fresh water supplies reflects an ongoing loss of potential productivity in the economy. Air and water pollution represent significant burdens on human health in a number of locations, lowering productivity and diverting scarce resources into the treatment of pollution-related illnesses. These externalities are rarely adequately internalized into the decisions made by the users of these resources, so addressing them can have potentially significant impacts on growth and the distribution of gains from growth. However, curbing environmental damage *now* to reduce natural resource degradation and improve the population's health and quality of life means diverting resources from other conventional growth-oriented opportunities.

To address these issues, various strategies have been discussed to provide a blueprint for 'green' growth, in which the need to protect the environment is internalized while leaving sufficient opportunities for economic growth (OECD 2011). This option appears particularly attractive when viewed from a poverty angle. Importantly, it retains and may even provide further impetus to a growth focus, which is essential for poverty reduction in low-income economies, while contributing to their resilience in the face of environmental problems. Furthermore, because environmental damage is currently not equity neutral, it could help ameliorate the consequences of environmental costs for the poor under current growth trajectories.

Much of the discussion on 'green growth' remains relatively vague in terms of specifics, including for poor countries or the poor in general. However, more recent reports (e.g., World Bank 2012) begin to amend this discussion in a careful, nuanced way. Just as not all growth leads to the same degree of poverty reduction or to the same environmental impacts, it is likely that not all efforts to maximize growth given environmental constraints will maximize poverty reduction. The question rarely asked is how various green growth strategies and resilience-enhancing investments interact with poverty. To what extent is green growth good for the poor?

Under what conditions can certain green growth strategies lead to unwelcome adverse impacts on the poor or even to ‘green poverty’?

In this paper, we first provide a stylized discussion of the nature of poverty, particularly its dynamics and interaction with growth. We focus not only on the various assets and capital sources of the poor but also on the sectoral and spatial dimensions of the dynamics of poverty reduction. We then discuss how global and local environmental change affect these dynamics and whether current patterns of change may persist. We introduce a number of stylized examples of ‘green’ growth initiatives and assess their impact on poverty based on their interaction with the patterns identified earlier in the paper. Finally, we present our conclusions.

The perspective that we take is to examine the *consequences* of possible greener versions of growth processes on poverty reduction. Although growth is important for poverty reduction, our focus is not on the *possibility* of sufficiently high growth from green growth strategies in developing countries. There is more evidence on this subject, and it is well reviewed (World Bank 2012). We focus on *current* poverty, not on the important issues of intergenerational equity. This is an important caveat. However, this is a deliberate choice; we have in mind the extreme poverty that is experienced in many low and lower middle income countries. Current growth patterns that do not internalize the depletion of natural capital may, of course, reduce opportunities in the future and risk some future poverty. However, those who are currently poor could be considered to have high discount rates shortening their horizon. Their current dire circumstances leave little room for the use of long planning horizons. Furthermore, most work on resilience, including resilience to future climate change, stresses the need for development, including higher incomes and wealth, as key factors in the ability of the poor to withstand this future (World Bank 2010b). Therefore, a strong focus on the current poor has considerable legitimacy in these debates. Towards the end of this paper, we revisit these two issues.

Framing Poverty and the Environment

In this section, we offer a stylized discussion of poverty and how the current environmental pressures may affect the poor. First, we provide the standard microeconomic analysis, as implicitly embedded in much of the basic writing on the welfare costs of environmental damage, and the correction of market failures. This analysis also underlies some of the writings that use a simple livelihoods framework to discuss environmental pressures (Ellis 2000). Then, we expand the framework to consider the sectoral, spatial, and dynamic (intertemporal) dimensions that are relevant in a growth context, and we show that these dimensions affect ways of thinking about poverty and environmental linkages. This discussion forms the basis for a

more in-depth discussion of how a 'green growth' strategy may affect poverty in troublesome ways.

Profile of the Poor and the Environment

Much of the analysis of poverty and its links with environmental change is rooted in a micro-level analysis of household and community livelihoods. In a stylized way, households are viewed as having access to various assets, such as financial, physical, and human capital, and importantly, environmental (or natural) capital, such as land, air quality, or water resources. Markets define the opportunities for earning a living by combining these assets and the benefits that households can obtain from them. Ownership and control of the use of these assets is not always well defined. Whereas financial or human capital is typically private with well-defined property rights, rights to environmental capital are not always clear, with a mixture of private, communal, and often contested or undefined rights. Obvious examples are the use of water resources, fisheries, and forests, whose use and management are rife with collective action problems. Furthermore, the use of environmental capital often involves externalities on others, from the local-level effects of local air or water pollution, soil degradation, and increasing global impacts, including effects on production opportunities and welfare.

These problems lead to some of the standard economic and welfare impacts of environmental pressures, with emphasis placed on market failures (stemming from externalities and coordination failures). Overcoming these market failures will improve efficiency in the economy (in the Pareto sense). This does not mean that the poor will be better off; starting from a particular distribution of income and environmental impacts with market failures, there will be winners and losers. The possibility of efficiency gains only means that the winners, in principle, would be able to compensate the losers sufficiently given the size of the gains. However, such a redistribution of gains certainly does not occur automatically. We return to this issue later in the discussion of green growth.

Labor Demand, Sectoral, and Spatial Dimensions

Beyond a standard treatment of environmental externalities and policy responses, it is helpful to add specifically sectoral, spatial, and dynamic elements that are relevant for understanding poverty. These elements are crucial in setting the stage for how environmental pressures and green growth strategies affect the poor. First, there are sectoral elements to poverty and the way in which it interacts with growth. Most of the poor are living in rural areas and are engaged in agriculture, either as smallholders or farm workers. Their labor productivity and, therefore, their incomes are low. At the same time, a substantial and likely increasing number

of the poor are living in urban areas, as in large parts of Latin America (Ravallion et al. 2007), and are working in low-paying jobs or are self-employed in the informal sector. A high dependence of the poor on agriculture or self-employment implies risky livelihoods, with limited wealth or sources of informal or formal insurance to protect themselves against these risks. Rural livelihoods are directly linked to environmental and agro-climatic risks, with particularly important implications given the pressures of climate change.

Nevertheless, this is only a static picture. Poverty reduction will require structural and sectoral change, with a decline in the dependence on agriculture for the poor, a vast reduction in the number of peasants, a reduction in informal sector employment, and increasing wage employment in other sectors. Furthermore, in addition to a change in livelihoods, this structural change will have significant spatial dimensions, implying a substantial migration from rural areas into urban areas (Dercon 2009). Large-scale migration is a standard feature of the process of large-scale poverty reduction. For example, over the last two decades, poverty reduction was accompanied by substantial migration in China, with more than 170 million people moving into cities from rural areas since 1990 (Chan 2012). Internal migration is also closely linked to welfare increases (Beegle et al. 2011).

Successful poverty reduction has other spatial dimensions, with a massive increase in the coastal population (which is where economic activity will increasingly be located due to comparative advantage) and a vast reduction of the size of the population living in areas relatively far from urban areas and the coast (because incomes in agriculture can only keep up with other incomes where demand is located or where transport is cheap).

Of course, this is not just a deterministic process of the poor uprooting, changing livelihoods, and migrating as a successful strategy to move out of poverty. With their limited capital sources and often only their labor as an asset, the poor themselves will not be the engine of growth and transformation. At best, they can be responsive to changing opportunities. Even then, these processes are fraught with problems and risks, even in a context of rapid growth. The speed of these sectoral and spatial dimensions of poverty change during growth is largely dependent on the evolution of the demand for labor during the growth process: are the growth sectors in the economy labor intensive? The higher the growth of labor demand, especially for the lower-skilled labor with which the poor are largely endowed, and the more the poor are able to respond to it via sectoral and spatial mobility, the faster poverty is reduced in the face of growth (Loayza and Raddatz 2010).

Poverty Dynamics

There is no need for this transformation to happen always and everywhere. Others have argued for the existence of combinations of economy-wide processes that

could result in low growth and high poverty ‘traps’, such as those linked to the natural resource curse, conflicts, governance, geography, or even aid (Sachs 2005; Collier 2007). All of these processes could stifle growth and the economic transformation needed to lift large populations out of poverty. Here, our focus is on the constraints that prevent particular poor populations from *benefitting* from growth. A framework allowing for various market failures can help illustrate the risks and constraints for poor populations that limit their ability to move out of poverty and potentially even trap them in persistent poverty.

The first constraint is the failure of credit markets to offer capital to the poor. Credit market failures tend to lead to collateral requirements for access to credit, resulting in those without collateral being frozen out of the market. If livelihood transformation requires at least some threshold level of capital, then a trap may occur whereby some can take advantage of opportunities, whereas others remain trapped in livelihoods with low returns. This phenomenon provides much of the justification for microcredit interventions, even if their transformative success is not clearly proven (Armendariz and Morduch 2010). The second constraint relates to risk and the lack of insurance for the poor. In this case, the poor may be induced to choose safety over higher returns, resulting in limited investment in high-return activities and technologies but choosing to hold safe assets and activity portfolios, implying a choice of poverty to reduce their exposure to even more dramatic risks (Dercon 2002; Carter and Barrett 2006). A variation of both credit and insurance market failures relates to the impact of large shocks, which, if uninsured, could push the poor back to lower levels of assets from which recovery is slow or even impossible due to limited credit, again leading to poverty traps.²

A third source of market failure relates to spatial externalities. In a positive form, these are the basis of increasing returns to scale linked to location and agglomeration and are central to much of the thinking about growth and geography (Fujita et al. 2001). In a negative form, they imply that the areas that stayed behind, such as those linked to particular poor geographical features or very limited human, physical, or social capital, may find it increasingly difficult to keep up with progress elsewhere. In turn, this may make poverty persistent (Ravallion and Jalan 1997). Migration may be a way of overcoming these spatial traps; however, interactions of various types of market failures conspire to make this difficult. Migration tends to be costly, and often only the relatively better off will manage to leave when opportunities are better elsewhere. Successful migration also tends to require networks in destination areas (Massey 2002), and those from poorer geographical areas may find it even more difficult to set up these ‘chains’ of migration. The result is that some may end up trapped in locations with limited income opportunities.

Taken together, poverty and its reduction are not only about the assets that the poor currently own or have access to, the activities in which they are engaged, and where they are currently living. Growth and poverty reduction takes place as part

of dynamic processes with important spatial and sectoral dimensions. As a result, the poverty impacts of considering environmental capital, such as by appropriately pricing environmental costs, should not be assessed only via its consequences on the value of the assets of the poor, their current activities, and their location. It will also be important to assess the sectoral, spatial, and other dynamic consequences of any changes to the growth process. In what follows, the key question is therefore: how will environmental change and greener policies not just affect the poor directly but also the particular spatial or sectoral growth incentives and the extent to which growth is intensive in the production factors owned by the poor? Before turning to green growth policies, we discuss how environmental change, not least climate change, affects poverty including via these dynamic processes.

Environmental Change and Poverty

How does environmental change affect the poor? There are various dimensions of change, such as the processes of deforestation, soil degradation, depletion of water resources, and environmental pressures linked to urbanization and economic change. There are also risks linked to climate change, which magnify other environmental pressures and appear likely to have an increasingly negative impact in coming decades, with mean temperature rises, sea-level rises, and spatial change in climatic conditions as well as an increased frequency of extreme weather events. Although the poor will no doubt be strongly affected because they have limited resources to protect themselves, it will be helpful to provide a structure for the patterns of consequences by revisiting some of the features of the poor and poverty reduction processes.

A first feature of much environmental change is the gradual erosion of the environmental capital base in many rural settings that affects the livelihoods of the poor, such as loss of forests, soil erosion, depletion of fish stocks, and water scarcity. Lower environmental capital makes income generation more difficult and affects wealth accumulation. Given credit market imperfections, this situation will affect entry into more profitable activities, including the potential exit from agriculture or diversification into other higher-return activities. Climate change is likely to accelerate these pressures, with some winners and many losers in terms of the potential for agricultural production and other climate-dependent activities. A higher frequency of extreme weather events and disasters, including droughts and floods, will put further pressure on rural livelihoods and will contribute to the poverty persistence cycles described above as investments focus on minimal livelihood security rather than higher returns and assets are lost without scope for recovery. In urban settings, livelihoods are also affected by environmental change, not least in many industrial or informal sector activities that are dependent on water and local fuel

sources, such as wood. The lives of the urban poor are further blighted by increased scarcity of clean water and air, pressures on sanitation, and the risk of disease. In both rural and urban areas, climate change and extreme events also erode infrastructure and other types of public capital.

Although higher incomes and growth could provide a route to economic diversification, investment in more productive or less environmentally damaging capital, and greater resilience in the face of environmental change, those who miss the boat may end up trapped in lower return activities that perpetuate their poverty. Importantly, the *economic and sectoral transformation* required for rising living standards is likely to be negatively affected. For example, the negative impacts of climate change on agriculture are likely to affect growth and demand for labor, thereby slowing poverty reduction. Furthermore, unlike wealthier farmers, the poor may not have the means and capital to make the necessary investments in agriculture to adequately adapt to new circumstances (including adjusting output patterns to take advantage of likely higher food prices), trapping them in low productivity agriculture. Of course, environmental degradation (specifically, climate change) may lead to winners among some of the poor in areas where agricultural opportunities increase or where adaptation investments by richer parts of society provide jobs and higher labor demand.

The *spatial* consequences of economic activity and the location of the poor are also considerably affected, not least when climate change takes hold. Some have argued that climate change would create large numbers of displaced international 'climate migrants', with figures of 200–300 million suggested (Myers 2002). However, the empirical basis for this scale of displacement is all but nonexistent (Gemenne 2011). On the contrary, lower wealth accumulation in rural settings is likely to hinder large-scale migration from marginal areas, contributing to 'spatial poverty trap'-like processes in which populations may remain trapped in marginal and vulnerable areas (Black et al. 2013; UK Government Office for Science 2011).

In a limited number of island and other locations, migration may be the only option, but its scale will be more modest. Nevertheless, there will be migration pressures, although recent reviews suggest that drivers of migration other than environmental drivers (such as those linked to economic opportunities and socio-political pressures) may still dominate (Connell 2011). Economic growth and transformation, especially as experienced in Asia in recent decades, has meant a rapid urbanization in low-lying coastal areas, and these areas will be especially vulnerable to extreme events and sea-level rises. In other words, migration *into* vulnerable areas has been the pattern and is likely to continue (Blake et al. 2011). These relatively recent and poor settlers may end up becoming more marginalized, thus undoing some of the progress that they experienced previously and making them unable to move into better areas of cities. This situation may expose them to poverty and to poor water quality and sanitation and lead to deprivation and higher disease

burdens ([Black et al. 2013](#)). Environmental change may even cause the emergence of further spatial poverty traps that are linked to higher threshold costs to move to better areas because these settlers lack the wealth to invest in necessary adaptation.

All of these processes will be exacerbated as the pressures to reduce greenhouse emissions lower global growth, with impacts on the export demand of these transforming economies and therefore on jobs and income growth. These poor and emerging economies will also face pressures to reduce their own emissions, thus also reducing growth opportunities. Indirect effects could also arise as pressures for the global reduction of greenhouse gas emissions impose costs on the world economy, thereby limiting global GDP growth and affecting both the demand for poorer countries' exports and their income growth. Furthermore, although poor countries generally have much lower greenhouse gas emissions than do higher income countries, there is likely to be increasing pressure to force low-income countries to curb their emissions as well. This would further increase the costs of transformation to a higher-return economy. Finally, climate change is only one source of environmental pressure that has serious impacts in many developing countries and elsewhere. Curbing environmentally damaging activities to safeguard natural resources for the future as well as the population's health and the quality of life will divert resources from growth-oriented opportunities and impose further costs on their economies and current economic growth opportunities.

Green Growth and Poverty

Can alternative growth paths be designed to avoid these negative consequences linked to climate change and other long-term environmental pressures? 'Green growth' alludes to patterns of growth that are consistent with internalizing some or all environmental costs but leaving sufficient opportunities for economic growth. Internalizing the social costs linked to negative environmental consequences makes the allocation of resources more efficient in a static efficiency framework. [Hallegate et al. \(2011\)](#) offer a careful discussion of when and why green growth could improve overall growth. Using an output (frontier-expanding) growth equivalent of the static efficiency arguments, properly assigning values to environmental capital would raise potential output by unlocking production factors. The less substitutable environmental capital is by other sources of capital, the higher the output gains. Furthermore, there could be efficiency gains on other production factors and more scope for technological progress. The key to these results is that we use an output measure that accounts for the environmental capital and uses appropriate (shadow) prices. In that case, appropriately valued overall output should be able to compensate any losses in efficiency gains on other factors.³

As Hallegate et al. (2011) helpfully note, growth in conventionally measured output or GDP may not necessarily increase. Environmental regulation could *reduce* conventionally measured output growth if other growth-benefiting efficiency gains or technology changes are discouraged or are not possible and the net return from those investments exceeds the net return from the environmental measure.⁴ If particular green policies reduce overall growth, then given the close link between GDP growth and poverty, poverty reduction may be slowed. Much of the overall assessment of the beneficial impact of ‘green’ growth will depend on the overall welfare objectives, the way it is measured, and whether returns to environmental capital and ecosystem services are valued directly and appropriately.

Even if welfare gains are valued properly by considering returns from environmental capital, the conclusion that there will necessarily be welfare gains from this internalization is only correct ‘on average’ for the welfare of country as a whole because any (appropriately valued) output gains create the possibility of compensation. However, among a heterogeneous population with different people owning different endowments and supplying or using factors of production, when we begin from a particular allocation that did not account for these environmental valuations, there will be distributional effects that do not necessarily imply Pareto improvements for everyone *unless* there are also (lump sum) transfers to compensate the losers. Given that this compensation rarely happens, it will be important to identify those policies that will favor or hurt the poor, even if they increase the overall output or welfare measured in particular aggregate or global ways.

Identifying Distributional Linkages

Our key concern in the remainder of the paper is the following: under what conditions do different green growth policies favor the poor? We are *not* considering further whether any overall growth trade-off exists between the greening of growth and growth itself, although this issue will remain crucial for poverty reduction. Other research, such as the study of Hallegate et al. (2011), has focused on this issue, and its implications (and opportunities) for policy are well summarized (World Bank 2012). As discussed above, under certain assumptions, green growth that appropriately values environmental capital will increase the overall (appropriately valued) output, potentially allowing any losers from this new growth path to be compensated. However, this does not mean that this compensation will occur; as a result, it does not follow that the poor gain from this move to green growth. In other words, we *only* focus on the distributional costs and benefits of green growth, with an emphasis on current poverty. Because green growth is framed in a growth context, the appropriate counterfactual to consider is how a green growth path would *change* the relationship between growth and poverty reduction. As argued

previously, this must include a consideration of the spatial, sectoral, and dynamic consequences of greening growth.

To do so, we must make ‘green growth’ more specific by articulating a number of possible families of policies that could have spatial, sectoral, and dynamic consequences. We consider green growth to consist of three types of (linked) policies. First, it relates to policies that aim to change the *prices or shadow prices* of environmental capital to internalize the externalities and other market failures that are inherent in the use and management of environmental capital. Key examples would be fuel prices, products with high intensity of fossil fuels (such as inorganic fertilizer), or water charges. Because correct pricing, including via taxation and subsidies, may not always be feasible, it could also involve other non-price interventions to affect production processes, typically via *regulation*, such as the nature of technology allowed in production processes. Examples include environmental controls on vehicles or manufacturing technologies linked to the use of water or air.

A second set of policies considered are interventions that focus directly on *investments* in low carbon or otherwise less environmentally damaging production processes. The main instruments considered are public investments and financing deals to encourage private investment or other forms of joint ventures between the public and private sector. Examples include the location and nature of transport or water infrastructure.

A third set of green growth policies might be considered, in principle, a subset of the other two, but this set of policies often is considered separately, particularly in the context of *climate change adaptation efforts* and *climate-resilient investments*. Examples include efforts to make growth more resilient to factors such as sea-level changes or increased risks in production linked to extreme weather events. This class includes infrastructure investment to reduce the impact of sea-level rises, urban planning in flood plains, or the development of lower-risk crops for (increasingly) drought prone areas. We do not consider policies that are generally good for growth but make growth more resilient in the face of climate or other environmental change and extreme events. Examples are human capital investments, flexible market access including by the poor, appropriate macroeconomic policies, and the development of better savings and insurance mechanisms (DFID 2010; World Bank 2010b).

Although we do not yet have high-quality empirical evaluations of well-defined examples of ‘green growth’ in action, we can think conceptually through its consequences and offer suggestive evidence. In principle, there are several channels through which green growth could favor the poor. For example, it could ensure that the negative costs on the poor’s livelihoods are properly internalized by producers, such as halting unmanaged commercial deforestation or the pollution of water and air, which has serious health and sanitation consequences. This is likely to have positive consequences for the poor’s living conditions, although there may be a

reduction in incomes or GDP growth. It could make growth more labor-intensive so that labor demand would rise more quickly, with real income benefits. It could facilitate the structural transformation from agriculture into other activities by, for example, increasing the prices and returns to the agricultural activities in which the poor are more involved. It could increase the connectivity between poorer and richer areas and facilitate the migration into less vulnerable areas, such as if green growth investments are used to make the roads connecting marginal and richer areas more resilient to extreme events.

These examples reflect four central dimensions to assess various ‘green growth’ measures and their consequences on the poor, as developed previously: first, (static) efficiency gains, whereby internalizing externalities may offer potential welfare gains to the poor; second, the way in which green growth could contribute to poverty reduction via growth in employment (i.e., the labor intensity of green growth); third, the livelihood and sectoral transformational dimension of green growth and the extent to which it allows the poor to move into higher return activities; and finally, the spatial connectivity and mobility dimension of green growth in terms of migration opportunities and the linkages between poorer and richer areas.

It is possible to test a number of core examples from each of the three categories of green growth policies, environmental pricing and regulation, low carbon investment (as an example of clean investment more generally), and adaptation investment (as an example of risk-mitigation investment) with regard to their potential impact on the four dimensions above. The result is that certain elements of the ‘green growth’ policy set may have far less positive (or even negative) impacts on current poverty reduction than others. The list is not exhaustive, but the overall patterns will become clear. These examples do not make assumptions about their effectiveness to increase growth; this is left to other work to assess. We only assume that they may offer reasonable growth prospects that are ‘green’. At most, a simple narrative is offered for how this may happen.

Environmental Pricing and Regulation

Let us first consider charging prices for natural resources such as fuel or water that more closely reflect the full social opportunity costs of their production and use. Considering first the poor as consumers, such charges would be paid in absolute terms more by the rich than the poor simply because they use more fuel and water. However, the welfare effects on the poor are more appropriately assessed by the share of spending on these commodities, which is likely to be relatively high for the poor. In any case, the poor will be affected as consumers irrespective of whether they are relatively less or more affected by the charges than are the rich. These impacts can be considerable (Coady et al. 2006). In theory, the efficiency gains should allow for enough resources for lump-sum transfers to compensate the poor, but in

practice, ex-post redistributive measures have been difficult to institute. This reflects an important general principle in *any* pricing-based green growth strategy: without complementary actions, the poor will be harmed as consumers. *Compensatory social protection must be part and parcel of any attempt to internalize shadow prices of natural capital for the poor not to suffer reduced real income.*⁵

Forms of regulation could have similar impacts as pricing, even in cases in which the poor may not be forced to pay more directly. In terms of distributional consequences, it is easy to imagine gains from regulatory changes for poor slum dwellers who are exposed, for example, to air and water pollution. For their part, richer groups already have more resources to adapt to environmental pressures (such as via private clean-up of local environmental damage or better garbage collection). Provided that the improvement in environmental consequences for consumers does not substantially or even totally exclude the poor and the benefits that they obtain are not outweighed by increased payments for public or private services (such as garbage collection), this may even be seen as pro-poor regulation. In general, the design and enforcement of pricing policies or other regulation as part of green growth policies will determine the extent to which the poor benefit.

Bias in favor of richer groups in society is also a distinct possibility. Regulation in the form of planning restrictions is a good example. As part of the introduction of green growth in some middle and low-income countries, stricter planning restrictions may exist for industries to be forced not to locate near residential areas to protect air quality. However, if the poor live in unplanned city settlements, then they may suffer more because these settlements may not be covered by the regulatory protection, and industries may relocate toward these areas (Blake et al. 2011).

Thus far, we have only considered the poor as consumers of ecosystem services from environmental capital. Poor households are also dependent for their incomes on environmental capital. Whether poor households are losing or gaining will therefore also depend on how regulation or pricing affects their production and job opportunities. Many of the poor use environmental capital directly, as occurs in agriculture or fisheries. Internalizing environmental capital costs is predicted to have a positive impact on sustaining this source of capital. However, a longer-run improvement can be accompanied by lower returns to activities in the nearer term. If the poor have relatively high discount rates (as seems reasonable; see Deaton 1990), then they will put a relatively higher weight on nearer-term costs compared to the rich. It is possible that the policy could redistribute wealth away from the current poor.

The poor may also be employed in industries with strong impacts on the environment. If the cost of using environmental capital rises, then there would be incentives to move to production processes that are less intensive in environmental capital and more intensive in alternative production factors, such as physical or human capital (e.g., linked to more costly and sophisticated technologies that

require more capital and technical support). Although the size of these effects would depend on the substitutability of environmental capital with these other capital sources, the poor tend to have less access to human and physical capital, so it may be more difficult for the poor to enter these now more profitable activities and compete with wealthier incumbents or more skilled workers.

A similar concern is related to labor demand of the type the poor can supply and the consequences for labor demand from rising relative costs of using environmental capital, such as the impact of a relative increase in fuel or water prices or environmental regulation. The key for the poor would be the low-skilled-labor intensity of this 'new' growth trajectory. Arguments are often made for 'green jobs', but a priori, the expectation that industries need to find more energy efficient ways of production may lead to higher intensity in human and physical capital with sophisticated technologies, which are not *necessarily* labor intensive. In any case, it will be crucial to explicitly assess the labor intensity of these alternative technologies for cleaner production. Their efficiency and ability to sustain longer-term growth is not sufficient to make a judgment about whether these green growth changes are good for jobs in the nearer term, or even in the longer term, if the poor continue to face barriers to acquiring additional human capital.

There is as yet no evidence for developing countries, so the views expressed on this are mainly conjecture. Although it is often taken for granted, the evidence on green jobs in rich economies in response to reducing incentives for the use of environmental capital is not clear-cut at best, and positive effects – if any – are dependent on local circumstances (Huberty et al. 2011). Some of the highest quality evidence on reducing incentives for the use of environmental capital gives one pause. Using very detailed plant-level growth data, Greenstone (2002) shows that 590,000 jobs were displaced between 1972 and 1987 in the United States due to a particular key environmental regulation: the US Clear Air Act. Recent work on more cost-effective incentive-based policy instruments finds milder negative effects (Harrington et al. 2012). In any event, the way in which this would play out in the developing world is difficult to judge, but maintaining labor-intensive growth that internalizes environmental capital costs is not automatic. The above discussion indicates that the implications of the latter for the former depend on numerous adjustments of factor proportions and technology within sectors as well as adjustments of output shares across sectors.

If the labor intensity of growth cannot be maintained, this situation leads to further consequences in terms of structural transformation. Growth that is less labor intensive will slow the labor absorption from agriculture. Moreover, inputs into agricultural growth, such as fertilizer, water, or transport, will become more expensive. Although lower intensification of agriculture may preserve the environmental capital required for agriculture in the long run, it will slow the agricultural and labor productivity growth that feeds structural transformation in the nearer

term. If policies reduce incentives for mobility (due to fuel costs or transport regulation), the spatial development effects could also be considerable because longer-distance trade may become less profitable and because populations in more remote areas are at risk of losing their connectivity with urban centers.

Low Carbon and Other Environmentally Friendly Public Investments

The second set of green growth policies we consider here is the direct implementation or encouragement by the public sector of environmentally friendly investments. In the transport sector, for example, some capital may be allocated away from long-distance travel (such as highways connecting towns over large distances), making the movement of goods and people more expensive and affecting the integration of economies. Investments may instead focus on local-level development (such as supporting growth focused on local linkages, including agriculture near cities and other local product focuses). More marginal areas would find it more difficult to catch up because the integration of economies would be affected. In both examples, the dynamic transformation of economies across sectors and space would be affected. Whether the local growth effects would compensate for these negative effects is not clear a priori. However, it is highly unlikely that such plans would survive any serious economic scrutiny.

Among the more serious options, some have argued that investment in low carbon energy production could raise employment considerably because it would require considerable labor. This view would lead to green growth as an engine of pro-poor job creation. For example, [Engel and Kammen \(2009\)](#) provide data on the labor intensity in years per GWh for various means of producing energy, including gas, coal, nuclear, solar, geothermal, and biomass. The data suggest that much more labor is required per GWh for low carbon alternatives than for gas or coal. The implication of this higher labor intensity is that moving resources to invest more in low carbon alternatives would create more jobs for the same energy supply; therefore, moving to low carbon would be pro-poor. However, this reasoning is incorrect because this shift does not necessarily keep the cost of fuel constant. At present, low carbon energy sources may well be more expensive, leading to some of the effects discussed above. The impact on jobs and the poor requires a proper accounting of the impact on labor demand both directly, in energy production, and indirectly in the rest of the economy as a result of changes of costs and price incentives.

Of course, investments in these sectors may be subsidized to ensure that the delivered energy costs themselves are not affected (while energy production is made relatively greener); these jobs may then seem 'secure'. [Huberty et al. \(2011\)](#) reviewed evidence from richer economies and correctly remarked that these jobs are effectively 'Keynesian' demand boost jobs whose sustainability can be questioned. Even if 'Keynesian' consequences are expected (in that these jobs are securely created due

to the growth impacts of these investments), there is still an opportunity cost to creating jobs using this investment capital. If these investments occur with subsidies or public capital, then the net impact on pro-poor jobs should be compared with alternative ways of stimulating positive impacts on the poor. It remains to be seen whether spending resources on greening energy production is necessarily the best option. In many circumstances, a trade-off between poverty reduction and greening energy production would seem most plausible. [Strand and Toman \(2010\)](#) offered a further discussion of trade-offs between 'green' investments beyond energy production with employment and poverty reduction.

Adaptation and Other Resilience-Enhancing Investments

A particular set of investments includes those intended to adapt economies to the new realities of climate change, including providing more resilience against extreme events. These are examples of a broader class of investments that can increase resilience (e.g., a greater capacity to mitigate the impacts of extreme weather with the current climate). It is sometimes difficult to distinguish these from the previous two categories of intervention, although they differ conceptually: the policies considered previously largely attempt to reduce the pressures on the environment (via pricing, regulation, and environmentally sensitive investments), whereas here we focus on green growth investments that are intended to reduce the socioeconomic hazards that are consequences of climate and other large-scale environmental change.

Again, the range of investments involved can be considerable. Their relative poverty impact will depend on similar considerations to those discussed above. How will they affect prices for natural resources, such as fuel or water? How will they affect the quality of resources available and the efficiency with which they are used? What are the labor demand consequences of these changes? Are these investments themselves labor intensive? It is worthwhile to reflect briefly on two further aspects: the consequences for structural and spatial transformation and thus for poverty.

There is considerable discussion on how to make agriculture, the main sector in which the poor are involved, and especially smallholder agriculture more resilient to extreme events and adapted to shifts in potential climate conditions (e.g., [Howden et al. 2007](#)). In a context of poverty in potentially affected areas, it is not surprising that the response is largely focused on increasing the local food security and self-sufficiency. Furthermore, to reduce the consequences of extreme events, the use of drought-resistant or salt-tolerant crops is promoted. These may be sensible policies, although they can also reduce mean returns. For example, many drought-resistant crops have low returns, leading to more security but also less poverty reduction ([Morduch 1995](#); [Dercon 1998](#)). Some alternative investments may reduce risks without affecting expected returns, such as flood protection infrastructure, although they may require substantial upfront investment.

More crucially, these policies appear to start from the premise that for poverty reduction, the best investment is to ensure that adaptation occurs where the poor are currently located. However, because this adaptation investment has opportunity costs, for a dynamic process of poverty reduction, investing in agricultural resilience for marginal, increasingly drought-prone areas may not be effective or efficient. Instead, investment could be used to speed diversification out of agriculture for affected populations, including via migration, in line with well-established routes out of poverty (Blake et al. 2011). Without such careful weighing of different alternatives for adaptation for the poor, the risk may be exacerbated that the poor will remain trapped in unsuitable areas and with low-value livelihoods and that they will find it increasingly difficult to move out of agriculture as part of the economic transformation. In short, processes consistent with spatial poverty traps, as discussed previously, are also a risk with forms of adaptation investments that focus disproportionately on marginal rural areas.

Related concerns should be highlighted when considering adaptation or other resilience-enhancing investments in urban settings. An example is infrastructure investment to avert the consequences of sea-level rises or to protect assets against extreme weather events in one of the many large coastal cities at risk. The greatest direct economic returns for these investments would be from protecting the business districts or the residential areas of the rich because these assets are of the highest value. Similarly, if relocation investments are necessary, they would most easily be accomplished by protecting businesses and the highest value assets in cities. Many of the poor are located in flood plains and unplanned settlements, and they would find it far costlier to move permanently, partly because their main assets are the houses in which they live without legal title. Large-scale infrastructure investments to protect them are hardly sensible because these marginal areas should not be places of urban settlement. Sensible urban resettlement policies would need to be designed with sensitive relocation strategies to ensure that poor populations are not spatially trapped (see, for example, Patel et al. 2002, for a description of sensitive resettlement in the case of Mumbai; other examples are in World Bank 2011).

Conclusion: Is Green Poverty a Possible Consequence of Green Growth?

There is no doubt that environmental change is affecting the poor disproportionately, whereas growth is essential for poverty reduction. Green growth is offered as a recognition of the need to sustain the growth required for poverty reduction while ensuring that environmental costs are internalized. Until recently, however, discussions of green growth have said little about *how* it is realized other than conventional measures for externality internalization and innovation. Even less has been said

about the potential consequences on poverty reduction from policies that steer an economy onto an environmentally sustainable trajectory.

In this paper, we do not take issue with whether growth can be sustained, although environmental costs must be internalized. We argue that internalizing environmental constraints may change the patterns of growth with distributional effects that are not necessarily pro-poor growth. We choose to focus on those currently in extreme poverty as experienced in some of the poorest countries. A focus on the currently extreme poor is not only relevant but can also be justified because their plight leaves little scope for long-term horizons (another way of saying that their discount rate is likely to be relatively high), whereas building resilience for future climate and environmental pressures requires higher wealth and incomes for the poor. A stronger sensitivity to intergenerational concerns could change much of what has been said in this paper, not least if the focus were on relative poverty or inequality rather than on extreme absolute poverty.

In developing this focus on the implications for the currently poor, the paper focuses on three elements of green growth policies: pricing and regulation to internalize environmental capital costs (such as via fuel or water pricing or regulation on water and air pollution); low-carbon and other environmentally sensitive public (or publicly stimulated) investments; and 'green' adaptation and other resilience-enhancing investments, particularly to effectively address the consequences of climate change. Four elements for assessing a green growth strategy with regard to its effectiveness to reduce poverty are offered: first, the efficiency gains from internalizing environmental externalities; second, its labor intensity because labor is the main asset of the poor; third, whether it contributes to a transformation of the livelihoods and sector of employment of the poor because most of the poor are either engaged in agriculture or in low-return informal sector self-employment; and finally, how it contributes to the spatial transformation of economies during growth and how it affects the opportunities for poverty reduction from internal migration and urbanization.

We argue that green growth could have important negative consequences for the currently poor that may outweigh the benefits for the poor from growth. In particular, *environmental pricing and regulation* may have considerable negative consequences for the poor as consumers and would require specific social protection measures to compensate for price rises. In terms of regulation, there is a risk of a bias in favor of the rich, which could exclude the poor from the benefits of regulation or even make them worse off, such as by displacing pollution.

Environmental pricing and regulation also affect the poor as producers because they may not have sufficient access to the wealth or human capital required to substitute for more expensive energy or other natural resources in their production processes. Furthermore, because the poor often only have their labor to sell, they depend on the labor intensity of growth for rapid poverty reduction linked to

growth. With higher costs of natural resources and other services of environmental capital, incentives are likely to be present to substitute human and physical capital for fuel and environmental capital. The labor intensity of such 'green' growth is crucial, however. More technology and capital intensive growth are unlikely to favor the poor.

Low carbon and environmentally sensitive investments can also have impacts that are not pro-poor. For example, although low carbon energy production may be more labor intensive, the size of the subsidy and/or public investment required may crowd out more pro-poor ways of spending resources. Other environmentally sensitive investments, such as promoting local food self-sufficiency or discouraging the movement of goods and people, do not necessarily benefit the poor because poorer areas may face increased risk of becoming trapped in low incomes and disconnected from higher growth areas.

The trade-off is even more stark when considering adaptation and other resilience-enhancing investments. In rural areas, these investments may induce the poor to adhere to lower returns and lower-risk livelihoods with little chance of escaping, even in increasingly marginal areas. This would follow if local adaptation is seen as the main option rather than also considering the critical need for economic transformation, including options related to migration and investing in urbanization. In urban settings, creating climate resilience may be targeted toward the most important economic assets, whereas the poor may end up trapped in environmentally marginal and unsuitable areas with little hope of being included in infrastructure plans for climate resilience.

Green growth is not necessarily bad for the poor. However, the key message of this paper is that promises that green growth will offer a rapid route out of poverty are not plausible, although there may be a less rapid exit than with more conventional growth strategies. To sustain growth, green growth also needs to be weighed in terms of its ability to reduce poverty. To sustain poverty reduction, green growth may involve giving up some environmental benefits to keep the growth-poverty elasticity high. Because poverty reduction remains at the top of the agenda, different shades of green may be needed. In particular, poverty reduction is a powerful force for giving those who are currently poor more resilience to the increasing risks of climate change. They should not be asked to pay the price for greening the planet.

Notes

Professor of Development Economics, University of Oxford and Chief Economist, Department for International Development, the UK government department involved in aid and development. Paper prepared for the World Bank project on Inclusive Green Growth. I am grateful for helpful comments by Mike Toman and the editor and referees of this journal. All views and errors are mine and should not be attributed to the UK's Department for International Development.

1. Although always a difficult exercise, some serious attempts have been made to quantify the costs of climate change for the poor, combining microeconomic evidence with modeling. World Bank (2010b) puts the price tag for adaptation at \$70-\$100 billion per year for the developing world, well outweighing the likely future welfare costs of climate change.

2. We must be careful with the evidence for poverty traps narrowly defined. In fact, it is difficult to find strong evidence. However, there is considerable evidence of poverty persistence in the sense of very slow escape and recovery and actions that lead to a perpetuation of poverty to avoid risk (Dercon 2009).

3. Note, however, that in an endogenous growth framework with growth externalities from the accumulation of certain factors of production, this may not *necessarily* be the case. For example, in some models, temporary reductions in growth due to increases in costs of production may actually have permanent consequences. For example, increasing environmental costs may reduce short-run resources for human capital investment, and overall growth may slow. Of course, there is no necessity to this either; it depends on the specific endogenous growth model applied (Aghion and Howitt 1998).

4. This same reason could also be the case for *appropriately* valued output growth if only *some* environmental regulations are put in without a full accounting of environmental capital such that not all factors are appropriately accounted for. A well-known welfare economics result from the theory of second best is that if only *some* and not all distortions are removed, then the allocation may actually become *less* efficient, both statically and dynamically. Applying this result, environmental regulations on some but not all forms of fuel-intensive transport may make the allocation less efficient and slow *appropriately* valued output growth.

5. The exception would be if the benefits for the poor that result from the new pricing of scarce resources, such as cleaner water or air, have a value to them that is larger than the costs of increased payments for public services (or for private goods whose costs are higher to reflect internalization of externalities outweigh the cost). This is possible, but they are likely to be potential benefits in the long run for higher short-run costs in the form of reduced real income.

References

- Aghion, P., and P. Howitt. 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Armendariz de Aghion, B., and J. Morduch. 2010. *The Economics of Microfinance*, 2nd Edition. Cambridge, MA: MIT Press.
- Black, R., N. W. Arnell, W. N. Adger, D. Thomas, and A. Geddes. 2013. "Migration, immobility and displacement outcomes of extreme events in nature and society." *Environmental Science and Policy* 27 (1): S32–43.
- Blake, R., A. Grimm, T. Ichinose, R. Horton, S. Gaffin, D. Jiong, L. Bader, and D. Cecil. 2011. "Urban Climate: Processes, Trends and Projections." In C. Rosenzweig, W. D. Solecki, S. A. Hammer, and S. Mehrotra, eds., *Climate Change and Cities: First Assessment Report of the Urban Climate Change Research Network*. Cambridge: Cambridge University Press.
- Carter, M., and C. Barrett. 2006. "The Economics of Poverty Traps and Persistent Poverty: An Asset-based Approach." *Journal of Development Studies* 42 (2): 178–99.
- Chan, K. W. 2012. "China, Internal Migration." In I. Ness, and P. Bellwood, eds., *The Encyclopedia of Global Migration*. Hoboken, NJ: Blackwell Publishing.
- Coady, D., M. El-Said, R. Gillingham, K. Kpodar, P. Medas, and D. Newhouse. 2006. "The magnitude and distribution of fuel subsidies: evidence from Bolivia, Ghana, Jordan, Mali and Sri Lanka." Working Paper WP/06/247. International Monetary Fund, Washington DC.
- Collier, P. 2007. *The Bottom Billion. Why the Poorest Countries are Failing and What Can Be Done About it*. Oxford: Oxford University Press.

- Connell, J. 2011. "Quantifying change in ecosystem services and exposure to hazards, and non-environmental drivers of change, in low-lying islands and small island states over the next fifty years that might be relevant to migration." Paper DR16, *Migration and Global Environmental Change*. UK Government Office for Science, Final Project Report, Foresight, Government Office for Science, United Kingdom.
- Deaton, A. 1991. "Savings and Liquidity Constraints." *Econometrica* 59 (5): 1221–48.
- Dell, M., B. Jones, and B. Olken. 2008. "Climate Change and Economic Growth: Evidence from the Last Half Century." Working Paper 14132. National Bureau of Economic Research, Cambridge, MA.
- _____. 2009. "Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates." *American Economic Review Papers and Proceedings* 99 (2): 198–204.
- Dercon, S. 1998. "Risk, Crop Choice and Savings: Evidence from Tanzania." *Economic Development and Cultural Change* 44 (3): 485–514.
- _____. 2002. "Income Risk, Coping Strategies and Safety Nets." *World Bank Research Observer* 17 (2): 141–66.
- _____. 2009. "Rural Poverty: Old Challenges in New Contexts." *World Bank Research Observer* 24 (1): 1–28.
- Department for International Development (DFID). 2010. "Climate-Resilient Growth: Preparing Growth Strategies for Climate Change." How To Note, A DFID Practice Paper. Department for International Development UK, London.
- Ellis, F. 2000. *Rural Livelihoods and Diversity in Developing Countries*. Oxford: Oxford University Press.
- Engel, D., and D. Kammen. 2009. "Green Jobs and the Clean Energy Economy." Thought Leadership Series #4, Copenhagen Climate Council, Copenhagen.
- Fomby, T., Y. Ikeda, and N. Loayza. 2012. "The Growth Aftermath of Natural Disasters," *Journal of Applied Econometrics* 28 (3): 412–34.
- Fujita, M., P. Krugman, and T. Venables, 2001. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.
- Gemenne, F. 2011. "Why the numbers don't add up: a review of estimates and predictions of people displaced by environmental changes." *Global Environmental Change* 21 (1): S41–9.
- Greenstone, M. 2002. "The Impacts Of Environmental Regulations On Industrial Activity: Evidence From The 1970 And 1977 Clean Air Act Amendments And The Census Of Manufactures." *Journal of Political Economy* 110 (6): 1175–219.
- Hallegate, S., G. Heal, M. Fay, and D. Treguer. 2011. "From Growth to Green Growth." Policy Research Working Paper WPS 5872. World Bank, Policy Research Department, Washington DC.
- Harrington, W., R. Morgenstern, and D. Velez-Lopez. 2012. "Tools for assessing the costs and benefits of green growth: the US and México." Paper presented at the Green Growth Knowledge Platform conference, January 12–13, Mexico City.
- Howden, M., J.-F. Soussana, F. N. Tubiello, N. Chhetri, M. Dunlop, and H. Meinke. 2007. "Adapting agriculture to climate change." *Proceeding of the National Academy of Science USA* 104 (50): 19691–6.
- Huberty, M., H. Gao, and J. Mandell. 2011. "Shaping the Green Growth Economy: a review of the public debate and the prospects for green growth." The Berkeley Roundtable on the International Economy, University of California at Berkeley.
- IPCC. 2007. "Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Core Writing Team." R. K. Pachauri, and A. Reisinger, eds., IPCC, Geneva, Switzerland.

- Loayza, N., and C. Raddatz. 2010. "The composition of growth matters for poverty alleviation." *Journal of Development Economics* 93 (1): 137–51.
- Massey, D., J. Durand, and N. J. Malone. 2002. *Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration*. New York: Russell Sage Foundation.
- Morduch, J. 1995. "Income smoothing and consumption smoothing." *Journal of Economic Perspectives* 9 (3): 103–14.
- Myers, N. 2002. "Environmental Refugees: A Growing Phenomenon of the 21st Century." *Philosophical Transactions of the Royal Society B* 357 (1420): 608–13.
- Noy, I. 2009. "The Macroeconomic Consequences of Disasters." *Journal of Development Economics* 88 (2009): 221–31.
- OECD. 2011. *Towards Green Growth*. Paris: Organisation of Economic Cooperation and Development.
- Patel, S., C. d’Cruz, and S. Burra. 2002. "Beyond evictions in a global city: people-managed resettlement in Mumbai." *Environment and Urbanization* 14 (1): 159–72.
- Raddatz, C. 2009. "The wrath of God: macroeconomic costs of natural disasters." Policy Research Working Paper Series WPS 5039. World Bank, Policy Research Department, Washington DC.
- Ravallion, M. 2000. "Growth, Inequality and Poverty: Looking Beyond Averages." *World Development* 21 (11): 1803–15.
- Ravallion, M., S. Chen, and P. Sangraula. 2007. "New Evidence on the Urbanization of Global Poverty." Policy Research Working Paper 4199. World Bank, Policy Research Department, Washington DC.
- Ravallion, M., and J. Jalan. 1997. "Spatial Poverty Traps?" Policy Research Paper no 1862. World Bank, Policy Research Department, Washington DC.
- Sachs, J. 2005. *The End of Poverty: Economic Possibilities for Our Time*. London: Penguin Press.
- Strand, J., and M. Toman. 2010. "Green stimulus, economic recovery, and long-term sustainable development." Policy Research Working Paper Series 5163. World Bank, Policy Research Department, Washington DC.
- UK Government Office for Science. 2011. *Migration and Global Environmental Change*. Final Project Report, Foresight. Government Office for Science, United Kingdom, London.
- Bank, World. 2010a. *World Development Report 2010: Development and Climate Change*. Washington DC: World Bank.
- . 2010b. *Economics of Adaptation to Climate Change*. Washington DC: World Bank.
- . 2011. *Building Resilience to Disasters—Delivering Results*. Global Facility For Disaster Reduction And Recovery Annual Report 2011. Washington DC: World Bank.
- . 2012. *Inclusive Green Growth – The Road to Sustainable Development*. Washington DC: World Bank.

Entry Regulation and the Formalization of Microenterprises in Developing Countries[#]

Miriam Bruhn and David McKenzie

The majority of microenterprises in most developing countries remain informal despite more than a decade of reforms aimed at making it easier and cheaper for them to formalize. This paper summarizes the evidence on the effects of entry reforms and related policy actions to promote firm formalization. Most of these policies result in only a modest increase in the number of formal firms, if there is any increase at all. Most informal firms appear to not benefit on net from formalizing. As a consequence, ease of formalization along will not induce most of them to become formal. Increased enforcement of rules can increase formality. Although there is a fiscal benefit of doing this with larger informal firms, it is unclear whether there is a public rationale for attempting to formalize subsistence enterprises. Informality, Entry regulation, Self-employment, Regulatory enforcement. JEL codes: O12, O17, D22, L26

Making it easier to formally register a business is the area of business regulatory policy that has seen most attention from policymakers over the past decade. The *Doing Business* project of the World Bank found that 368 reforms took place in 149 economies between 2003 and 2012. As a result of these reforms, the world average time to start a business has decreased from 50 to 30 days, and the cost of starting a business is one-third of what it was previously ([World Bank 2013](#)).

However, despite these reforms, the majority of businesses in most developing countries remain informal. By formality, we mean whether a business is registered for relevant municipal licenses and with the tax department. For example, Brazil simplified its taxation system for small businesses, and the state of Minas Gerais introduced a “one-stop shop”, which has become a popular way of streamlining the registration process for firms worldwide. After these efforts, a government survey found that 72 percent of enterprises in the state remained informal ([Andrade et al. 2013](#)).

Informality rates are highest for low-scale enterprises; [de Mel et al. \(2013\)](#) report that only 39 percent of Sri Lankan businesses with one worker are registered for taxes, compared to 68 percent of those with five workers and 78 percent of those with 10 workers. Most small firms in developing countries have only one or two workers. For example, in Mauritius, of the 91,980 units with nine or fewer workers, 70.6 percent have only one or two workers, and only 6.2 percent have five to nine workers ([McKenzie 2011](#)). By virtue of the sheer quantity of one- and two-person firms and their much greater likelihood of being informal, most of the stock of informal firms in developing countries is low-scale enterprises consisting of the owner and, at most, one other paid worker.

The key question for policymakers is then what, if anything, they should attempt to do about this vast quantity of small-scale informal firms. One view, popularized by [Hernando de Soto \(1989\)](#), is that these informal firm owners would like to be formal but costly regulations and bureaucracy prevent them from doing so, causing a productivity loss for these firms. The implication is that policymakers should attempt to reduce these barriers and bring these firms into the formal sector. A competing view, associated with [Maloney \(2004\)](#) and others, is that firms that would benefit from formalizing do so, whereas smaller and less productive firms rationally opt out of the formal sector because they perceive little benefit from becoming formal. A related view is the segmented labor market perspective, in which the informal sector is seen as “the free entry sector of last resort” catering to individuals who would prefer paid employment but are unable to find formal wage jobs ([Fields 2004](#)). The implication is that it is not privately optimal for these firms to formalize, so policymakers should only attempt to bring them into the formal sector if there is a compelling public rationale for doing so.

In practice, the large number of informal firms with a few workers at most is likely to include a mix of different types of firms, including subsistence enterprises and individuals who would rather be wage workers as well as potential entrepreneurs with the desire and, sometimes, the ability to grow their businesses over time. [De Mel et al. \(2010\)](#) classify the self-employed in Sri Lanka using statistical methods designed for species classification in biology and find that between one-quarter and one-third of the self-employed appear to share characteristics that give them the potential to be owners of somewhat larger firms, whereas the remaining two-thirds to three-quarters look much more like wage workers. Our focus is the role of business regulations on the formalization decisions of this group of small-scale enterprises before we turn to a discussion of the extent to which policy can differentiate among them. We note that although our focus is regulatory policy, governments have a number of other policy tools, such as credit and labor market policies, that also affect the decisions of individuals to operate firms at all and, if so, to do so formally or informally.¹

Discussion of Key Findings

We divide our discussion of key findings into two subsections. The first considers the extent to which business regulations appear to be barriers to firms *becoming formal* and whether efforts to make it easier for firms to formalize or more difficult to remain informal have succeeded in bringing low-scale enterprises into the formal sector. The second subsection examines the extent to which *being formal* benefits these types of firms. Attempts to answer both sets of questions must address the fact that places that change regulations may differ systematically from those that do not and that firms that choose to become formal are likely to differ in many unobservable ways from those that choose to remain informal.

Measuring the Impacts of Formalization Efforts and of Becoming Formal

Simple comparisons of formal and informal firms usually reveal that formal firms are more productive and profitable. A series of high-profile sector studies by the McKinsey Global Institute comparing the operation of formal and informal firms in several countries around the world concluded that informality has a very negative impact on productivity, accounting for nearly 50 percent of the overall productivity gap between countries such as Portugal and Turkey and the United States (Farrell 2004). La Porta and Shleifer (2008) use World Bank firm surveys to compare the productivity of informal and formal firms and find the formal firms to have substantially higher productivity levels.

However, such estimates assume that firms do not choose whether to be formal. In practice, however, firms are likely to weigh the potential benefits of being formal (greater access to credit, greater scope for marketing, participation in government contracts and programs, reputation effects, less risk of fines) with the costs of becoming formal (the time and money costs of registering) and of being formal (ongoing taxes, accounting costs, paperwork). As a result, individuals who see large benefits to becoming formal will do so, whereas less productive and smaller firms that see little benefit in formalizing will likely choose to remain informal. As a result, lower productivity is likely to be as much a cause as a consequence of informality, and simple comparisons of formal and informal firms will greatly overstate the productivity benefits of formalizing.

To account for the likelihood that firms that choose to become formal differ from those that remain informal, ideally, one would randomly choose some firms to make formal and leave similar firms informal before comparing the two groups. This approach is known as the randomized experimental approach and has been used in a number of attempts to formalize firms by providing them with information and lower costs of registering (Jaramillo 2009; Alcázar et al. 2010; de Mel et al. 2013; De Giorgi and Rahman 2013; Andrade et al. 2013).

When randomization has not been possible, studies have used a variety of non-experimental econometric methods designed to ensure that they can measure the causal effect of policy efforts to spur formality or the consequences of formalizing. For example, [McKenzie and Sakho \(2010\)](#) use an instrumental variables method to measure the consequences of being formal in Bolivia, arguing that firms that are closer to the municipal tax office have more information and lower costs of registering than similar firms located elsewhere in the same city. [Bruhn \(2011\)](#) uses a difference-in-differences method, comparing changes in the formalization rates of firms in municipalities in which entry regulation was simplified earlier to those in similar municipalities in which the reform came later. These non-experimental methods require more assumptions to establish causality than are required for randomized experiments, but they also represent serious attempts to address the self-selection of firms into formal or informal status.

The following section discusses findings from both randomized and non-experimental studies on the causal effect of policies to promote firm formalization (see table 1 for a summary of the key results from these studies).

Becoming Formal

The Effect of Business Entry Reforms on the Number of Firm Registrations. A number of studies have examined the impact of business entry reforms on firm registration in different Latin American countries, exploiting cross-time and cross-municipality variation in the implementation of these reforms. A common element in these reforms is that they opened one-stop shop service points and thus eliminated the need to visit several different government offices for completing the registration process, lowering the time and/or cost needed to register a business. These studies find that a large reduction in the cost and time taken to register a firm leads to a modest increase in the number of formal firms.

A reform in Mexico, which was implemented in some of the most populous and economically developed municipalities starting in 2002, reduced the number of days needed to start a business from 30.1 to 1.4. [Bruhn \(2011\)](#) uses data on individuals' employment status from the Mexican Labor Market Survey (ENE) to show that this reform increased the number of registered business owners by 5 percent. [Kaplan, Piedra, and Seira \(2011\)](#) use administrative data and find that the same reform increased the number of new firm registrations with the Mexican Social Security Institute (IMSS) by 5 percent.

The Colombian government implemented a one-stop shop that reduced the time required to register a business from 55 to less than 9 days and lowered registration fees by 30 percent. [Cárdenas and Rozo \(2007\)](#) use administrative data from Chambers from Commerce in six major cities to show that the reform led to a 5 percent increase in business registrations.

Table 1. Summary of Studies on the Causal Effect of Policies to Promote Firm Formalization

<i>Study</i>	<i>Policy or program studied</i>	<i>Econometric Approach</i>	<i>Main results</i>
Panel A: Effect of business entry reforms on number of firm registrations			
Bruhn (2011), Kaplan, Piedra, and Seira (2011)	One-stop shop (combining municipal, state, and federal business registration procedures) in urban areas in Mexico	Difference-in-Differences	Reform increased business registrations approximately 5 percent and increased employment; Bruhn shows that the increase in registered businesses was mainly due to previous wage earners opening new businesses
Bruhn and McKenzie (2013)	One-stop shop in less populous municipalities in Minas Gerais, Brazil	Difference-in-Differences	Program led to a reduction in number of firms registering during the first two months of implementation, with no subsequent increase
Cárdenas and Rozo (2007)	One-stop shop in six major cities in Colombia	Difference-in-Differences	Reform increased business registrations by 5 percent
Panel B: Effect of business entry reforms on formalization of informal firms			
Bruhn (2013)	One-stop shop in urban areas in Mexico	Difference-in-Differences	Some informal business owners become wage workers due the reform, some register their business, but these effects are small
Mullainathan and Schnabl (2010)	Municipal licensing reform in Lima, Peru	Before-After	Reform increased number of provisional licenses issued to informal firms, but many firms do not renew their license later
Panel C: Effect of information, waived costs, and enforcement on formalization of informal firms			
Alcázar et al. (2010)	Offer a subsidy for the cost of obtaining a municipal license to informal firms in Lima, Peru	Randomized Experiment	Subsidized cost offer led to 10 to 12 percent of informal firms obtaining a municipal license
Andrade et al. (2013)	Three interventions for informal firms in Belo Horizonte, Brazil (a) deliver brochures with information about registration process and potential benefits; (b) waive registration costs; (c) receive municipal inspector	Randomized Experiment	Information and waived registration costs had no effect on formalization rate; municipal inspections increased formalization rate by 22 to 27 percentage points
de Mel et al. (2013)	Provide information and reimburse registration costs for informal firms in Sri Lanka	Randomized Experiment	Information and cost reimbursement had no effect on formalization rate
De Giorgi and Rahman (2013)	Deliver brochures with information to informal firms in Bangladesh	Randomized Experiment	Information had no effect on formalization rate

Source: Authors' summary of the literature

Entry reforms appear to have less of an effect on business registrations in less populous and more remote areas. [Bruhn and McKenzie \(2013\)](#) analyze the impact of the Minas Fácil Expresso program in the state of Minas Gerais, Brazil. This program aimed to extend the benefits of a one-stop shop to less populous municipalities, thereby removing the need for firms to register separately at their municipality and then travel to register with state and federal governments in a larger town or city. They find that this program actually led to a reduction in the number of firms registering during the first two months of implementation (perhaps due to officials learning to use the new system), with no subsequent increase.

The majority of these studies cannot disentangle whether the increase in registrations is due to informal firms registering or to new firms being created by individuals who did not previously run businesses. The labor market survey used by [Bruhn \(2011\)](#) is an exception; she finds that the increase in registered businesses was due to wage earners opening new businesses.

In contrast to the literature on business entry reforms, we are aware of very few studies that rigorously examine the impacts of reforms in the ongoing regulations facing formal firms on the decision to become formal. An exception is the Brazilian SIMPLES reform, which reduced the complexity and rates of taxes facing small firms without changing the entry procedures. [Monteiro and Assunção \(2012\)](#) use a difference-in-differences approach to compare the legal status of firms in sectors affected and unaffected by the reform that were created before and after the program. They find no significant overall effect, but they do find that retail firms created after the reform appear to be more likely to be formal. Complementary evidence comes from [Waseem \(2013\)](#), who examines the converse policy reform: dramatically increasing the tax burden facing small partnerships in Pakistan. He finds a large response: 57 percent of taxpayers leave the formal system, and those remaining substantially reduce their reported earnings.

The Effect of Business Entry Reforms on the Formalization of Informal Firms. A follow-up study by [Bruhn \(2013\)](#) examines the impact of the business registration reform in Mexico on informal business owners in more detail. This paper takes into account that informal business owners are heterogeneous and that some may be looking for a wage job instead of aiming to continue or to grow their business. Bruhn divides the group of informal business owners into those who have personal characteristics similar to formal business owners and those who have the characteristics of wage earners and examines the impact of the reform for each of these groups. The results show that wage earner types are less likely to register their businesses due to the reform but are more likely to become wage workers because the reform created jobs. In contrast, informal business owners with characteristics similar to formal business owners are more likely to register, but only in municipalities with high pre-reform constraints on formal entrepreneurship. However, the

effects are relatively small, and most informal business owners remain informal even after the reform.

Some countries experience a large spike in firm registrations after the introduction of a business registration reform. For example, [Mullainathan and Schnabl \(2010\)](#) document that the number of firm registrations increased from 1,758 in the year before a municipal licensing reform in Lima, Peru, to 8,517 in the year after a reform. They conduct a survey to show that approximately 75 percent of the firms that registered in the year after the reform were previously informal and that 25 percent were new firms. However, the majority of newly issued licenses were provisional, meaning that they had to be renewed every year. Many firms do not seem to have renewed their licenses after the year was over because the number of registrations dropped to 3,500 in the second year after the reform, suggesting that the increase in formalization was temporary.

The Effect of Providing Information about How to Register or Lowering the Cost of Formalizing on Formalization Rates

Surveys of microenterprises typically show that many informal firms are not very well informed about either the process of formalizing or the costs involved. They often overestimate the time and cost involved, particularly following efforts to simplify the registration process. For example, in Sri Lanka, [de Mel et al. \(2013\)](#) report that only 17 percent of informal firm owners know the cost of registering. Most believe that the process of registering takes more than a month compared to one week or less in practice, and only 2 percent know that lower incomes are not liable for business income taxes. In Bolivia, [McKenzie and Sakho \(2010\)](#) report that only one-third of informal owners know the location of the tax office (which is the registration location) and that only 10 percent have even heard of the commerce registry.

A natural policy response to this lack of information is to attempt to provide information to firms about how to register and about the possible benefits of formalizing. However, two randomized experiments which have done this have found no resulting increase in formalization. In Belo Horizonte, Brazil, [Andrade et al. \(2013\)](#) gave a glossy brochure prepared by a State Government marketing team to 208 informal firms and found that firms that received this brochure were not more likely to register over the following year. In Bangladesh, [de Giorgi and Rahman \(2013\)](#) had field staff deliver information and brochures to 1,500 informal firms. They found that fewer than 5 percent registered and that this number was not larger than the number among the control group firms not given the information.

Two randomized experiments have also investigated the impact of lowering the cost of formalizing along with this information. In Sri Lanka, [de Mel et al. \(2013\)](#) provided information and agreed to reimburse the full cost of registering for taxes

(approximately US\$10) to 104 firms, with only one firm accepting the offer and formalizing. In Belo Horizonte, Brazil, [Andrade et al. \(2013\)](#) provided information, had the government waive all registration costs and sanitary and inspection taxes for the first year (approximately US\$200), and provided free mandatory accounting services for the first year (approximately US\$1800). This package of information and reduced costs was offered to 255 informal firms, again with only one firm accepting the offer. Lowering the cost of registering for taxes thus appears to have very little impact on the registration of existing firms. One reason for this finding may be that the ongoing costs of being registered, including tax payments, accountant costs, and other fees, deter firms from registering, even if the process of doing so is simplified.

In many countries, the formalization process requires registration with more than one government agency and thus involves different ongoing costs of being registered with each agency. A municipal license is often the first license that informal firms receive. Because a municipal license typically requires only a small annual license fee but no sizeable local tax payments, it is less expensive for firms than registering with the national (and state) governments for business taxes. Therefore, we might expect firms to be more willing to obtain this license if its cost is lowered. Evidence along these lines comes from a randomized experiment conducted by [Alcázar et al. \(2010\)](#) in Lima, Peru. They follow up on a reform that reduced the time to obtain a municipal license from 160 days to 1.6 days by attempting to induce 300 informal firms to obtain a municipal license by offering a subsidy of between 27 and 35 percent of the cost of the license (or approximately US\$30). They find that the subsidized cost offer led approximately 10 to 12 percent of informal firms to obtain a municipal license.

Does Increased Enforcement Cause More Firms to Formalize?

Policy and research attention has focused on attempting to simplify the regulations facing firms, but much less attention has been given to the effects of better enforcement of the regulations which are in place. A sizeable minority of informal firms report having received some type of inspection visit over the past year (most often from the municipality), but in the countries where we have collected data, very few firms report being fined. For example, in Sri Lanka, 43 percent of informal firms were visited by a municipal official, but only 0.8 percent paid any fines (or bribes); in Brazil, 33 percent of firms were visited by a municipal official, but only 2.1 percent paid a fine; and in Bolivia, fewer than 2 percent of informal firms report having been fined.

A randomized experiment by [Andrade et al. \(2013\)](#) shows that more enforcement by inspectors can induce some informal firms to become formal. They randomly allocated 577 firms to receive inspection visits from municipal inspectors.

Their results point both to the difficulties in inspecting informal firms (the inspectors were unable to locate some of these firms or talk to their owners) and to the potential impact on formalization: the authors estimate that between 22 and 27 percent of firms that received an inspection as a result of their intervention registered with the municipality.

Enforcing Formality in the Tax and Labor Domains

In practice, even after firms formally register, they face a number of other regulations, and they choose the extent to which they comply with these regulations. In particular, formally registered firms may be partly informal with regard to their compliance with tax and labor regulations. Several studies examine the role of enforcement in increasing formality along these dimensions.

[Gordon and Li \(2009\)](#) provide an overview of tax policy in developing countries and note that the possibility of evasion by not reporting cash transactions helps explain why tax policies in developing countries take the structure they do. One approach to enforcement is to increase auditing. [Pomeranz \(2013\)](#) examines this approach in the context of the Chilean VAT system. The Chilean tax authority sent letters intended to generate a perceived increase in audit probability to randomly selected firms. Pomeranz finds that these letters increase the median declared income for firms by 12 percent, with this effect concentrated in firms that sell to final consumers (because other firms already effectively face enforcement through the VAT paper trail). The effect is larger for small firms, which is consistent with larger firms' lower likelihood of evading taxes.

A complementary approach is to establish incentives for customers to demand that firms be formal. One way to do this is a lottery tax receipt system in which each tax receipt is printed with a lottery number. A drawing for monetary prizes gives customers an incentive to ask for receipts. Such a system has been used in Taiwan, Korea, China, Brazil, and Puerto Rico. [Wan \(2010\)](#) compares changes in tax revenues in districts in China that introduced this reform to those that did not and finds that the introduction of this tax receipt lottery increased sales tax revenue by 17 percent. [Naritomi \(2013\)](#) uses a difference-in-differences approach to compare changes in tax revenue in Sao Paulo from retail stores (which were subject to consumer monitoring) to those in wholesale stores (which were not) and finds that the introduction of a tax receipt lottery increases tax revenue by 23 percent.

Increased enforcement has also been found to affect the extent to which firms comply with labor regulations. [Almeida and Carneiro \(2012\)](#) study the impact of greater enforcement of labor regulations in Brazil by using geographical variation in the likelihood of enforcement. They find that an increase in the enforcement of mandated benefits in the formal sector leads to a reduction in formal wages and an increase in formal sector employment. However, the authors have data on workers,

not firms, so they are not able to determine how much of this impact comes from partially formal firms that now register more of their existing workforce. [Kumler et al. \(2013\)](#) examine another dimension of labor informality in which firms register workers, but underreport the wages they pay them, to reduce their payroll tax burden. They find that a social security reform that gave workers incentives to monitor what firms reported led to less evasion among firms reporting the wages of affected workers.

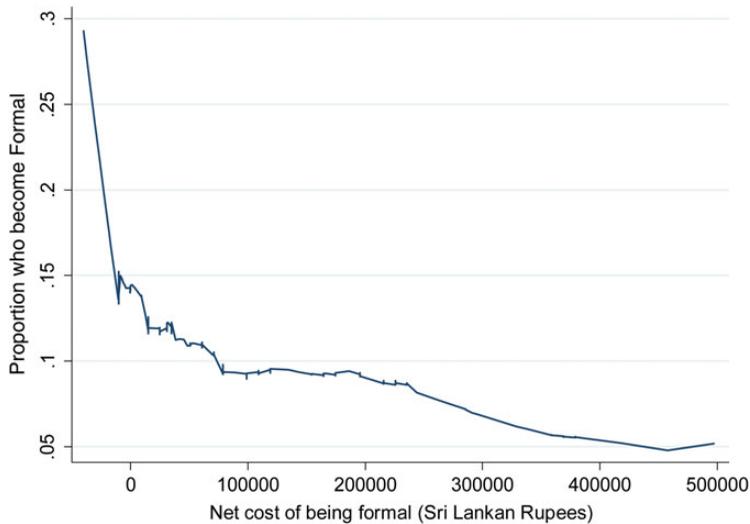
Despite Formalization Policies, the Majority of Small-scale Firms in Developing Countries Remain Informal

As noted in the motivation section, the majority of self-employed individuals and firms with only one or two workers remain informal in many developing countries, despite the decade of reforms intended to simplify business registrations. All of the studies described above that have provided information and lowered the cost of formalizing have found that, at most, a small fraction of informal firms choose to formalize when given help to do so. This finding contrasts with the view of [de Soto \(1989\)](#) and suggests that in many cases, firms decide that the benefits of formalizing do not exceed the costs. An experiment in Sri Lanka provides the strongest evidence for this view, showing that more firms become formal as the benefits of doing so increase.

[De Mel et al. \(2013\)](#) offered informal firms different amounts of money to become formal, ranging from just enough to cover the cost of tax registration to 40,000 Sri Lankan Rupees (US\$350). The largest amount was equivalent to approximately two months' profits for the median firm. They find that although no additional firms registered when only the cost of registering was covered, 20 percent of firms registered when offered 10,000 Sri Lankan Rupees, and 47 percent registered when offered the maximum amount of 40,000. By combining these incentives to register with the initial registration costs and the discounted value of future tax payments facing these firms, these authors are able to examine how the demand for formalizing varies with the net cost of registration. We observe a steep downward demand curve (Figure 1), with the proportion of firms that are willing to become formal dropping quickly as the cost of formality increases. We expect firms to formalize if they view the net benefits as exceeding these costs, so another way of reading this figure is to note that fewer than 15 percent of informal firms seem to view the net benefits of being formal as greater than zero.

[De Mel et al. \(2013\)](#) note that many of the firms that did not formalize when offered the maximum amount cited other barriers, such as land titling issues (the costs of overcoming these other barriers are not included in the net cost), whereas others said that they did not believe that the benefits would outweigh the costs even

Figure 1. The Demand for Formality in Sri Lankan Firms as a Function of the Net Cost of Formalizing



Source: Based on [de Mel et al. \(2013\)](#), Figure 3.

with this monetary incentive. Next, we consider the benefits to firms of being formal.

Being Formal

Informal firms are often referred to as “operating in the shadows”, afraid to market themselves widely or locate in visible locations for fear of attracting the attention of the law. In addition, firms that do not issue tax receipts may not be able to sell to certain customers, such as the government or larger firms. As a result, one of the channels through which formalizing may help firms is by expanding their customer base. This is the channel that has the most support from various studies in the literature. In Bolivia, [McKenzie and Sakho \(2010\)](#) find that firms that become formal as a result of being located slightly closer to the tax office issue more tax receipts and have greater sales. In Brazil, [Fajnzylber et al. \(2011\)](#) find that firms that opened just after the introduction of the SIMPLES program, which simplified regulation by combining different types of tax registrations and payments, were more likely to operate in a permanent location, although this finding could reflect a change in the types of individuals that decide to open firms after the reform rather than a causal impact of formalizing. In Sri Lanka, firms that formalize as a result of experimental inducement are 26 percentage points more likely to advertise ([de Mel et al. 2013](#)).

Two other main channels through which formalization is often hypothesized to benefit firms are improved relationships with the financial sector and increased access to government benefits. Existing studies find no evidence of these effects. In Sri Lanka, [de Mel et al. \(2013\)](#) find that firms that formalize are not more likely to obtain a business bank account or business loan, make sales to the government, or participate in a government program. In Bolivia, [McKenzie and Sakho \(2010\)](#) find no impact on the likelihood of a bank loan or of sales to the government. A likely explanation is that many informal firms would not receive credit even if they did register, whereas those for whom registration is the only barrier preventing them from obtaining credit have already chosen to become formal. The same situation is likely to apply to government contracting: most small-scale firms are unlikely to sell to the government, and those that do formalize when this need arises.

Consistent with the viewpoint that most firms rationally choose whether to formalize by comparing the expected benefits of becoming formal to the costs of doing so, studies have found that informal firms that have been induced to become formal by subsidized costs have experienced little benefit from doing so, on average. [Alcázar et al. \(2010\)](#) find no significant impact of obtaining a municipal license on firm profits or revenues in Lima. [McKenzie and Sakho \(2010\)](#) find that firms that would formalize if they were slightly closer to the tax registration office would appear to increase profits by doing so, but the average informal firm would not. [De Mel et al. \(2013\)](#) find a significant increase in average profitability, but they show that this average increase is driven by a handful of firms growing substantially after formalization, whereas most firms show no improvement.

Limitations and Gaps

The past five years have seen a number of innovative experimental studies measuring the impacts of policy efforts to increase formality along with several relatively rigorous non-experimental studies. However, a number of limitations and gaps remain:

- The existing research comes from only a few countries concentrated in Latin America and Southeast Asia. Thus, it is unclear to what extent the findings can be generalized to regions such as Africa, the Middle East, and Eastern Europe, where little evidence is currently available.
- Existing research has largely focused on the registration and tax regulations facing small-scale enterprises, with little work on the importance of other regulatory barriers for firm growth. The experimental studies by [de Mel et al. \(2013\)](#) and [Alcázar et al. \(2010\)](#) note that land zoning and titling regulations are key barriers which that appeared to prevent firms formalizing, and more work is needed on the impacts of policy efforts to improve these types of regulations.

In addition, more evidence is needed on the role of reductions in the ongoing costs and complexity of being formal on the initial formalization decision.

- There is a need for better measurement of both the regulations facing small-scale enterprises and their interactions with these regulations. For example, the *Doing Business* project attempts to measure key regulatory burdens facing businesses around the world, but it focuses on limited liability companies, not small-scale enterprises. The ease of starting a business measure considers a company with start-up capital of 10 times per capita income, turnover of at least 100 times per capita income, and between 10 and 50 employees ([World Bank 2013](#)). This measure may not reflect the regulations facing small businesses. Few countries have data on the characteristics of firms that register and whether they are new or existing firms.
- Existing research has largely focused on the private costs and benefits of formalizing for informal firms. Two key public policy rationales for formalization require more research. The first is the claim that informality leads to inefficient resource allocation, since informal firms who don't pay taxes can compete away customers from more efficient tax-paying formal firms. The second is the view that formalization is socially optimal because it increases government revenues and a culture of respect for the rule of law.

Summary and Policy Advice

Efforts to dramatically lower the cost of registration and simplify the registration process have left most small-scale enterprises operating informally in many developing countries. Although these firms would gain some benefits from formalizing in terms of the ability to advertise more and issue tax receipts, they typically view these benefits as smaller than the costs of becoming formal and the ongoing taxes they must pay for being formal. Existing evidence suggests that most of these firms are making what is a privately optimal decision, and so the policy rationale for attempting to bring small-scale informal firms into the formal sector should not rely on the belief that doing so will provide benefits to these firms. This does not mean that there is no cost of inefficient and costly regulations; requiring firm owners to spend extra time and effort in registering their firms is usually a purely social cost. Therefore, efforts to simplify regulations will benefit those firms that choose to go through this process

A key issue for policymakers is whether there is a public rationale for attempting to formalize small-scale firms. There are several compelling reasons to attempt to bring larger and more profitable informal firms into the formal system. First, the need of most developing countries to widen the tax base is likely to include a public rationale for collecting taxes from relatively well-off owners of informal firms and for the revenue collected from them to justify the costs of formalizing. Indeed,

[Andrade et al. \(2013\)](#) suggest that inspecting informal firms in Brazil that earn an average of US\$1000 a month in profits would formalize more than enough firms to pay for the costs of such enforcement. Second, these larger and more successful informal firms are more likely to be the ones competing with formal firms for customers. Therefore, ensuring that such firms also become formal may reduce unfair competition that prevents more efficient formal firms from growing faster. The challenge is then how to encourage the formalization of such firms. Based on the evidence reviewed here, reducing the cost and complexity of registration seems a necessary but not sufficient step. Policymakers also need to increase enforcement of the simplified rules and perhaps to experiment with innovative approaches to encourage suppliers or customers to demand formality. One approach that has been attempted in several countries is to link each tax receipt number to a lottery so that customers have an incentive to demand a tax receipt for each transaction.

What about subsistence enterprises? Existing evidence seems to suggest that such firms see no benefits from formalizing and would typically contribute very little to taxes if they did formalize. They may still compete with larger firms, but in the absence of other job opportunities for these individuals, the government may prefer to leave them alone rather than to have them close down. The only remaining public rationale for attempting to bring them into the formal sector is that the presence of so many informal firms may send a message to the public that obeying the law is optional and may dissuade more prosperous informal firms from formalizing. This “broken-windows” theory of crime has some support in developed countries ([Keizer et al. 2008](#)), but we are unaware of any such studies or efforts applied to informality in developing countries. An alternative approach used in some countries is to write the law in a way that does not require firms with income below a certain threshold to register, making them in compliance with the law. However, unless such a threshold is set very high, it is still likely that many firms above the threshold will choose not to register.

Finally, given the limited existing evidence base and examples of policies that have not accomplished their intended effects, policymakers attempting to innovate in this area would be well served to build in rigorous impact evaluations, in order to measure whether any new efforts are achieving their desired goals. In addition to regulatory policy changes, this should include evaluations of the effects of other government policy choices that seek to change the relative costs and benefits of being formal, such as credit and labor market policies.

Note

[#]Miriam Bruhn and David McKenzie are Senior Economist and Lead Economist respectively in the Finance and Private Sector Unit of the World Bank’s Development Research Group.

Email: mbruhn@worldbank.org and dmckenzie@worldbank.org (corresponding author). Address: MSN MC3-307, 1818 H Street N.W., Washington D.C. 20433, U.S.A. The authors gratefully acknowledge funding for their work on informality from DFID, the Ewing Marion Kauffman Foundation, the Knowledge for Change Trust fund, the State Government of Minas Gerais, and the World Bank. We thank the editor and three anonymous reviewers for helpful comments.

1. Levy (2008) also highlights the role of social policies such as the provision of health insurance in determining the relative attractiveness of the formal versus the informal sectors to workers, which, in turn, affects incentives to enter self-employment.

References

- Alcázar, L., R. Andrade, and M. Jaramillo. 2010. "Panel/tracer study on the impact of business facilitation processes on enterprises and identification of priorities for future business enabling environment projects in Lima, Peru – Report 5: impact evaluation after the third round." Report to the International Finance Corporation. World Bank Group, Washington, DC. Processed.
- Almeida, R., and P. Carneiro. 2012. "Enforcement of labor regulation and informality." *American Economic Journal: Applied Economics* 4 (3): 64–89.
- Andrade, G. H., M. Bruhn, and D. McKenzie. 2013. "A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms." World Bank Policy Research Working Paper 6435. World Bank, Policy Research Department, Washington, DC.
- Bruhn, M. 2011. "License to sell: the effect of business registration reform on entrepreneurial activity in Mexico." *Review of Economics and Statistics* 93 (1): 382–6.
- . 2013. "A Tale of Two Species: Revisiting the Effect of Registration Reform on Informal Business Owners in Mexico" *Journal of Development Economics* 103: 275–83.
- Bruhn, M., and D. McKenzie. 2013. "Using Administrative Data to Evaluate Municipal Reforms: An Evaluation of the Impact of Minas Fácil Expresso" World Bank Policy Research Working Paper 6358. World Bank, Policy Research Department, Washington, DC.
- Cárdenas, M., and S. Rozo. 2007. "La informalidad empresarial y sus consecuencias: ¿Son los CAE una solución?" Documento de Trabajo 38. Fedesarrollo, Bogota, Colombia.
- De Giorgi, G., and R. Rahman. 2013. "SME's Registration: Evidence from an RCT in Bangladesh." Stanford University, Stanford, CA, and World Bank, Washington, DC. Processed.
- De Mel, S., D. McKenzie, and C. Woodruff. 2010. "Who are the Microenterprise Owners?: Evidence from Sri Lanka on Tokman v. de Soto." In J. Lerner, and A. Schoar, eds., *International Differences in Entrepreneurship*. Boston, MA: National Bureau of Economic Research.
- . 2013. "The demand for, and consequences of, formalization among informal firms in Sri Lanka." *American Economic Journal: Applied Economics* 5 (2): 122–50.
- De Soto, H. 1989. *The Other Path*. New York: Harper and Row Publishers.
- Farrell, D. 2004. "The hidden dangers of the informal economy." *The McKinsey Quarterly* (Number 3). http://www.mckinseyquarterly.com/article_page.aspx?ar=1448&L2=7&L3=10&srid=6&gp=1
- Fajnzylber, P., W. Maloney, and G. Montes-Rojas. 2011. "Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program." *Journal of Development Economics* 94: 262–76.
- Fields, G. 2004. "A guide to multisector labor market models." Cornell University, School of Industrial and Labor Relations, Ithaca, NY. <http://digitalcommons.ilr.cornell.edu/workingpapers/86/>. Retrieved 23 May 2013.

- Gordon, R., and W. Li. 2009. "Tax structures in developing countries: many puzzles and a possible explanation." *Journal of Public Economics* 93 (7-8): 855–66.
- Jaramillo, M. 2009. "Is there demand for formality among informal firms? Evidence from microfirms in downtown Lima." German Development Institute Discussion Paper 12/2009. German Development Institute, Bonn, Germany.
- Kaplan, D. S., E. Piedra, and E. Seira. 2011. "Entry regulation and business start-ups: Evidence from Mexico." *Journal of Public Economics* 95 (11-12): 1501–15.
- Keizer, K. S. Lindenberg, and L. Steg. 2008. "The spreading of disorder." *Science* 322 (5908): 1681–5.
- Kumler, T., E. Verhoogen, and J. Frías. 2013. "Enlisting employers in improving payroll tax-compliance: Evidence from Mexico." Mimeo. Columbia University.
- La Porta, R., and A. Shleifer. 2008. "The unofficial economy and economic development." *Brookings Papers on Economic Activity* 47 (1): 123–35.
- Levy, S. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Washington, DC: Brookings Institution.
- Maloney, W. 2004. "Informality Revisited." *World Development* 32 (7): 1159–78.
- McKenzie, D. 2011. "How can we learn whether firm policies are working in Africa? Challenges (and solutions?) for experiments and structural models." *Journal of African Economics* 20 (4): 600–25.
- McKenzie, D., and Y. S. Sakho. 2010. "Does it pay firms to register for taxes? The impact of formality on firm profitability." *Journal of Development Economics* 91 (1): 15–24.
- Monteiro, J., and J. Assunção. 2012. "Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises." *Journal of Development Economics* 99: 105–15.
- Mullainathan, S., and P. Schnabl. 2010. "Does Less Market Entry Regulation Generate More Entrepreneurs? Evidence from a Regulatory Reform in Peru." In J. Lerner, and A. Schoar, eds., *International Differences in Entrepreneurship*. Chicago: National Bureau of Economic Research.
- Naritomi, J. 2013. "Consumers as tax auditors." Mimeo. Harvard University.
- Pomeranz, D. 2013. "No taxation without information: Deterrence and self-enforcement in the value-added tax." Working Paper 13-057. Harvard University, Business School, Cambridge, MA.
- Wan, J. 2010. "The incentive to declare taxes and tax revenue: The lottery tax receipt experiment in China." *Review of Development Economics* 14 (3): 611–24.
- Waseem, M. 2013. "Taxes, informality, and income shifting: Evidence from a recent Pakistani tax reform." IGC Working Paper. International Growth Centre, London, UK.
- World Bank. 2013. *Doing Business 2013: Smarter Regulations for Small and Medium-Size Enterprises*. Washington, DC: World Bank.

Cash Transfers and Child Labor

Jacobus de Hoop* and Furio C. Rosati*[†]

Cash transfer programs are widely used in settings where child labor is prevalent. Although many of these programs are explicitly implemented to improve children's welfare, in theory their impact on child labor is undetermined. This paper systematically reviews the empirical evidence on the impact of cash transfers, conditional and unconditional, on child labor. We find no evidence that cash transfer interventions increase child labor in practice. On the contrary, there is broad evidence that conditional and unconditional cash transfers lower both children's participation in child labor and their hours worked and that these transfers cushion the effect of economic shocks that may lead households to use child labor as a coping strategy. Boys experience particularly strong decreases in economic activities, whereas girls experience such decreases in household chores. Our findings underline the usefulness of cash transfers as a relatively safe policy instrument to improve child welfare but also point to knowledge gaps, for instance regarding the interplay between cash transfers and other interventions, that should be addressed in future evaluations to provide detailed policy advice. Cash Transfers, Child Labor, Impact Evaluation, Review. JEL codes: I28, I38, O20

The [International Labor Office \(2013a\)](#) estimates that over 120 million 5- to 14-year-old children were involved in labor in 2012, accounting for approximately 10 percent of the children in this age group. Child labor constitutes a violation of children's basic rights. It may have detrimental short- and long-run effects on children's lives, potentially lowering their school participation and learning in school and possibly affecting their mental and physical health. Beyond its direct effects on the children concerned, child labor may result in negative externalities. The economic literature assumes that parents determine the child labor supply by weighing its harmful effects against its potential benefits (primarily contributions to current household income). This literature shows that child labor supply is likely to be inefficiently high, even if parents are altruistic and externalities are absent, thus

The World Bank Research Observer

© The Author 2014. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
doi:10.1093/wbro/lku003 Advance Access publication April 1, 2014 29:202–234

providing a rationale for policy intervention (see, among others, [Cigno and Rosati 2005](#); [Edmonds 2007](#); and [Udry 2006](#)).

This paper aims to add to our understanding of the role of policy interventions by systematically reviewing the evidence on the impact of cash transfers, both unconditional and conditional, on child labor in developing countries.¹ Unconditional cash transfers provide households with an income transfer that can, for example, help reduce poverty, address household vulnerability, and encourage household investment in the human capital of their children. Conditional cash transfer programs also provide an income transfer, but on the condition that the members of the households receiving the transfer adhere to specific behavioral requirements. The behavioral conditions are typically in the area of health (e.g., health checkups and attendance at health-related seminars) and human capital accumulation (e.g., regular school attendance of children in the household).²

Cash transfers are a particularly relevant category of interventions. They are widely used in settings where child labor is prevalent and are often explicitly implemented to improve children's welfare. Nonetheless, in theory, their impact on child labor is undetermined. In fact, even increases in school participation do not necessarily translate to reductions in child labor because school participation and child labor are not mutually exclusive activities. The extensive evidence on the beneficial effects of cash transfers on school participation (see, e.g., [Baird et al. 2013](#); [Fiszbein and Schady 2009](#); [Rawlings and Rubio 2005](#); [Saavedra and Garcia 2012](#)) thus provides little guidance regarding their effects on child labor.

A key issue that must be emphasized from the outset is that child labor is not a concept that easily translates into statistical indicators. Child labor can affect children in different ways, and its consequences for children's welfare cannot be captured by a single indicator. Damage to child health, for instance, depends on the participation in - and the length of exposure to - hazardous activities and occupations. Similarly, the effects on the accumulation of human capital depend on participation in work, work schedule and hours worked, sector of employment, and occupation.³ Thus, detailed information is necessary to fully understand how policy interventions affect children's welfare.⁴ Such detailed information, however, is seldom available. Instead, most studies focus on participation in work, with a few studies discussing the impact on working hours. To complicate matters further, there is substantial variation in the way "work" is defined across different studies. Some studies focus on specific activities (such as work in agriculture), whereas others use a more general definition (such as work in economic activities or household chores).⁵

To ensure consistency and incorporate evidence from the largest possible number of studies, we mainly discuss the impact of cash transfer programs on children's participation in economic activities. To the extent possible, we complement this discussion with a separate analysis of the impact of cash transfer

programs on hours worked. Despite its limitations, the literature's focus on participation in work provides important insights. First, impact on participation serves as a useful first-order approximation to examine whether cash transfers affect the working behavior of children. Second, keeping children out of work is a sufficient condition to prevent detrimental effects on outcomes such as health and human capital accumulation. Third, participation in work is typically the key outcome in policy discussions and is an important outcome from a legal perspective (most countries in the world have adopted child labor laws that set a minimum age below which children are not allowed to work). Moreover, by focusing on participation in work, we are able to highlight some of the limitations of the research conducted to date and the main questions that remain to be answered in future evaluations.⁶

The remainder of this review is organized as follows. Section 1 provides the necessary background. It heuristically describes why the effects of cash transfers on child labor are theoretically undetermined, and it introduces the procedure that we used to identify the relevant studies for this review. Section 2 discusses the impact of two subsets of *unconditional* cash transfers: programs designed to support poor households' investments in children's human capital and old age pension schemes. Section 3 discusses the average impact of *conditional* cash transfer schemes on the intensive and extensive margin of child labor, the impact of conditional cash transfers on child labor compared to their impact on school participation, heterogeneity by poverty, age, and gender, spillover effects, long-run effects, determinants of program effects, protection from shocks, and variations on the basic conditional cash transfer scheme. Section 4 discusses and concludes.

Background

Theoretical Framework

The effects of cash transfers on child labor cannot be determined a priori, as we briefly discuss below. There are two main channels through which cash transfers may affect child labor: (i) by modifying the propensity to attend school and (ii) by changing the returns to child labor. Because a detailed theoretical discussion of these issues is beyond the scope of this paper, we present a heuristic explanation in what follows. This discussion draws on the theoretical framework presented in [De Hoop and Rosati \(2014\)](#).⁷

Consider a unitary household in which parents maximize utility over current consumption (which depends partly on children's work), children's education, and leisure. For simplicity, and because the cash transfer programs we consider are targeted at the poor, assume that households are credit constrained. As a result, they may under-invest in children's education and in productive assets. Assume also

that the adult labor supply is fixed, fertility is exogenous and fixed at one child, and school participation is dichotomous (i.e., the child either does not attend school or spends a fixed amount of time in school). The basic characteristics of an altruistic overlapping generation model that are essential for our analysis are captured by this very simple set of assumptions.

Within this framework, the effects of cash transfers, both conditional and unconditional, depend on whether the transfer program affects school participation.⁸ If the school participation of the child is *not* affected (e.g., in case of children who would be in school in the absence of a cash transfer program) and the household consumes the transfer (i.e., the transfer results only in an income effect), then child labor should decrease.⁹ However, if the household invests (part of) the transfer in productive assets, the returns to child work may increase, thereby counterbalancing the income effect and possibly resulting in increased child labor.¹⁰

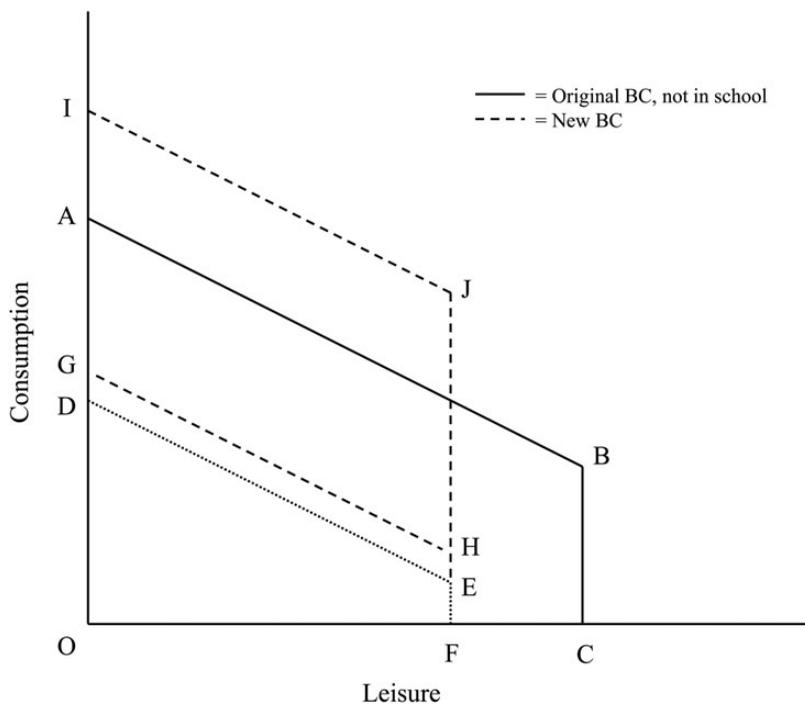
Figure 1 helps explain what happens when school participation is affected by the transfer. Household consumption is plotted on the vertical axis, and child leisure is plotted on the horizontal axis. The budget constraint for a household that does not receive a transfer and does not send its child to school is given by the solid line ABC. The downward slope of the budget constraint reflects the decrease in consumption as the child works less (i.e., consumes more leisure). Household consumption is positive even if the child does not work and spends all available time on leisure (point B) because the household also relies on income from other household members.

If the child begins to attend school following a transfer, the time available to the child for leisure and work is reduced (here, from C to F). Moreover, the household incurs the cost of education (e.g., school fees). If the transfer does not fully compensate the monetary cost of attending school, the budget constraint shifts downwards (for instance, to the dashed line GHF). Here, the distance DA represents the monetary costs of attending school, and DG represents the transfer. The change in child labor is undetermined because both feasible consumption and leisure are reduced, and the final outcome depends on the relative change in the (utility) value of the two. If the transfer exceeds the monetary cost of attending school, the budget constraint shifts upwards (for instance, to the dashed line IJF). Child labor should unambiguously decrease if the household does not invest the transfer in productive assets. However, in both cases, the effect on child labor also depends on whether the household decides to invest (part) of the transfer in productive assets, which may increase the returns to child labor.

Identifying Relevant Studies

To identify potentially relevant studies that evaluate the impact of cash transfer interventions on children's work outcomes, we began with a literature search

Figure 1. Changes in the Budget Constraint when the Transfer Affects School Participation



Source: Authors' own work.

Notes: Household consumption is plotted on the vertical axis, and child leisure is plotted on the horizontal axis. The solid line ABC represents the budget constraint for a household that does not receive a transfer and does not send its child to school. The dotted and dashed lines represent the budget constraints for households that do send their child to school and (i) receive no cash transfer (DEF), (ii) receive a cash transfer smaller than the cost of education (GHF), and (iii) receive a cash transfer that exceeds the cost of education (IJF). The segments AD, DG, and DI respectively represent the monetary cost of education, a transfer smaller than the cost of education and a transfer exceeding the cost of education. The segment FC represents the fixed amount of time a child spends in school.

covering Google Scholar's electronic bibliographical database, the World Bank Development Impact Evaluation Initiative (DIME) database, the Poverty Action Lab, the Social Science Research Network (SSRN), Network of Networks for Impact Evaluation (NONIE), and the International Initiative for Impact evaluation (3IE).¹¹ To determine which papers to include in our discussion, we split the identified papers into two subgroups: peer-reviewed papers and non-reviewed papers. We included all of the peer-reviewed papers in our discussion. We included non-reviewed studies if they applied a plausible strategy to address endogenous program placement and self-selection into the program. We did not limit our overview to randomized controlled trials (RCTs) and considered other methodologies, such as regression

discontinuity designs, natural experiments, and propensity score matching studies.¹²

Table 1 provides an overview of the studies included in our primary analysis: seven studies of unconditional cash transfer schemes and 23 studies of conditional cash transfer schemes. The majority of studies, 23 in total, focus on cash transfer programs implemented in Latin America and the Caribbean (column 1). Five studies have focused on Mexico's flagship conditional cash transfer scheme, which is called Oportunidades.¹³ Over half of the studies have appeared in a peer-reviewed journal (column 4). Fourteen studies are based on an RCT (column 5). The age range covered differs substantially across studies, an issue that should be considered when interpreting our results (column 6). In most studies, the main outcome variable on which we focus includes economic activities for pay or for the household, although in some cases, the impact estimate refers to a narrower (economic activities for pay) or a broader outcome variable (economic activities or household chores) (column 7).¹⁴

Unconditional Cash Transfers

Unconditional Cash Transfer Programs Targeted at Poor Households with Children

We discuss three unconditional cash transfer programs for which rigorous evidence is available: Ecuador's Bono de Desarrollo Humano, Malawi's Social Cash Transfer Scheme, and South Africa's Child Support Grant (the results are summarized in figure 2).¹⁵ The evaluations of these programs suggest that unconditional cash transfers that aim to encourage investment in human capital tend to lower participation in economic activities. IV estimates using randomly assigned eligibility status as an instrument indicate that Ecuador's Bono de Desarrollo Humano had a particularly strong effect. It lowered children's participation in economic activities by 17 percentage points for 6- to 17-year-old children (Schady and Araujo 2006) and by 25 percentage points for 11- to 16-year-old children (Edmonds and Schady 2012), and it resulted in substantial reductions in work for pay. It appears that South Africa's Child Grant did not affect the time allocation of 10-year-olds. However, propensity score dose-response estimates indicate that it did affect adolescents aged 15 to 17; the probability that they worked outside the home was 21 percent if they started receiving the grant at the age of 14 and 13 percent if they started receiving the grant at the infant or pre-school age (DSD, SASSA, and UNICEF 2012).¹⁶

Covarrubias, Davis, and Winters (2012) find that Malawi's Social Cash Transfer Scheme increased household investment in productive agricultural assets. Perhaps as a result, there is evidence of reductions in child labor, especially in domestic work, outside the household; however, "the time freed seems to be replaced with greater

Table 1. Programs and Studies Discussed in Main Text

	Country	Program	Studies	Published in peer-reviewed journal	Methodology	Main age-range	Activities included in main outcome variable
ID	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Unconditional Cash Transfers</i>							
1.	Brazil	Old age pensions	de Carvalho Filho (2012)	Yes	Triple difference	10–14	Economic activities for pay
2.	Malawi	Social Cash Transfer Scheme	Covarrubias, Davis, and Winters (2012)	Yes	PSM (see footnote)	0–17	Multiple underlying activities
3.	Malawi	Social Cash Transfer Scheme	Miller and Tsoka (2012)	Yes	Dif in dif (see footnote)	6–18	Economic activities for pay
4.	Nicaragua	Bono de Desarrollo Humano	Schady and Araujo (2006)	No	Household level RCT	6–17	Economic activities for pay or for household
5.	Nicaragua	Bono de Desarrollo Humano	Edmonds and Schady (2012)	Yes	Household level RCT	11–16	Economic activities for pay or for household
6.	South Africa	Child Support Grant	DSD, SASSA, and UNICEF (2012)	No	PSM Dose response analysis	10 and 15–17	Economic activities for pay or for household
7.	South Africa	Old age pensions	Edmonds (2006)	Yes	RDD	13–17	Economic activities for pay or for household
<i>Conditional Cash Transfers</i>							
8.	Brazil	Bolsa Escola	Ferro, Kassouf, and Levison (2010)	In edited volume	PSM	6–15	Economic activities for pay or for household
9.	Brazil	PETI	Yap, Sedlacek, and Orazem (2002)	No	Regression (see footnote)	7–14	No definition of work given
10.	Cambodia	CESSP Scholarship Program	Ferreira, Filmer, and Schady (2009)	No	RDD	7–18	Economic activities for pay
11.	Colombia	Familias en Accion	Attanasio et al. (2010)	Yes	Probit and regression (see footnote)	10–17	Economic activities for pay
12.	Colombia	Familias en Accion	Fitzsimons and Mesnard (Forthcoming)	Yes	Regression (see footnote)	10–17	Economic activities for pay or for household or chores

13.	Colombia	Subsidios Condicionados a la Asistencia Escolar	Barrera-Osorio et al (2011)	Yes	Child level RCT	grades 6–11	Work last week as a primary activity
14.	Honduras	PRAF-II	Galiani and McEwan (2013)	Yes	Cluster RCT	6–12	Economic activities for pay or for household
15.	Honduras	PRAF-II	Glewwe and Olinto (2004)	No	Cluster RCT	6–12	Economic activities for pay or for household
16.	Indonesia	Jaringan Pengaman Social	Sparrow (2007)	Yes	Natural experiment IV	10–18	Economic activities for pay or for household
17.	Jamaica	Path	Levy and Ohls (2007)	No	RDD	6–17	Economic activities for pay or for household or chores
18.	Mexico (rural)	Oportunidades	Buddelmeyer and Skoufias (2004)	No	Cluster RCT	6–12	Economic activities for pay or for household
19.	Mexico (rural)	Oportunidades	De Janvry et al. (2006)	Yes	Cluster RCT	8–18	Economic activities for pay or for household
20.	Mexico (rural)	Oportunidades	Skoufias and Parker (2001)	Yes	Cluster RCT	8–17	Economic activities for pay or for household
21.	Mexico (rural, long-run)	Oportunidades	Behrman, Parker, and Todd (2011)	Yes	Cluster RCT and PSM	15–16	Economic activities for pay or for household
22.	Mexico (urban)	Oportunidades	Behrman et al. (2012)	Yes	PSM	12–14	Economic activities for pay
23.	Nicaragua	Atención a crisis	Del Carpio and Macours (2010)	In edited volume	Cluster and household level RCT	6–15	Hours worked last week for pay or for household
24.	Nicaragua	Atención a crisis	Del Carpio and Loayza (2012)	No	Cluster and household level RCT	7–14	Hours worked last week for pay or for household or in chores
25.	Nicaragua	Red de Protección Social	Maluccio and Flores (2005)	No	Cluster RCT	7–13	Economic activities for pay or for household
26.	Nicaragua	Red de Protección Social	Dammert (2009)	Yes	Cluster RCT	7–13	Economic activities for pay or for household

Continued

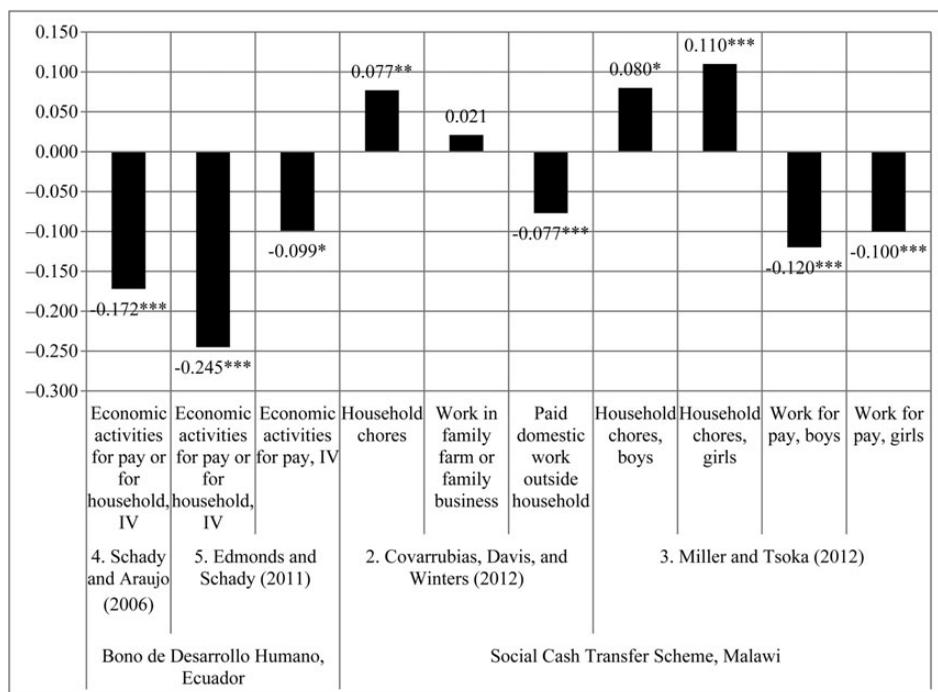
Table 1. *Continued*

	<i>Country</i>	<i>Program</i>	<i>Studies</i>	<i>Published in peer-reviewed journal</i>	<i>Methodology</i>	<i>Main age-range</i>	<i>Activities included in main outcome variable</i>
<i>ID</i>	<i>(1)</i>	<i>(2)</i>	<i>(3)</i>	<i>(4)</i>	<i>(5)</i>	<i>(6)</i>	<i>(7)</i>
27.	Nicaragua	Red de Protección Social	Gee (2010)	Yes	Cluster RCT	7–13	Hours worked last week (no further definition given)
28.	Pakistan	Female School Stipends	Alam, Baez, and Del Carpio (2011)	No	RDD	12–19	Economic activities for pay or for household
29.	Uruguay	PANES	Amarante, Ferrando, and Vigorito (2011)	No	RDD	6–17	Working (no further definition given)
30.	Uruguay	PANES	Borraz and González (2009)	Yes	PSM	6–15	Economic activities for pay or for household or chores

Source: Authors.

Note: PSM stands for propensity score matching, RCT stands for randomized controlled trial, and RDD stands for regression discontinuity design. The identification strategy of [Covarrubias, Davis, and Winters \(2012\)](#) and [Miller and Tsoka \(2012\)](#) is enhanced by a randomization procedure in which 8 groups of villages (containing 23 villages in total) were randomly divided into an equally sized treatment group and control group (4 groups of villages each). The age range for the study by [Covarrubias, Davis, and Winters \(2012\)](#) is not explicitly mentioned and we deduce the reported age range from the descriptive statistics. [Yap, Sedlacek, and Orazem \(2002\)](#) compare children in nine treatment municipalities to children in nine control municipalities that were either scheduled for incorporation in the program or had applied to be incorporated. The identification strategy of [Attanasio et al. \(2010\)](#) and [Fitzsimons and Mesnard \(Forthcoming\)](#) is enhanced by the fact that control households are drawn from purposefully chosen similar non-program municipalities.

Figure 2. Unconditional Cash Transfers Tend to Reduce Child Labor, Although the Results Are Not Uniform



Source: Authors' own work.

Note: Change in the probability of working as a result of the unconditional cash transfer programs displayed on the horizontal axis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

involvement in within-household tasks". Participation in household chores, for instance, increased significantly by 8 percentage points according to the authors' propensity score matching estimates. Although participation in work in the family farm did not change significantly, *hours* worked in the family farm or family business increased.¹⁷ Miller and Tsoka (2012) provide difference-in-difference estimates using the same data, confirming that participation in household chores increased significantly, whereas participation in work for pay decreased.

Old Age Pensions and the Role of Credit Constraints

In our theoretical discussion, we assumed that households are credit constrained. If they are not credit constrained, then household investment in education and productive assets (and, concomitantly, child labor) should not change upon the receipt

of a fully anticipated unconditional cash transfer. Two studies examine whether credit constraints matter for child labor by studying the impact of old-age pensions. Pensions are highly institutionalized and represent an *anticipated* and relatively certain future income stream for the household. Economic theory suggests that in the absence of credit constraints, households will follow the optimal smooth path of consumption and investment. Therefore, we should not observe any discontinuity in the behavior of households, and the fraction of working children in particular, just above and just below the pension age.

Edmonds (2006) uses data from South Africa's old-age pension scheme to test this proposition. South Africa's pensions are means tested and, as a result, primarily cover the comparatively deprived black population of South Africa. The benefits provided by the pension scheme are large: in 1999, they represented approximately 125 percent of median per capita income of South Africa's black population. School participation of 13- to 17-year-old children in the household increases substantially when an eligible elderly person reaches the pension age, an effect that is especially relevant for male pensioners. Children's participation in economic activities does not decline significantly when an elderly person in the household becomes eligible for the old-age pensions. However, there is evidence of a significant decline in daily hours worked.¹⁸ Boys experience larger reductions in time spent on economic activities, whereas girls appear to experience larger reductions in time spent on household chores.¹⁹

De Carvalho Filho (2012) confirms these findings for Brazil's social pension scheme. The author exploits a 1991 social security reform that increased the minimum benefit provided to pension beneficiaries in rural areas and reduced the minimum eligibility age, comparing households that became eligible to receive old-age benefits as a result of the reform to households that were nearly eligible after the reform.²⁰ Reduced form estimates indicate that the reform significantly increased girls' school participation and reduced their participation in economic activities for pay. IV estimates that disentangle this effect (using eligibility status as an instrument) show that increasing pension benefits by 100 Reais (approximately US\$50) increases girls' school enrollment by nearly 10 percentage points and reduces the probability that girls work for pay by 3.6 percentage points. It appears that effects of the program differ by the gender of the pension beneficiary, as only girls living in a household where a female received the pension experienced a significant reduction in economic activities for pay.

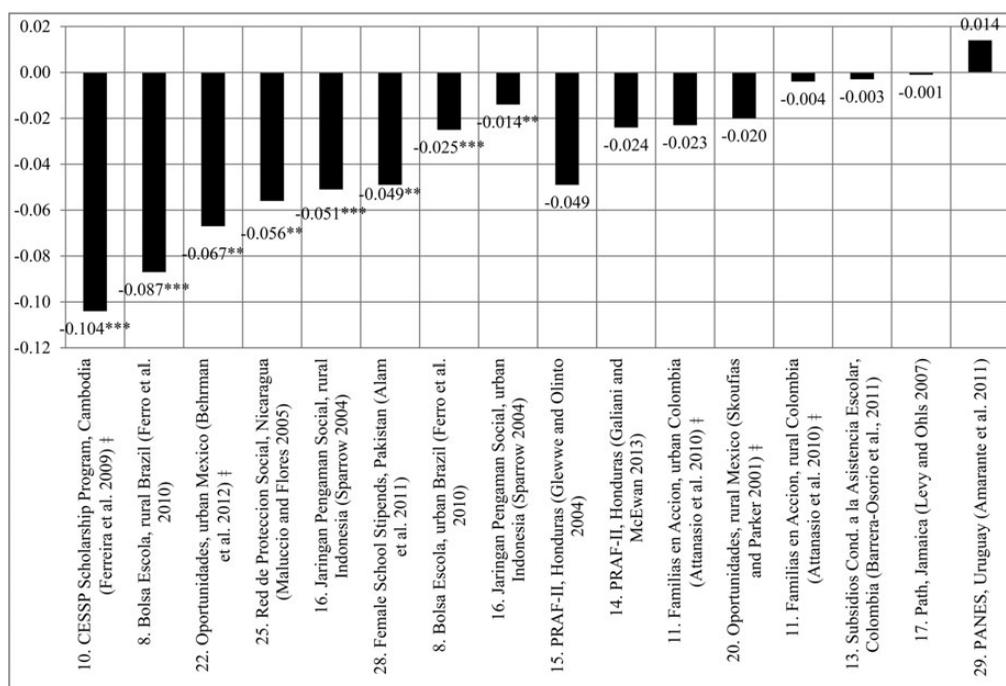
In summary, the evidence from studies of pension schemes points to the relevance for child labor of transfer schemes that do not have human capital accumulation as one of their objectives. Moreover, it lends support to the hypothesis that unconditional cash transfers affect child labor at least in part because they mitigate the effect of credit constraints.

Conditional Cash Transfers

Impact on Participation in Work and Hours Worked

Conditional cash transfers have been evaluated extensively, allowing us to discuss their impact in detail. Figure 3 synthesizes the evidence on the average impact of the conditional cash transfer programs on participation in child labor. For some programs, we only have disaggregated estimates by gender and/or age. Because we want to begin with a comparison of the average program effects before discussing heterogeneous effects, we impute the average program effect for those programs by taking the unweighted mean of the impact estimates given for different age and gender groups (these imputed estimates are marked with a † in the table).²¹ For other programs, we have estimates of average program effects from more than one study, in which case we show all of the estimates.²² To the greatest extent possible,

Figure 3. Conditional Cash Transfer Programs Tend to Reduce the Prevalence of Child Labor



Source: Authors' own work.

Note: Change in the probability of working as a result of the conditional cash transfer programs displayed on the horizontal axis. † indicates that the estimate is a weighted average of multiple age and gender groups. To minimize the text on the horizontal axis, we only display the first author of the study if the study has more than 2 authors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

we separately show impact estimates for urban and rural areas.²³ For brevity, we do not provide a discussion of the identification strategy employed in each individual paper, but Table 1 provides a basic description.

The results indicate that although there is considerable variation across programs, conditional cash transfer programs tend to reduce child labor. Impact estimates range from no statistically significant change in eight of the 16 studies to a reduction in child labor of 10 percentage points for Cambodia's CESSP scholarship program (Ferreira, Filmer, and Schady 2009). We do not observe statistically significant *increases* in child labor for any of the programs, an important finding given the theoretically ambiguous effect of conditional cash transfers on child labor discussed above. The results suggest that the effects of any household investments in productive assets and activities that draw children into work are offset by a stronger income and substitution effect that keeps children in school and out of work.²⁴

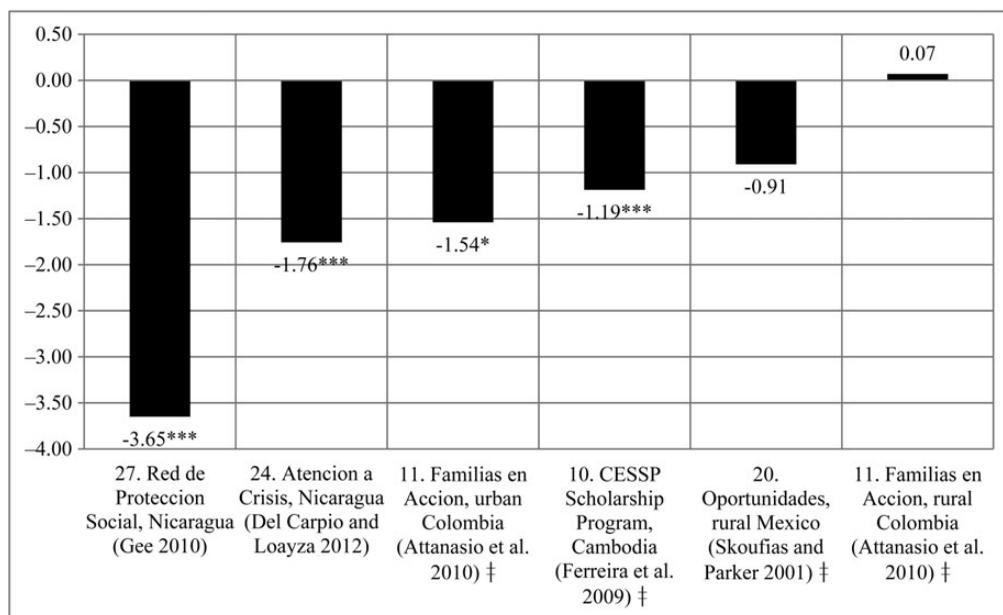
Figure 4 displays the impact of cash transfer programs on weekly hours worked.²⁵ Because each of the included studies sets hours worked equal to zero for children who do not work, these results effectively represent the combined effect of the included conditional cash transfer programs on the intensive and extensive margin of child labor.²⁶ Although fewer studies focus on hours worked than on participation in work, the results are qualitatively similar: conditional cash transfer programs tend to reduce hours worked, and none of the studies finds evidence of a significant increase in hours worked. In general, reductions in hours worked are modest (the average reduction is approximately one hour and a half a week).

Comparing Impacts on School Participation and Child Labor

The impact of cash transfer programs on child labor is correlated with their impact on school participation. Of the eight studies that find a significant reduction in child labor, six also find a significant increase in school participation (enrollment or attendance, depending on the outcome examined in the study), and of the eight studies that find no significant reduction in child labor, only three find a significant increase in school participation. However, the correlation between program impact on child labor and school participation is not perfect. When we regress the child labor impact estimates on the school participation impact estimates (results not displayed), we find evidence of substantial and statistically significant co-movement. Each percentage point increase in school participation is associated with a reduction in child labor of 0.31 percentage points, suggesting that child labor is a key part of households' human capital investment decisions. However, the coefficient on school participation impacts is significantly different from -1 , indicating that changes in school participation are not fully mirrored in changes in child work.

In fact, some of the differences in program impacts on school participation and participation in work are remarkable. As noted above, there are cases in

Figure 4. Conditional Cash Transfer Programs Tend to Reduce Weekly Hours Worked by Children



Source: Authors' own work.

Note: Change in hours worked as a result of the conditional cash transfer programs displayed on the horizontal axis. ‡ indicates that the estimate is a weighted average of multiple age and gender groups. To minimize the text on the horizontal axis, we only display the first author of the study if the study has more than 2 authors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

which conditional cash transfers had a significant effect on education but no effect on child labor (e.g., Path in Jamaica). Perhaps more unexpectedly, there are cases in which conditional cash transfers had a significant negative effect on child labor but no effect on school participation (e.g., female school stipends in Pakistan). It is, therefore, evident that complex adjustments in household behavior occur when cash transfers are received (changes in working hours and/or in leisure time, changes in the type of activities performed) and that the impact of a cash transfer on child labor is not necessarily the reciprocal of its impact on education.

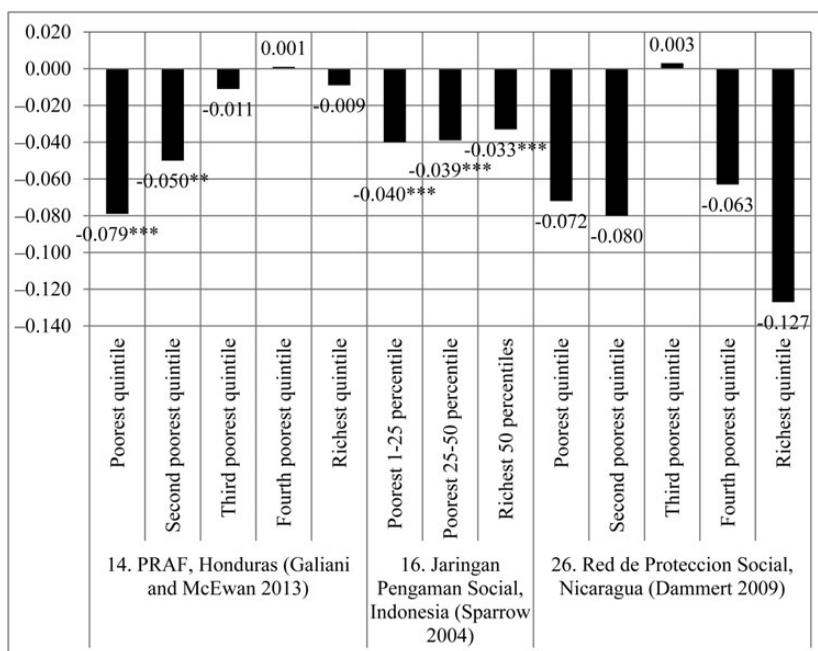
Heterogeneous Effects

Heterogeneity by Income. Above, we presented the results for the average impact of conditional cash transfer programs. We now examine possible differences associated

with some characteristics of the beneficiaries: income (in this subsection) and age and gender (in the next).²⁷ A priori, there are reasons to expect that the impact of conditional cash transfers will differ depending on the level of household income.²⁸ Most importantly, because poor households are more likely to be affected by credit constraints, transfers targeted to these households are more likely to affect investment in human capital and to reduce inefficiently high levels of child labor.

Two studies of the PRAF conditional cash transfer scheme in Honduras find that although the program did not significantly reduce child labor overall (figure 3), the effects of the program differed depending on household income and were statistically significant for the poorer households (figure 5). [Galiani and McEwan \(2013\)](#) divide their sample of municipalities into height-for-age quintiles (arguing that height for age provides an indication of poverty in the municipality) and find that the program reduced children’s participation in economic activities by eight percentage points (significant at the 5 percent level) in the poorest quintile. In the richest quintile, no statistically significant change in child labor could be observed.

Figure 5. The Impact of Conditional Cash Transfers on Child Labor Tends to Be Stronger in Poorer Households and Communities



Source: Authors’ own work.

Note: Change in the probability of working as a result of the conditional cash transfer programs displayed on the horizontal axis by poverty quintile or percentile. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Glewwe and Olinto (2004) interact their treatment indicator with the log of per capita expenditure and find a qualitatively similar result (not displayed in figure 5): as the per capita expenditure decreases by one percentage point, the impact of PRAF-II on children's participation in economic activities increases significantly by 0.45 percentage points.

Sparrow (2007) divides his sample into four per capita consumption quantiles to examine the heterogeneous effects of Indonesia's Jaringan Pengaman Sosial. He too finds that program impact on child labor increases with poverty. In the lowest consumption quantile, children's participation in economic activities decreased by 4 percentage points, whereas in the combined upper two quantiles, it decreased by only 3.3 percentage points (both are significant at the 1 percent level). Only Dammert (2009) does not identify a differential impact of cash transfers on child labor by income level. She investigates the effects of Nicaragua's Red de Protección Social by interacting a treatment dummy with marginality quintiles (based on a locality level marginality index).²⁹ The impact estimates by marginality quintile are volatile, and the estimates for the four richest quintiles are not significantly different from the impact coefficient for the poorest quintile.

In conclusion, the impact of conditional cash transfers on child labor generally appears to be larger for the poor. This finding is in accordance with the findings of Fiszbein and Schady (2009), who observe that "numerous studies have shown larger [conditional cash transfer] program effects among households that are poorer at baseline" on school participation. It is also in accordance with Edmonds and Schady (2012), who find that reductions in child labor as a result of Ecuador's Bono de Desarrollo Humano *unconditional* cash transfer scheme are concentrated in the poorest households. These results lend further support to the hypothesis that the mitigation of credit constraints, which are more likely to be binding for poor households, is a key channel through which cash transfers reduce child labor.

Heterogeneity by Age and Gender. It is not obvious a priori whether we should expect stronger changes in work participation for older or younger children. On the one hand, children's participation in work, and hence the margin for improvement in child labor outcomes, increases with age. On the other hand, as children become older, their returns to work are likely to increase, making work a more attractive alternative to compliance with a conditional cash transfer schooling requirement. To examine this issue, we use data from all studies that examine the effect of a conditional cash transfer scheme on child labor for two or more age groups. Using a regression, we test whether the impact on the younger and older age group differs (Column (1) of Table 2).³⁰ We find no evidence of a significant correlation, suggesting that the effect of conditional cash transfer schemes on child labor is not clearly heterogeneous by age.

Table 2. Heterogeneity of Conditional Cash Transfer Child Labour Impacts

	(1)	(2)
Dummy for older age group	0.003 (0.015)	
Dummy for boys		-0.033 (0.017)*
Number of observations	24	32

Source: Authors.

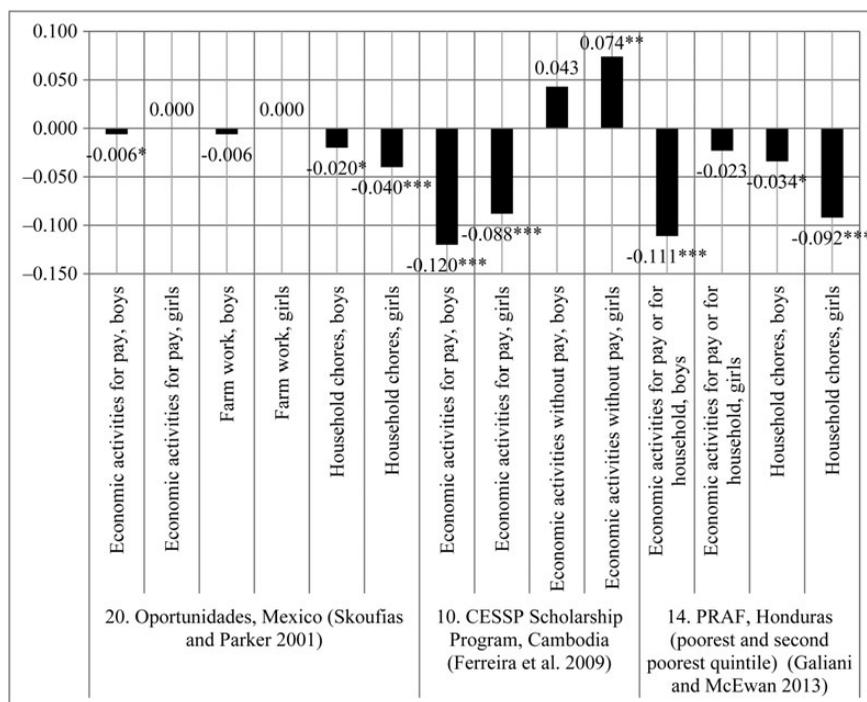
Note: Column 1 shows the results of regressing changes in child labor for 2 age groups per study (within the 7–14 age range, if possible) on a dummy taking the value 1 for the older age group (and a constant). Column 2 shows the results of regressing changes in child labor for boys and girls on a dummy taking the value 1 for boys (and a constant). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In Column (2) of Table 2, we test whether the effect of conditional cash transfers is heterogeneous by gender. To do so, we use the results from all of the studies that show impact results separately for boys and girls. We find that conditional cash transfer schemes result in a 3.3 percentage point stronger reduction in child labor for males than females. In fact, a substantial number of studies finds a significant impact on child labor among boys and no significant impact on child labor among girls. [Borraz and González \(2009\)](#), who examine the impact of the PANES conditional cash transfer program in Montevideo, Uruguay, are the only authors to find a significant decrease in child labor for girls, but no significant decrease for boys.³¹

To better understand the differential impact of conditional cash transfer interventions on the work of boys and girls, figure 6 provides results from three studies that disaggregate the overall impact of conditional cash transfers across different work activities by gender. Boys primarily experience reductions in economic activities for pay. Oportunidades, for example, resulted in a significant reduction in economic activities for pay conducted by boys, but not for girls ([Skoufias and Parker 2001](#)). Similarly, the CESSP scholarship program in Cambodia ([Ferreira, Filmer, and Schady 2009](#)) and the PRAF program in Honduras ([Galiani and McEwan 2013](#)) appear to have had a stronger impact on work for pay and work outside the home, respectively, for boys than for girls.³² Girls, in contrast, appear to experience larger reductions than boys in household chores (Oportunidades in Mexico),³³ economic activities without pay (the CESSP program in Cambodia), and work at home (PRAF in Honduras).³⁴

Similar results are obtained by [Del Carpio and Macours \(2010\)](#), who focus on Atención a Crisis, a one-year randomized pilot building on the Red de Protección Social cash transfer scheme in Nicaragua. The authors test for differences in reductions in hours worked by boys and girls in different activities during the week before the interview.³⁵ Fixed effects estimates indicate that the reduction in participation in economic activities of boys as a result of the basic conditional cash transfer program exceeded that of girls by more than one hour a week.³⁶ For household

Figure 6. Boys Experience Stronger Reductions in Economic Activities, Girls in Household Activities



Source: Authors' own work.

Note: Change in the probability of involvement in economic and household activities as a result of the conditional cash transfer programs displayed on the horizontal axis by gender. To minimize the text on the horizontal axis, we only display the first author of the study if the study has more than 2 authors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

chores (cooking, cleaning, washing, and caring for younger siblings), there was no significant difference between boys and girls.³⁷

It appears that a focus on economic activities does not reveal the full impact of conditional cash transfers on girls. Girls are more likely to participate in household chores than in economic activities; consequently, the impact of cash transfers on work performed by girls is likely to be underestimated if we focus only on participation in economic activities. More elaborate survey modules on children's time use are necessary to fully understand the impact of cash transfers, particularly on girls' activities.

Spillover Effects and General Equilibrium Effects

Conditional cash transfer programs may affect children who are not direct beneficiaries. At the household level, for example, income effects may alter the probability

that the siblings of the beneficiary work. Conditional cash transfer programs may also result in spillover effects in the local labor market. For example, pulling a substantial number of children out of work might affect conditions in the local labor market and increase the marginal returns to child labor.³⁸

Two evaluations suggest that such spillover effects are not highly relevant. [Ferreira, Filmer, and Schady \(2009\)](#) and [Galiani and McEwan \(2013\)](#) compare the siblings of eligible children in their treatment group to the siblings of eligible children in their control group for the CESSP program in Cambodia and PRAF in Honduras, respectively. Neither study finds compelling evidence that the siblings of eligible children in the treatment group altered their participation in work in comparison with the siblings of eligible children in the control group.³⁹ However, a different result is found by [Barrera-Osorio et al. \(2008, working paper version\)](#), who examine the spillover effects of Colombia's *Subsidios Condicionados a la Asistencia Escolar* conditional cash transfer scheme on the intensive margin of child labor. Within the sample of households that registered two children for the program, there was evidence that children, especially girls, spent more hours in work if they were not assigned to the program while their sibling was assigned to the program.

[Buddelmeyer and Skoufias \(2004\)](#) investigate whether village-level spillover effects can be observed for the *Oportunidades* conditional cash transfer scheme in Mexico. They exploit the fact that within randomly selected intervention villages, only poor households were eligible to participate in the program. Non-poor households in intervention villages were not eligible and can thus be compared to non-poor households in control villages to identify spillover effects. The authors find no significant village-level spillover effects on children's participation in economic activities.

Long-run Effects

The impact evaluations discussed thus far examine the impact of conditional cash transfers at one particular point in time (mainly shortly after the program began). However, the impact of a program may vary significantly over its lifetime ([King and Behrman 2009](#)). For instance, if program operation improves as providers become more experienced, the impact of the program may be amplified over time. If the impact of the program depends on the duration of exposure to the program, impact estimates based on data collected relatively soon after participants enter the program may differ substantially from estimates based on data collected at a later stage.

[Behrman, Parker, and Todd \(2011\)](#) examine the impact of *Oportunidades* in the longer run and discuss whether these estimates differ from the short-run results. Propensity score estimates suggest that the probability that boys who were 14 to 16 years old in 2003 (5.5 years after the program was first implemented) worked

was 14 percentage points lower in Oportunidades communities than in communities that had never received benefits of the cash transfer scheme.⁴⁰ There is no evidence that work participation changed for girls in this age group (who were less likely to work in the first place). The strong reduction in work by boys in the long run compared to the modest impact in the short run (Skoufias and Parker (2001) found that boys aged 14–15 reduced work participation by approximately 4 percentage points) suggests that the beneficial impact of the Mexican conditional cash transfer program is compounded over time. It is possible that reduced probabilities of dropping out of school in individual grades (an issue alluded to in Schultz, 2004) begin to add up. The latter interpretation appears to be confirmed by Behrman, Parker, and Todd (2011), who also register strong improvements in school participation for girls up to 18 and boys up to 21 years old.

Differences in Program Impact

We have observed that the estimates of the impact of conditional cash transfers on child labor vary substantially. Based on the available information and without developing a new in-depth quantitative and qualitative analysis for each program, it is impossible to fully identify the reasons behind this variation. However, we can obtain some basic insight into the relationship between program characteristics and program impact on child labor. In this section, we examine the role of schooling conditionality and the size of the transfer.

Effect of the Conditionality. A key question is, of course, whether the impact of conditional cash transfers on child labor exceeds that of unconditional cash transfers. This question is not easy to answer by comparing the effects of the unconditional and conditional cash transfers presented thus far. The decision to attach a condition might be endogenous depending on the expected impact of the program in the target population. Moreover, the studies included in this review do not always discuss the exact conditions of the programs, how these conditions are communicated to beneficiaries, and to what extent conditions are enforced. Hence, it may be the case that programs that are nominally conditional are unconditional in practice.

However, a few recent studies allow us to shed some light on the effects of schooling conditions. Schady and Araujo (2006) and Edmonds and Schady (2012) exploit a glitch in the rollout of the Bono de Desarrollo Humano program, which resulted in some beneficiary households incorrectly believing that the cash transfers were provided conditional on school attendance. Although the conclusions of the two papers are not entirely uniform, it appears that the effect of the program on child labor was similar in the households that believed that the program was conditional on school participation and in the households that did not.

At the time of writing, preliminary evidence was also available from an experiment in which households in rural Morocco were randomly selected to receive conditional and unconditional cash transfers (Benhassine et al. 2012). Although we caution that these results are not conclusive, they too suggest that changes in the time spent working in the household business are similar (and insignificant) for conditional and unconditional cash transfers.⁴¹ The similar impact of conditional and unconditional cash transfers on child work is in contrast with more abundant evidence suggesting that conditional cash transfers have a stronger impact on school participation than do unconditional cash transfers (for a review, see Baird et al. 2013). It is evident that more information is necessary to understand the effect of the schooling condition on child labor.

Amounts Transferred. There is some experimental evidence on the impact of the size of the transfer on school participation (e.g., Baird, McIntosh, and Özler 2011). No such evidence is available for child labor. To examine the association between amounts transferred and changes in child labor, we regressed the average impact of conditional cash transfer programs on the amount of money transferred as a percentage of average household income (results not displayed).⁴² We found no evidence that larger transfer amounts are associated with a larger reduction in child labor. The lack of a negative relationship between amounts transferred and changes in child labor is also apparent from the individual studies.⁴³ The CESSP scholarship program in Cambodia, for instance, resulted in the second strongest decrease in child labor of all evaluated conditional cash transfer programs even though it provided only very modest transfers (equal to 2 to 3 percent of the total expenditures of the average recipient household). Uruguay's PANES, in contrast, provided income transfers equal to approximately 50 percent of average self-reported pre-program household income and yet does not appear to have lowered child labor. This finding is in line with the evidence synthesized by Kremer and Holla (2009), suggesting that relatively small costs (for example, the costs of school uniforms) are sufficient to keep children out of school and that relatively small subsidies can generate "sizeable movements" in the take-up of health and education interventions. Together with the finding that cash transfers have a stronger impact on the poor, this finding lends further support to the hypothesis that the mitigation of credit constraints is a key channel through which cash transfers affect child labor. Once sufficient cash has been transferred and credit constraints are no longer binding, transferring additional cash to the household may have no effect on child labor.

Protection from Shocks

There is substantial evidence that households in developing countries use child labor to cope with income shocks. Beegle, Dehejia, and Gatti (2006) use data from

a household panel survey in Tanzania to show that households increase child labor in response to transitory income shocks. Similarly, [Duryea, Lam, and Levison \(2007\)](#) find that unemployment shocks experienced by male household heads in Brazil significantly increase the probability that a child will enter the labor force. [Guarcello, Mealli, and Rosati \(2010\)](#) show that in Guatemala, exposure to negative shocks strongly influences household decisions and pushes children into work. An important question is whether social protection programs, such as conditional cash transfers, can serve as a safety net preventing income shocks from causing children to drop out of school and enter work.

[De Janvry et al. \(2006\)](#) test whether the conditional cash transfers provided by the Oportunidades program in Mexico protect children from household level shocks, including illness of the head of household, loss of employment by the head of household, and natural disasters such as drought and harvest failure. The authors find evidence of state dependence in schooling: children may leave school as the result of a shock (particularly illness of the household head and locality-level natural disasters) and there is a substantial risk that they will not return to school at a later stage. Oportunidades strongly reduced the risk that pupils would leave school as a result of such shocks. For instance, the risk that a child will drop out of school as a result of illness of the household head or a locality-level disaster (1.7 and 3.2 percentage points in the absence of Oportunidades, respectively) was virtually reduced to zero by the conditional cash transfer scheme. No such protective effect could be observed for child labor, however. The same two shocks, illness of the household head and locality-level natural disasters, increase children's participation in economic activities by 2.2 and 4.7 percentage points, respectively, in the absence of the program. This increase was not significantly different in Oportunidades villages, suggesting that the protective effect of the conditional transfer program was "not sufficient to reduce the use of child work as a crucial element of [household] risk-coping strategies."

Following a similar approach, [Fitzsimons and Mesnard \(Forthcoming\)](#) examine whether Colombia's Familias en Acción conditional cash transfer scheme protected children from the detrimental effects of the permanent departure of their father from the household due to death or divorce. The authors show that the departure of the father is a quasi-random event accompanied by a substantial reduction in household income. Children who experience the departure of their father are approximately five percentage points less likely to attend school and approximately three percentage points more likely to work. In contrast with [De Janvry et al. \(2006\)](#), [Fitzsimons and Mesnard \(Forthcoming\)](#) find that conditional cash transfers strongly mitigate both the detrimental effect of the departure of the father on school enrollment *and* the detrimental effect on participation in work.⁴⁴

Variations on the Basic Conditional Cash Transfer Scheme

Some studies examine whether changes in the basic setup of the conditional cash transfer program affect their impact. [Del Carpio and Loayza \(2012\)](#) rely on a randomized experiment in Nicaragua to compare the impact of simple conditional cash transfers to that of conditional cash transfers combined with a “grant for productive investments” to start a new income-generating non-agricultural activity. The grant was provided to households conditional on the development of a business plan and was accompanied by technical assistance and training in basic business skills. The two variants of the conditional cash transfer scheme do not appear to have had the same impact on child participation in economic activities or household chores: the conditional cash transfers in combination with the productive investment grant reduced child labor by 0.94 hours a week, whereas the basic conditional cash transfer reduced child labor by 1.76 hours a week (the difference between the point estimates is statistically significant at the 5 percent level).

[Barrera-Osorio et al. \(2011\)](#) rely on a randomized experiment in Colombia to compare the impact of simple conditional cash transfers to that of two variations on the traditional conditional cash transfer scheme. The first variation provided a regular cash transfer that was equal to two-thirds of the basic conditional cash transfer. The remaining accumulated one-third was transferred shortly before the start of the following school year, thus potentially helping households cope with savings constraints that keep students from proceeding to the next grade. The second variation also lowered the regular cash transfer payment, but instead of transferring the accumulated funds prior to the new school year it provided a large bonus for graduating from secondary school. This bonus was provided earlier to students who enrolled for tertiary education (upon graduation) than students who did not. The effect of the first variation on child work appears to be similar to that of the simple cash transfer scheme. The second variation, however, resulted in a more pronounced reduction in child work, in particular for graduating pupils (who were also markedly more likely to enroll for tertiary education).

[Glewe and Olinto \(2004\)](#) and [Galiani and McEwan \(2013\)](#) compare the impact of receiving Honduras’ PRAF-II conditional cash transfers in isolation (discussed in more detail above) to the impact of receiving the conditional cash transfer in combination with direct investments in the communities’ health and education facilities. As shown in figure 3, neither of these studies found an overall significant reduction in child labor as a result of the pure conditional cash transfer scheme (although both found effects for the poorest households). Glewe and Olinto (2004) also find no overall effect of the cash transfers combined with supply-side interventions. [Galiani and McEwan \(2013\)](#), however, find that the conditional cash transfers in combination with investments in health and education facilities did result in a statistically significant decrease in economic activities outside the household and in household chores.⁴⁵

Finally, [Yap, Sedlacek, and Orazem \(2002\)](#) investigate the impact of Brazil's Programa de Erradicacao de Trabalho Infantil (PETI). This program was targeted at poor households in rural areas of the country where the prevalence rates of child labor are high. Similar to Brazil's Bolsa Escola program, PETI provided a cash transfer conditional on school participation. However, PETI was explicitly designed to reduce child labor. For this purpose, it required households to enroll their child in an after-school education program. The content of the after-school education program differed by community and could include academic and physical education components. After-school education essentially doubled the length of the school day for participating children. The authors find that PETI reduced child labor by 5 to 25 percentage points in different regions.⁴⁶ Taken at face value, these results suggest that the program resulted in strong reductions in child labor compared to, for example, Bolsa Escola, which provides pure conditional cash transfers and reduced participation in economic activities by 8.7 percentage points in rural areas ([Ferro, Kassouf, and Levison 2010](#)).

Taken together, the studies discussed in this section suggest that the impact of conditional cash transfers depends partly on their integration with other interventions. Interventions that aim to improve income-generating activities may reduce the impact of conditional cash transfers on child labor, possibly by generating increased demand for children's time within the household. In the case of Nicaragua, for example, a plausible explanation for the weaker program effect when the basic conditional cash transfer is combined with a grant for productive investments could be that children are employed in the newly developed household business.

Combining conditional cash transfers with supply-side interventions such as the provision of health and education facilities and after-school education may increase impacts on child work. This finding seems intuitive because these supply-side interventions reduce the incentives or the time available to the child for work. However, further research is needed to better understand how cash transfer and supply-side interventions interact. Is their combined effect simply equal to the sum of the effect of the individual interventions? Or is there a synergy, such that the effect of the cash transfers and supply-side interventions is mutually reinforcing? This is a key topic for future research, not only for child labor but also for outcomes such as school participation.

Conclusion

Cash transfer schemes are not often designed and implemented with the aim of reducing child labor, although social protection is recognized as one of the main instruments to address child labor ([International Labor Office 2013b](#); [Understanding Children's Work 2010](#)). However, as this review shows, cash transfers have a strong potential to address child labor. We have not identified a single program that

increased child labor (although indications of negative spillover effects warrant further research). On the contrary, there is broad evidence that cash transfers, conditional and unconditional, lower both the extensive and intensive margin of child labor. Moreover, cash transfers appear to cushion the effect of economic shocks that may lead households to use child labor as a coping strategy.

The effects of the cash transfer programs on child labor are heterogeneous. There are differences by the child's gender. Boys tend to experience a larger reduction in participation in economic activities, whereas girls experience relatively larger reductions in involvement in household chores. Moreover, reductions in child labor are particularly pronounced when beneficiaries are poor, signaling that the mitigation of credit constraints, which force households to use child labor as a consumption smoothing mechanism, are a key channel through which cash transfers lower child labor. This interpretation is confirmed by studies of old-age pension schemes that allow for a more direct test of the role of credit constraints.

Preliminary evidence suggests that the impact of conditional cash transfers depends partly on their integration with other interventions. Combining conditional cash transfers with supply-side interventions such as the provision of health and education facilities and/or after-school education may increase impacts on child labor. In contrast, interventions that aim to improve income-generating activities may reduce the impact of conditional cash transfers on child labor, possibly by generating increased demand for children's time within the household. These results indicate that to reduce child labor, careful attention should be paid to the integration of conditional cash transfers with other interventions.

Our findings are in accordance with those of [Edmonds \(2007\)](#) and [Fiszbein and Schady \(2009\)](#), who also find that cash transfer programs are generally a promising tool for reducing child labor, drawing on a subset of the studies discussed in this paper.⁴⁷ In conclusion, the use of cash transfers as an anti-poverty strategy seems to be effective to reduce child labor. The same does not necessarily hold for all anti-poverty and income-generating interventions. The encouragement of entrepreneurship through microfinance programs, for instance, *can* increase child labor (e.g., [Augsburg et al. 2012](#); [Nelson 2011](#)). This finding further underlines that cash transfers are a useful policy instrument to improve child welfare and suggests that they are unlikely to have detrimental effects on child labor, even when they are not implemented or designed to address it.

However, important knowledge gaps must be addressed to provide more detailed policy advice. The main gaps are because, as mentioned, cash transfers are seldom implemented with a reduction in child labor as one of their main objectives and are therefore typically not assessed in depth against this outcome. As a result, we know relatively little about the program characteristics that determine cash transfer programs' effects on child labor, and we do not clearly understand why some of these programs have no effect on child labor. The role of design elements that *have* been

tested appears to be limited. There is little evidence that schooling conditions affect program impact on child labor. The latter finding is surprising in light of recent research indicating that schooling conditions matter for school participation, and this finding warrants further research. The size of the transfer relative to household income also appears to have little influence on reductions in child labor. Some conditional cash transfer projects that transfer substantial sums of money have no effect on child labor, whereas other programs that provide only a small subsidy result in strong changes.

Beyond examining the program characteristics that determine the impact of cash transfers on child labor, there are more questions open for future research. A key issue, which we outlined in the introduction, is the measurement of child labor. Most impact evaluations focus on economic activities. This approach potentially results in the underreporting of program impact on activities performed by girls because they are more likely to be involved in household chores. Additionally, as a result of the focus on participation in economic activities (or in one of its subcomponents), we have little evidence on the extent to which the interventions prevent and reduce (i) the worst forms of child labor, including hazardous work, and (ii) long working hours that keep children from learning in school.

More systematic evidence on extensions of basic cash transfer schemes and the interplay between cash transfer schemes and other social protection and supply-side interventions would also be important. In practice, cash transfer interventions are rarely implemented in isolation, and interaction effects may well determine their impact. This review provides preliminary evidence suggesting that the impact of cash transfer schemes and education interventions may be mutually reinforcing, whereas combining cash transfer schemes with services that aim to foster income-generating activities may have a detrimental effect on child labor. However, much work remains to be done to understand the extent to which synergy effects drive the combined effect of cash transfer and supply-side interventions.

Finally, we know very little about the relative cost effectiveness of cash transfers in reducing child labor and how their cost effectiveness compares to other interventions. This issue has remained largely unexplored in most impact evaluations focusing on child labor. Few of the studies included in our review are explicit on the cost of implementing a cash transfer program (other than the transfer amounts). Evaluations of the impact of other categories of interventions on child labor are equally unlikely to discuss the cost of implementing the program per beneficiary. More explicit cost effectiveness analyses will be crucial in guiding governments in the elimination of child labor.

Notes

* Understanding Children's Work.

*† Understanding Children's Work, University of Rome Tor Vergata, IZA Funding for this paper was provided by the United States Department of Labor. This paper does not necessarily reflect the views or policies of the United States Department of Labor, nor does the mention of trade names, commercial products, or organizations imply endorsement by the United States Government. The paper benefited greatly from the valuable comments and suggestions of Marco Manacorda, the editor, and the anonymous referees. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of Understanding Children's Work or its partner organizations, the International Labor Organization, UNICEF, and the World Bank. Corresponding author: Jacobus de Hoop, email address: jdehoop@ucw-project.org; phone: +39-06-43412008; street address: International Labor Office for Italy and San Marino, Via Panisperna 28, 00184 Rome, Italy.

1. Previous work discussing evidence on the relationship between cash transfers and child labor, building on the earlier papers included also in this review, includes [Edmonds \(2007\)](#) and [Fiszbein and Schady \(2009\)](#). [Edmonds \(2007\)](#) also discusses the effect of a range of other interventions on child labor. [Fiszbein and Schady \(2009\)](#) also examine the effects of conditional cash transfers on various of other outcomes (including adult labor).

2. Over the past decade, developing country governments have begun to adopt conditional cash transfers as social protection instruments at a rapid pace. [Fiszbein and Schady \(2009\)](#) present a comprehensive review of the recent proliferation of conditional cash transfer schemes in developing countries.

3. The three principal international conventions on child labor recognize this complexity and set the legal boundaries that define children's work that is targeted for elimination. The ILO Convention No. 138 determines the minimum age below which children should not work and the minimum ages for light and hazardous work. The ILO Convention No. 182 specifies the worst forms of child labor that are prohibited for all children under the age of 18. The United Nations Convention on the Rights of the Child aims to protect children from economic exploitation and from performing any work that is likely to be hazardous or to interfere with the child's education or to be harmful to the child's health or physical, mental, spiritual, moral, or social development. The international legal standards contain a number of flexibility clauses left to the discretion of the competent national authority in consultation (where relevant) with worker and employer organizations (e.g., minimum ages, scope of application). There is no single legal definition of child labor across countries, and concomitantly, there is no single standard statistical measure of child labor consistent with national legislation across countries. (Text in this footnote is adapted from standard Understanding Children's Work description.)

4. Moreover, we must keep in mind that schooling, work, and other activities, including leisure and sleep, are jointly determined.

5. There is also variation in the reference period. Some studies consider work in the seven days prior to the household survey, whereas some studies examine work in the past 12 months. Finally, some studies present results for a few separate categories of activities instead of focusing on a comprehensive indicator for participation in work.

6. Initially, most evaluations of the impact of cash transfer interventions on human capital accumulation focused on self-reported school enrollment or attendance. A new generation of studies focuses on more elaborate outcomes, such as performance on standardized tests. A similar development has not taken place for child labor, and we hope that this review can serve as a starting point.

7. For a formal model, see for instance [Ravallion and Wodon \(2000\)](#).

8. For a theoretical discussion of the effects of cash transfers on school participation see, inter alia, [Fiszbein and Schady \(2009\)](#).

9. If leisure is a normal good.
10. For a discussion of investment of cash transfers in productive assets, see [Fiszbein and Schady \(2009\)](#).
11. The full literature search was conducted in early 2011 by Understanding Children's Work (UCW) to build a comprehensive database of child labor impact evaluations. The database is updated regularly and can be found at <http://www.ucw-project.org/impact-evaluation/inventory-impact-evaluations.aspx>. We considered works identified to the end of 2012. We considered only papers written in English and did not consider dissertation chapters that did not appear as separate (working) papers. Of course, relevant impact evaluations continue to appear after this date. Authors of relevant papers not included in this database are invited to share their papers.
12. We exclude simulation studies, such as [Bourguignon, Ferreira, and Leite \(2003\)](#), studies that capture the effect of multiple programs at once, such as [Cardoso and Portela Souza \(2003\)](#), and studies of in-kind transfers, such as [Ravallion and Wodon \(2000\)](#) and [Kazianga, de Walque, and Alderman \(2009\)](#).
13. We do not discuss [Rubio-Codina \(2010\)](#) and [Schultz \(2004\)](#) in the main text because these studies rely on the same data and find results comparable to those of [Skoufias and Parker \(2001\)](#).
14. A table with the exact definition of the outcome variable for each individual study is available on request.
15. Figure 2 does not include the dose-response estimates for South Africa's Child Support Grant ([DSD, SASSA, and UNICEF 2012](#)) discussed in the text, which differ in nature from the "binary" treatment results discussed in the remainder of this paper.
16. This result should be interpreted with some care because the authors do not show whether differences between individuals treated from an early age and those treated from the age of 14 are significant.
17. The program did not affect children's participation in care for other children in the household, care for adults in the household, or work outside the household for income (other than domestic work). The estimates are not always robust for different specifications.
18. That is, -0.6 hours for children living in a household with a male pensioner and -0.5 hours for children living in a household with a female pensioner.
19. [Edmonds \(2006\)](#) discusses several alternative hypotheses that could explain the observed changes in child labor and school participation in response to the realization of anticipated income but concludes that credit constraints are the most plausible explanation.
20. To correct for age-specific trends not related to the reform, the author compares difference-in-differences estimates for rural households to difference-in-differences estimates for urban households (which were not affected by the reform) in a triple-difference framework.
21. To calculate the standard errors associated with these estimates, we assume that the covariance between the individual estimates is zero. Note that the overall impact of the program may not be statistically significant even if some of the underlying estimates for age and gender subgroups are statistically significant.
22. If we have an estimate of average program effects from one study and estimates of disaggregated program effects from another study, we show only the estimate of the average program effect.
23. Although it seems reasonable to take the unweighted means of impact estimates across age and gender groups, the same procedure cannot be applied across urban and rural areas that may differ drastically in terms of population size. When we have estimates for both rural and urban areas for the same program this is indicated after the program name.
24. We examined whether the average the impact estimates of randomized controlled trials are different from the impact estimates of quasi-experimental studies and find that this is not the case. This finding is interesting in light of the ongoing debates regarding the validity of randomized evaluations vis-à-vis quasi-experimental approaches ([Deaton 2010](#); [Duflo and Kremer 2005](#); and [Ravallion 2009](#)).
25. To ensure comparability, we multiply the hours worked by seven for studies that consider daily working hours.

26. To correct for the censoring of the outcome variable, [Attanasio et al. \(2010\)](#), [Dammert \(2009\)](#), [Del Carpio and Loayza \(2012\)](#), and [Ferreira, Filmer, and Schady \(2009\)](#) use Tobit regressions. [Skoufias and Parker \(2001\)](#) use a Heckit procedure.

27. [Bando, Lopez-Calva, and Patrinos \(2005\)](#) examine whether the impact of Mexico's Oportunidades was stronger for indigenous households.

28. See [Fiszbein and Schady \(2009\)](#) for a further discussion.

29. In her paper, Dammert shows how impact estimates for the richest four quintiles differ from the impact estimate for the poorest quintile. To ensure that the results are comparable with those of the other studies, we imputed the impact of the program for each of the five quintiles. We followed a procedure similar to that discussed in footnote 21 to calculate standard errors for these estimates. [Dammert \(2009\)](#) also examines heterogeneity along household per capita expenditure quintiles. These results are not qualitatively different from the marginality index results; therefore, we decided not to present them in Figure 5.

30. For each study, we focus on two age groups, within the 7–14 age range, if possible.

31. The effects of South Africa's (unconditional) Child Grant also appear to be stronger for adolescent girls than for boys ([DSD, SASSA, and UNICEF 2012](#)).

32. For [Galiani and McEwan \(2013\)](#), this result refers to the two poorest quintiles.

33. Making purchases for the family, making clothes for family members, taking a family member to school, work, the health center, or the hospital, cleaning the house, washing and ironing clothes, cooking, fetching water or firewood, disposing of trash, and caring for small children, elderly family members, or sick individuals.

34. Moreover, [Dubois and Rubio-Codina \(2012\)](#) find that Oportunidades lowered teenage girls' participation in care for younger siblings.

35. Not displayed in Figure 6 because the study focuses on the intensive instead of extensive margin of child labor.

36. For this set of estimates, only the differential impact of the intervention on boys versus girls is given. The impact on boys and girls themselves is not available.

37. These estimates do not correct for truncation of the outcome variable, but other estimates provided in the paper suggest that most results are robust to corrections for censoring.

38. A similar argument underlies the well-known theoretical work of [Basu and Van \(1998\)](#), who argue that pulling all children out of work through a ban may jolt the labor market to another equilibrium in which adult wages are higher and children do not work.

39. [Galiani and McEwan \(2013\)](#) find some evidence of reductions in child work in the poorest quintile.

40. We do not consider the older age groups also discussed in the paper.

41. Because the results are preliminary, we have not included them in our discussion of the impact of unconditional and conditional cash transfers on the intensive margin of child labor.

42. Amounts partly based on the figures provided in [Fiszbein and Schady \(2009\)](#).

43. We have no estimate for the amount transferred as a percentage of household income for the Female School Stipends program in Pakistan.

44. Several other studies do not directly test whether conditional cash transfers protect children from participation in work when the household is hit by an economic shock. However, they do suggest that conditional cash transfers can reduce child labor during economic downturns. [Maluccio \(2005\)](#), for instance, shows that Red de Protección Social reduced participation in economic activities among children living in Nicaragua's coffee-growing regions during a sharp downturn in coffee prices in 2001 and 2002. [Sparrow \(2007\)](#) finds similar results in his study of Indonesia's Jaringan Pengaman Social emergency conditional cash transfer program. However, these results contrast with those of [Amarante, Ferrando, and Vigorito \(2011\)](#), who find no effect of Uruguay's anti-crisis conditional cash transfer program on child labor.

45. The point estimates for conditional cash transfers in combination with investments in health and education facilities exceed those of conditional cash transfers only, but the estimated coefficients are not significantly different from each other.

46. Moreover, there is some evidence of spillover effects at the locality level. The likelihood of working decreased for children from non-program households in treatment localities. However, for these same children the likelihood of working long hours (10 or more per week) increased.

47. Fiszbein and Schady (2009) also caution that potential detrimental spillover effects within households deserve further scrutiny.

References

- Alam, A., J. E. Baez, and X. V. Del Carpio. 2011. "Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan." Policy Research Working Paper 5669. World Bank, Policy Research Department, Washington, DC.
- Amarante, V., M. Ferrando, and A. Vigorito. 2011. "School Attendance, Child Labor, and Cash Transfers. An Impact Evaluation of PANES." PIERI Working Paper 2011–22. Partnership for Economic Policy, Québec.
- Attanasio, O., E. Fitzsimons, A. Gomez, M. I. Gutiérrez, C. Meghir, and A. Mesnard. 2010. "Children's Schooling and Work in the Presence of a Conditional Cash Transfer Program in Rural Colombia." *Economic Development and Cultural Change* 58 (2): 181–210.
- Augsburg, B., R. de Haas, H. Harmgart, and C. Meghir. 2012. "Microfinance, Poverty, and Education." Working Paper 18538. National Bureau of Economic Research, Cambridge, MA.
- Baird, S., C. McIntosh, and B. Özler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics* 126 (4): 1709–53.
- Baird, S., F. H. G. Ferreira, B. Özler, and M. Woolcock. 2013. "Relative Effectiveness of Conditional and Unconditional Cash Transfers for Schooling Outcomes in Developing Countries: A Systematic Review." *Campbell Systematic Reviews* 2013 : 8.
- Barrera-Osorio, E., M. Bertrand, L. L. Linden, and F. Perez-Calle. 2008. "Conditional Cash Transfers in Education: Design Features, Peer and Sibling Effects. Evidence from a Randomized Experiment in Colombia." Policy Research Working Paper 4580. World Bank, Policy Research Department, Washington, DC.
- . 2011. "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia." *American Economic Journal: Applied Economics* 3 (2): 167–95.
- Bando, R., L. E. Lopez-Calva, and H. A. Patrinos. 2005. "Child Labor, School Attendance, and Indigenous Households: Evidence from Mexico." Policy Research Working Paper 3487. World Bank, Policy Research Department, Washington, DC.
- Basu, K., and P. H. Van. 1998. "The Economics of Child Labor." *American Economic Review* 88 (3): 412–27.
- Beegle, K., R. H. Dehejia, and R. Gatti. 2006. "Child Labour and Agricultural Shocks." *Journal of Development Economics* 81 (1): 80–96.
- Behrman, J. R., J. Gallardo-García, S. W. P. P. E. Todd, and V. Vélez-Grajales. 2012. "Are Conditional Cash Transfers Effective in Urban Areas? Evidence from Mexico." *Education Economics* 20 (3): 233–59.
- Behrman, J. R., S. W. Parker, and P. E. Todd. 2011. "Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Followup of PROGRESA/Oportunidades." *Journal of Human Resources* 46 (1): 93–122.
- Benhassine, N., E. Devoto, E. Duflo, P. Dupas, and V. Pouliquen. 2012. "Unpacking the Effects of Conditional Cash Transfer Programs: Experimental evidence from Morocco." Unpublished Manuscript.

- Borraz, E., and N. González. 2009. "Impact of the Uruguayan Conditional Cash Transfer Program." *Cuadernos de Economía* 46 (Novembre): 243–71.
- Bourguignon, E., F. H. G. Ferreira, and P. G. Leite. 2003. "Conditional Cash Transfers, Schooling, and Child Labor: Micro Simulating Brazil's Bolsa Escola Program." *World Bank Economic Review* 17 (2): 229–54.
- Buddelmeyer, H., and E. Skoufias. 2004. "An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA." IZA Discussion Paper 827 and Policy Research Working Paper 3386. World Bank, Policy Research Department, Washington, DC.
- Cardoso, E., and A. Portela Souza. 2003. "The Impact of Cash Transfers on Child Labor and School Attendance in Brazil." Working Paper 0407. Vanderbilt University Department of Economics, Nashville, TN.
- Cigno, A., and F. C. Rosati. 2005. *The Economics of Child Labour*. Oxford University Press.
- Covarrubias, K., B. Davis, and P. Winters. 2012. "From Protection to Production: Productive Impacts of the Malawi Social Cash Transfer Scheme." *Journal of Development Effectiveness* 4 (1): 50–77.
- Dammert, A. C. 2009. "Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua." *Economic Development and Cultural Change* 58 (1): 53–83.
- Dubois, P., and M. Rubio-Codina. 2012. "Child Care Provision: Semiparametric Evidence from a Randomized Experiment in Rural Mexico." *Annales d'Économie et de Statistique* 105/106: 155–84.
- Deaton, A. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–55.
- De Carvalho Filho, I. E. 2012. "Household Income as Determinant of Child Labor and School Enrollment in Brazil: Evidence from a Social Security Reform." *Economic Development and Cultural Change* 60 (2): 339–435.
- De Hoop, J., and F. C. Rosati. 2014. "Does Promoting School Attendance Reduce Child Labour? Evidence from Burkina Faso's BRIGHT Project." *Economics of Education Review* 39: 78–96.
- De Janvry, A., F. Finan, E. Sadoulet, and R. Vakis. 2006. "Can Conditional Cash Transfer Programs Serve as Safety Nets in Keeping Children at School and from Working when Exposed to Shocks?" *Journal of Development Economics* 79 (2): 349–73.
- Del Carpio, X. V., and N. V. Loayza. 2012. "The Impact of Wealth on the Amount and Quality of Child Labour." Policy Research Working Paper 5959. World Bank, Policy Research Department, Washington, DC.
- Del Carpio, X. V., and K. Macours. 2010. "Leveling the Intra-household Playing Field: Compensation and Specialization in Child Labour Allocation." In *Child Labour and the Transition between School and Work (Research in Labour Economics, Volume 31)*, ed. R. K.Q. Akee, E. V. Edmonds, and K. Tatsiramos, 259–96. Bingley: Emerald Group Publishing Limited.
- DSD, SASSA, and UNICEF. 2012. *The South African Child Support Grant Impact Assessment: Evidence from a survey of children, adolescents and their households*. Pretoria: UNICEF South Africa.
- Duflo, E., and M. Kremer. 2005. "Use of Randomization in the Evaluation of Development Effectiveness." In *Evaluating Development Effectiveness*, ed. G. K. Pitman, O. N. Feinstein, and G. K. Ingram, 205–39. New Brunswick: Transaction Publishers.
- Duryea, S., D. Lam, and D. Levison. 2007. "Effects of Economic Shocks on Children's Employment and Schooling." *Journal of Development Economics* 84 (1): 188–214.
- Edmonds, E. V. 2006. "Child Labour and Schooling Responses to Anticipated Income in South Africa." *Journal of Development Economics* 81 (2): 386–414.
- . 2007. "Child Labour." In *Handbook of Development Economics, Volume 4*, ed. T. P. Schultz, and J. Strauss, 3607–709. Amsterdam: Elsevier Science.

- Edmonds, E. V., and N. Schady. 2012. "Poverty Alleviation and Child Labor." *American Economic Journal: Economic Policy* 4 (4): 100–24.
- Ferreira, F. H. G., D. Filmer, and N. Schady. 2009. "Own and Sibling Effects of Conditional Cash Transfer Programs." Policy Research Working Paper 5001. World Bank, Policy Research Department, Washington, DC.
- Ferro, A. R., A. L. Kassouf, and D. Levison. 2010. "The impact of conditional cash transfer programs on household work decisions in Brazil." In *Child Labour and the Transition between School and Work (Research in Labour Economics, Volume 31)*, ed. R. K.Q. Akee, E. V. Edmonds, and K. Tatsiramos, 193–218. Bingley: Emerald Group Publishing Limited.
- Fiszbein, A., and N. Schady. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. World Bank, Washington DC.
- Fitzsimons, E., and A. Mesnard. Forthcoming. "Can Conditional Cash Transfers Compensate for a Father's Absence?" *World Bank Economic Review*.
- Galiani, S., and P. J. McEwan. 2013. "The Heterogeneous Impact of Conditional Cash Transfers." *Journal of Public Economics* 103 : 85–96.
- Gee, K. A. 2010. "Reducing Child Labour through Conditional Cash Transfers: Evidence from Nicaragua's Red de Protección Social." *Development Policy Review* 28 (6): 711–32.
- Glewwe, P., and P. Olinto. 2004. "Evaluation of the Impact of Conditional Cash Transfers on Schooling: An Experimental Analysis of Honduras' PRAF Program." Working Paper. International Food Policy Research Institute, Washington, DC.
- Guarcello, L., F. Mealli, and F. C. Rosati. 2010. "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing, and Insurance." *Journal of Population Economics* 23 (1): 169–98.
- International Labour Office. 2013a. *Marking Progress Against Child Labour: Global Estimates and Trends 2000–2012*. Geneva.
- International Labour Office. 2013b. *World Report on Child Labour: Economic Vulnerability, Social Protection and the Fight against Child Labour*. Geneva.
- Kazianga, H., D. de Walque, and H. Alderman. 2009. "Educational and Health Impacts of Two School Feeding Schemes. Evidence from a Randomized Trial in Burkina Faso." Policy Research Working Paper 4976. World Bank, Policy Research Department, Washington, DC.
- King, E. M., and J. R. Behrman. 2009. "Timing and Duration of Exposure in Evaluations of Social Programs." *World Bank Research Observer* 24 (1): 55–82.
- Kremer, M., and A. Holla. 2009. "Pricing and Access: Lessons from Randomized Evaluations in Education and Health." In *What Works in Development: Thinking Big and Thinking Small*, ed. J. Cohen, and W. Easterly, 91–129. Washington, DC: The Brookings Institution.
- Levy, D., and J. Ohls. 2007. "Evaluation of Jamaica's PATH Program." Research Report No. 8966-090. Mathematica Policy Research, Inc., Washington, DC.
- Maluccio, J. A. 2005. "Coping with the 'Coffee Crisis' in Central America: The Role of the Nicaraguan Red de Protección Social." FCND Discussion Paper 188. International Food Policy Research Institute, Washington, DC.
- Maluccio, J. A., and R. Flores. 2005. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." Research Report 141. International Food Policy Research Institute, Washington, DC.
- Miller, C., and M. Tsoka. 2012. "Cash Transfers and Children's Education and Labour among Malawi's Poor." *Development Policy Review* 30 (4): 499–522.
- Nelson, L. K. 2011. "From Loans to Labour: Access to Credit, Entrepreneurship, and Child Labor." Paper presented at the 2011 NEUDC Conference, Yale University, Economic Growth Center, November 12–13. NEUDC, New Haven, CT.

- Ravallion, M. 2009. "Evaluation in the Practice of Development." *World Bank Research Observer* 24 (1): 29–53.
- Ravallion, M., and Q. Wodon. 2000. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy." *Economic Journal* 110 (March): C158–C175.
- Rawlings, L. B., and G. M. Rubio. 2005. "Evaluating the Impact of Conditional Cash Transfer Programs." *World Bank Research Observer* 20 (1): 29–55.
- Rubio-Codina, M. 2010. "Intra-household Time Allocation in Rural Mexico: Evidence from a Randomized Experiment." In *Child Labour and the Transition between School and Work (Research in Labour Economics, Volume 31)*, ed. R. K.Q. Akee, E. V. Edmonds, and K. Tatsiramos, 219–57. Bingley: Emerald Group Publishing Limited.
- Saavedra, J. E., and S. Garcia. 2012. "Impact of Conditional Cash Transfer Programs on Educational Outcomes in Developing Countries: A Meta Analysis." Working Paper WR-921-1. Rand Corporation, Santa Monica, CA.
- Schady, N. R., and M. C. Araujo. 2006. "Cash Transfers, Conditions, School Enrollment, and Child Work: Evidence from a Randomized Cash Transfer Experiment in Ecuador." Policy Research Working Paper 3930. World Bank, Policy Research Department, Washington, DC.
- Schultz, T. P. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74 (1): 199–250.
- Skoufias, E., and S. W. Parker. 2001. "Conditional Cash Transfers and their Impact on Child Work and Schooling: Evidence from the PROGRESA Program in Mexico." *Economía* 2 (1): 45–96.
- Sparrow, R. 2007. "Protecting Education for the Poor in Times of Crisis: An Evaluation of a Scholarship Programme in Indonesia." *Oxford Bulletin of Economics and Statistics* 69 (1): 99–122.
- Udry, C. 2006. "Child Labor." In *Understanding Poverty*, ed. A. V. Banerjee, R. Bénabou, and D. Mookherjee, 243–58. Oxford: Oxford University Press.
- Understanding Children's Work. 2010. *Joining Forces against Child Labour: Inter-agency Report for The Hague Global Child Labour Conference of 2010*. Geneva.
- Yap, Y. T., G. Sedlacek, and P. F. Orazem. 2002. "Limiting Child Labor through Behavior-Based Income Transfers: An Experimental Evaluation of the PETI Program in Rural Brazil." Working Paper 223. Inter-American Development Bank, Washington, DC.

The Promise of Performance Pay? Reasons for Caution in Policy Prescriptions in the Core Civil Service¹

Zahid Hasnain², Nick Manning³, and Jan Henryk Pierskalla⁴

There is a vast body of literature on performance-related pay (PRP), with strongly held views from opponents and proponents. This study reviews this literature, disaggregating the available evidence by the different public sector contexts, particularly the different types of public sector jobs, the quality of the empirical study, and the economic context (developing country or OECD settings), with the aim of distilling useful lessons for policy-makers in developing countries. The overall findings of the review are generally positive across these contextual categories. In particular, the findings from high quality studies, based on a simple scoring method for internal and external validity, of PRP in public sector-equivalent jobs show that explicit performance standards linked to some form of bonus pay can improve the desired service outcomes, at times dramatically. This evidence primarily concerns “craft” jobs, such as teaching, health care, and revenue administration, apparently negating (at least in the short term) the behavioral economics concern about the crowding out of intrinsic incentives. The available evidence suggests that if policy-makers are sensitive to design and vigilant about the risks of gaming, then PRP may result in performance improvements in these jobs in developing countries. However, it is difficult to draw firm conclusions from the review about the effect of PRP in core civil service jobs for three reasons. First, there are very few studies of PRP in these organizational contexts. The work of senior administrators in the civil service is very different from that of many private sector jobs and is characterized by task complexity and the difficulty of measuring outcomes. Second, although some studies have shown that PRP can work in even the most dysfunctional bureaucracies in developing countries, there are few cases illustrating its effectiveness or otherwise outside OECD settings. Finally, few studies follow PRP effects over time, providing little information on long-term effects and adjustments in staff

behavior. We conclude that more empirical research is needed to examine the effects of PRP in the core civil service in developing countries.

JEL codes: H11, H83, I18, I28

Performance-related pay (PRP) is increasingly used in public sector organizations across the world. Currently, 28 of the 34 OECD countries have introduced PRP in some form (OECD 2011). Similar movements are underway in middle-income countries and, more sporadically, in lower-income countries in the health and education sectors. A vast theoretical and empirical body of literature has analyzed various dimensions of PRP, and there is now a small but growing body of robust evidence on the impact of PRP that is shedding new light on what is achievable and under what specific conditions.

The objective of this paper is to provide a review of this literature on PRP with relevance for the public sector, spanning the fields of public administration, psychology, economics, education, and health, with the aim of distilling useful lessons for policy-makers in developing countries. This is by no means the first comprehensive review of this literature, but it is, to our knowledge, the first that aims to disaggregate the available evidence by (i) the quality of the empirical study; (ii) the different public sector contexts, particularly the different types of public sector jobs; and (iii) country context (developing country or OECD settings). Although large parts of the existing literature address performance pay in the private sector or the OECD country context, this disaggregation allows us to identify more nuanced lessons for the application of performance pay in developing countries' core civil services.

PRP can be defined as a compensation arrangement in which the final salary of an employee is a function of some form of measured "performance". How performance is measured, who measures it, and how it is linked to salary can all vary considerably. Performance can be based on qualitative assessments or quantitative measures of inputs, outputs, or outcomes. Salary can be either wholly a function of performance, such as piece-rate pay in a manufacturing setting, or a combination of base pay and one-off bonuses or merit increases of base pay. Bonuses and merit increases can be awarded on an individual, small team, or larger departmental basis. Evaluations can be implemented by direct supervisors, human resource specialists, peer panels, or outside agencies. Once performance has been measured, it must be evaluated against a performance standard. This standard can be based on individually pre-agreed goals, absolute performance against minimum or scaled standards, relative improvements against past performance, rank-order of performance in a tournament evaluation, or relative performance measured against co-workers, other teams of co-workers, or other schools or agencies nationally or regionally.

Drawing on contract theory and the problems of moral hazard, a significant part of the academic literature has examined the impact of PRP on increasing effort and

reducing shirking. In work environments where effort is unobservable, fixed pay contracts provide little ability for employers to influence employee effort. This is a particular problem in traditional civil service jobs that are characterized by uniform pay for jobs in similar grades, pay increases based largely on seniority, and negligible probability of termination. Contracts that tie observable outputs, which are correlated with unobservable effort, to desired pay incentives can theoretically mitigate these difficulties. Contract theory also suggests that PRP can help address the problem of adverse selection by encouraging high ability individuals who expect to do better under a performance pay scheme to join the agency and discouraging low ability individuals (the “sorting” effect).

Critics counter that PRP cannot work when outputs are difficult to measure and when tasks are multi-dimensional because it results in “gaming” behavior, whereby effort is only allocated toward activities that are observed and measured, which may not improve overall outcomes. The psychological and behavioral economics literature also argues that individuals are motivated by intrinsic concerns about the inherent social value of the job, particularly in the public sector, and that PRP, which explicitly focuses on extrinsic benefits, might crowd out intrinsic motivation and reduce worker productivity.

Our evaluation of existing studies finds a mounting body of evidence that PRP *can* increase worker effort in specific organizational contexts, particularly in those jobs where outputs are more easily measured. The evaluation proceeds in stages, first reviewing the evidence concerning PRP in general, regardless of whether the concerned jobs are in the public or private sectors. Then, drawing from the seminal work by J. Q. Wilson (Wilson 1989), a distinction is made between jobs with observable or unobservable production processes as a rough proxy for distinguishing between jobs that tend to be found in the private or public sectors. A number of increasingly rigorous studies document that performance pay can improve desired outcomes for teaching or health care provision in general, and in developing countries in particular, if it is carefully designed and implemented. Next, further distinctions are drawn between public sector jobs in which the outputs are measurable and those in which they are not as a further and likely more precise proxy for distinguishing between broad public service jobs and those in the core administration and between high-quality studies with high internal and external validity and others.⁵ Although the universe of studies under consideration draws on all relevant studies regardless of their location, this winnowing reveals that there are no high-quality studies of PRP relevant to tasks in the core civil service outside of the OECD. Thus, with the current state of knowledge, it is not possible to infer much about the effectiveness of PRP for core civil service functions in developing countries.

Theoretical Debates

Theoretical debates on PRP have been evolving in the context of private businesses (Prendergast 1998; Prendergast 1999), the general public sector (Dixit 1999; Burgess and Ratto 2003; Perry et al. 2006), and specific occupations, such as teaching (Neal 2011). These debates can be roughly divided into psychological theories on human motivation and training, which are popular in public administration research, and core economic theories on incentive structures and principal-agent problems, with behavioral economics building a bridge between the two.

Expectancy and Reinforcement Theory

Public administration research on PRP usually relies on what is often called “expectancy” (Vroom 1964; Porter and Lawler III 1968) and “reinforcement” theory (Skinner 1969; Luthans 1973). In its simplest form, the theory suggests that explicit incentives in the form of performance pay work under two conditions: first, employees need to believe that increased effort leads to increased performance; second, increased performance leads to desired outcomes and is recognized by management. If these two conditions are met, employees form a behaviorally salient expectation about a future reward and adjust their work effort upward. Reinforcement theory stresses the effect of cultivating a behavioral norm of high work effort by reinforcing behavior with positive rewards.

In addition to this direct link between PRP and individual effort, advocates in the field of public administration highlight the secondary effects of PRP: it helps recruit and retain highly skilled and/or motivated staff who presumably would do better under the scheme; it makes managers more committed to the strategic objectives and core organizational goals of the agency and increases the link between individual and organizational goals; it weakens the power of public sector unions to impose restrictive working practices; and it reduces the overall wage bill by moving away from automatic pay increases (Marsden 2004; OECD 2005b; Marsden 2009).

Critics of performance pay suggest that it is difficult to design performance pay schemes that meet the two conditions of expectancy and reinforcement theory (Kerr 1975). They argue that people do not always approach work effort and the assessment of salary in an entirely rational way. In addition, many core public servants perform services that are difficult to measure or are non-measurable, or they produce outputs that are not market-priced. For example, early critics of test-score-based school and teacher evaluations argued that teacher performance cannot be neatly summarized by mechanical student test scores and that such practices invite behavior that contradicts the overall goals of the teaching profession (Murnane and Cohen 1986). The use of explicit and objective performance measures can induce tunnel vision, myopia, and measure fixation (Propper and Wilson 2003).

Additionally, civil servants often work in large teams under the supervision of multiple managers, complicating the attribution of performance and the responsibility of performance evaluation. Some authors see the presence of high levels of trust and transparency between employees and management as a necessary condition for the effectiveness of PRP to avoid arbitrary implementation and worker dissatisfaction (Kellough and Lu 1993).

A different strand of criticism focuses on other motivations underlying public servants' efforts. Civil servants, it is argued, are motivated by notions of altruism, pro-social behavior, and commitment to institutional goals (Perry and Hondeghem 2008), which may compete or even conflict with explicit monetary incentives.

Incentive and Principal-Agent Theory

The most basic argument for incentive pay is based on a simple microeconomic principal-agent model of labor relations, in which a principal (the employer) wants to induce an agent (the employee) to perform a certain task. Such principal-agent relationships are commonly affected by two problems (Dixit 1999): moral hazard and adverse selection.

Moral hazard describes a scenario in which the agent's actions affect the principal's payoffs, but the action is not directly observable to the principal. In the workplace, the employee's effort is not directly observable, but it influences productivity and outcomes about which the employer cares. Under these conditions, offering fixed pay contracts to workers gives the employer little leverage to influence employee effort after hiring decisions have been made, a problem that is exacerbated if employees are difficult to fire. PRP can address this problem of moral hazard by tying observable outputs, which are correlated with unobservable effort, to pay.

In the case of adverse selection, the agent has access to private and valuable information at the time of contract signing. Adverse selection in the public sector plays an important role in recruitment, where low- and high-skilled applicants are difficult to distinguish based on public information. Public agencies need to offer contracts that induce high-quality applicants to apply and deter low quality applicants from misrepresenting their qualifications. PRP can alleviate this sorting problem because higher-quality workers will expect to perform better under this system of pay and therefore will be more likely to apply for a job opening (Delfgaauw and Dur 2008).

A well-known criticism of performance pay arrangements is that when tasks are multi-dimensional, incentivizing only tasks that are observable and measurable will not necessarily improve overall outcomes but rather will lead to a substitution of effort from unobservable to observable tasks, which can lead to worse outcomes (Holmstrom and Milgrom 1991). For example, the task of teaching can involve both instruction based on sound curricula and coaching on test-taking strategies, and

poorly designed incentive schemes can encourage teachers to re-allocate effort to the latter and away from the former (known as “teaching to the test”), to the detriment of human capital accumulation.

The problem of selecting appropriate performance measures to address this problem has spawned its own theoretical and empirical debates (Courty et al. 2005). A problem related to the multi-tasking argument addresses the issue of gaming or cheating incentive systems. Typical examples are the outright manipulation of results, whether through cream-skimming (i.e., the manipulative selection of clients to improve program effects; Heckman et al. 1997) or other forms of manipulation, such as the provision of high-caloric food to students on test days (Figlio and Winicki 2005).

The problem of gaming performance standards argues for ongoing adjustments in targets and metrics by the principal (Courty and Marschke 2003). To counteract the excessive gaming of incentive schemes, it has been suggested that evaluation systems should be used independently of output measurements (Neal 2011). Additionally, relative performance schemes in which employees are ranked against each other, potentially in a formal tournament setting, are much more difficult to manipulate (Barlevy and Neal 2011; Neal 2011). However, (Marsden 2009) notes that the gaming is not necessarily restricted to ingenious behaviors on the part of staff. Managers might be tempted to “collude with their subordinates: to go through the motions and fill in the forms for goal setting and appraisal, but not to worry about the reality” (Marsden 2009, 5)

Rewarding team performance can have certain advantages, ranging from reduced evaluation costs to the avoidance of harmful competition between employees. However, basing rewards on team outputs can also lead to problems of free-riding, where some team members willfully reduce their efforts in the expectation of relying on the work of others (Dixit 1999).

Choosing the correct bonus size brings its own challenges. Small bonuses will have little incentive effects and fall short of expectations, whereas large bonuses can lead employees to treat incentive schemes as pure lotteries, especially if outcomes are strongly stochastic (e.g., student test scores (Neal 2011)).

Behavioral Economics — Intrinsic versus Extrinsic Motivation

When staff perform well because of the inherent characteristics of the job they are doing, e.g., its interest or perceived social value, they are said to be “intrinsically” motivated. When staff performance stems from rewards that are unrelated to the nature of the job, staff are said to be “extrinsically” motivated. Intrinsic motivations are of particular significance in public service (Banuri and Keefer 2013). Behavioral economists have argued that PRP can lead to a reduction in effort by crowding out intrinsic motivation as employees change their perceptions about the

nature of their work. Several authors (Kreps 1997; Benabou and Tirole 2003; Benabou and Tirole 2006) have developed formal treatments of the trade-off between extrinsic and intrinsic motivation, building on the literature from psychology. Crowding out can be especially salient if performance pay is introduced using antagonistic framing and can stifle creativity and collaboration (Frey and Osterloh 1999). More generally, Pink (2009) argues that monetary and other extrinsic incentives are both counterproductive (because they frequently undermine intrinsic incentives) and unnecessary (because intrinsic incentives can be harnessed and used to maximize individual productivity). Drawing on empirical work (Ryan and Deci 2000; Chirkov et al. 2003; Sauermann and Cohen 2008; Niemiec et al. 2009), his theory suggests that tasks can be constructed to maximize an individual's sense of (i) *autonomy*; (ii) *mastery* (continuous incremental learning and improvements rather than distant targets); and (iii) *purpose*, improving overall performance.

Another psychological argument, known as the “Yerkes-Dodson law”, highlights the phenomenon of “choking under pressure” (Ariely et al. 2009). The argument is that performance has an “inverse U” relationship with the level of the incentive payment, with performance improving at low and moderate levels of incentive payments compared to no payments but becoming worse at very high levels of payment compared to moderate, low, or even no payments.

Finally, behavioral economists have identified the possibility of satisficing instead of maximizing behavior. Employees might exert effort until a certain minimum level of reward is reached and then substitute additional labor supply for increased leisure or idle time (Camerer et al. 1997).

Organizing the Empirical Evidence

The Type of Job

The hypothesized effects of PRP are heavily contingent on organizational context, particularly the nature of the job in which PRP is introduced. Borrowing James Q. Wilson's typology (Wilson 1989), jobs can be characterized by whether the job's outputs are easily measurable and whether the actions in the job to produce the output or the internal production process are observable.⁶ Table 1 provides a framework to organize the empirical evidence by job type, with the simplifying assumption that jobs with multiple dimensions are located within the cell that represents the most complex of those dimensions. The top left box describes “Production Jobs”, in which outputs are easily measurable, the production process consists of repeatable, mechanical tasks that are observable to an outside monitor, and controllability is likely to be high. Typical examples are manufacturing factory-floor jobs and municipal services, such as garbage collection. If the production process is not directly

Table 1. James Q. Wilson’s Classification of Job Types

		<i>Actions or internal production process of the job</i>	
		<i>Observable</i>	<i>Not observable</i>
Outputs from the job	Relatively easily measurable	Production job: Simple repetitive stable tasks, specialized skills. Examples: Manufacturing, sales, simpler municipal services (garbage collection).	Craft jobs: Application of general sets of skills to unique tasks, but with stable, similar outcomes. Examples: Auditing; revenue collection; teaching; medical practice; job placement work
	Not easily measurable	Procedural job: Specialized skills; stable tasks, but unique outcomes Examples: Military	Coping job: Application of generic skills to unique tasks, but outcomes cannot be evaluated in absence of alternatives Examples: Core civil service; managerial jobs in large private sector organizations

Source: Adapted from [Wilson \(1989\)](#).

observable but outputs remain measurable, such jobs are termed “Craft Jobs”. With recent advances in measuring learning outcomes, teaching can be classified as a job in which the exact process of production is difficult to ascertain, but, at least to a certain degree, desired outputs are quantifiable. Similarly, some of the outputs of healthcare, particularly in preventative services such as child immunization, are more measurable. Other examples include tax collection, job placement services, and auditing.

In the bottom row are “Procedural Jobs” and “Coping Jobs”. Both are characterized by difficult-to-measure outputs, but they differ in the observability of the production process to an outsider. Procedural jobs such as the military have clearly defined inputs, whereas policy jobs in the core civil service neither produce easily measurable outputs nor have transparent production processes. Coping jobs present the most challenging functional contexts for PRP.

We use the distinction between jobs with observable and unobservable production processes (center and right-hand columns of Table 1) as a rough proxy for distinguishing between jobs that tend to be found in the private or public sectors. Within the latter, we use the distinction between public sector jobs in which the outputs are measurable and those in which they are not (center and bottom rows in Table 1) as a further and likely more precise proxy for distinguishing between broad public service jobs and those in the core administration. The universe of studies considered for this evaluation were drawn from anywhere in the matrix, but we consider studies of PRP in coping jobs to be likely the best measure of the impact of performance-related pay in the primarily policy-jobs in the core civil service.

Using this classification system allows us to draw on results from studies of private sector jobs if the jobs themselves are essentially similar to those found in the public sector. The general literature on performance pay in the private sector has focused on activities classifiable as production jobs and has found fairly encouraging results (Stajkovic and Luthans 2003). The task of this evaluation was to determine what survived of this general finding when the studies were winnowed down to coping jobs, specifically to high-quality studies of PRP in coping jobs.

Methodological Approaches

The vast empirical literature on PRP utilizes a range of methodological approaches. Early studies on performance pay for administrative civil service, teaching, and healthcare were largely observational. Studies that focused on administrative jobs were either qualitative case studies or used convenience samples of employees and senior management to collect data on self-assessed motivation and satisfaction with newly introduced performance pay. Studies of performance pay for teachers were more quantitative, assessing the impact of bonus schemes on actual public service outcomes (e.g., drop-out rates, students' grades, or test scores).

Although an improvement, many of these quantitative observational studies and similar work on private companies and the public service fall short of an ideal research design for causal inference on program effects. The gold standard for program evaluation is the use of randomized-controlled trials (RCTs), in which treatment assignment to subjects is randomized and unrelated to other observable and unobservable characteristics. This randomization allows the estimation of the treatment effect by comparing the treated and control units. Observational studies, in contrast, rely on treatment and control groups created not by controlled random assignment but by real-world social processes. Issues of selection bias and confounding factors can undermine the internal validity of such studies.

Utilizing the power of a randomization framework, several behavioral economists have used laboratory experiments to test hypotheses with regard to incentive schemes. Laboratory experiments offer at least two distinct advantages over observational studies: the researcher can use randomization to ensure the identification of treatment effects, and researchers can design their experiments to directly relate to the theoretical questions at hand. However, laboratory experiments often use notoriously small samples and student subjects who share few characteristics with actual workers or public servants, and they cannot replicate real workplace settings or offer bonus schemes that remotely approach the bonus sizes common even in moderately incentivized performance pay schemes.

The most recent attempts to address the issue of proper causal inference on performance pay and to increase the representativeness of results are RCTs in the field. In a field RCT, researchers are able to randomize key features of an actual policy

program that services the population of interest. The advantage of such a field experiment is the similarity of the target population, the structure of the incentive program, and the randomization of treatment. Although field RCTs are time- and resource-intensive studies, several teams of researchers have implemented similar studies in different contexts around the world, adding considerably to the empirical understanding of performance pay.

Assessing the Evidence

Classifying the Universe of Studies

In total, 153 empirical studies of PRP were considered in this review (see [Hasnain et al. \(2012\)](#) for the full list), of which 110 were for craft and coping jobs (Table 2), our proxy for general public sector jobs, with 17 of these for coping jobs specifically (our proxy for jobs in the core public administration). The research to date on the subject has largely focused on advanced countries; in the review, 127 studies are in OECD contexts, and only 26 are in developing country settings. The literature has also focused largely on craft jobs and production jobs, with no experimental studies to date on coping jobs.

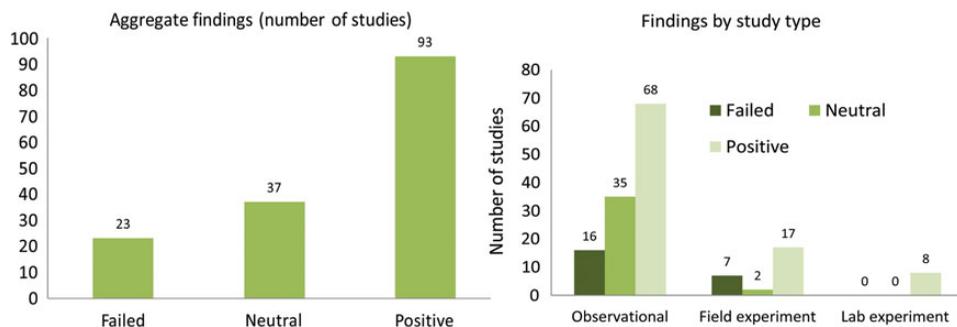
The 153 studies that were reviewed were grouped into three categories to capture the effect of PRP that they revealed: *positive* if their findings provided positive evidence of the effectiveness of incentive schemes;⁷ *neutral* if the study was largely descriptive or found contradictory evidence; and *failed* if the evidence indicated no effect or a negative effect of performance pay. Figure 1 shows the overall frequency of results. A majority of studies (93 of the 153) presented supportive evidence for

Table 2. Studies by Country Environment, Methodology, and Job Type

<i>Country and methodology</i>	<i>Types of Jobs</i>					<i>Total</i>
	<i>Production jobs</i>	<i>Procedural jobs</i>	<i>Coping jobs</i>	<i>Craft jobs</i>	<i>Unclassified</i>	
<i>OECD study</i>	27	0	16	71	13	127
Observational	14	0	16	58	13	101
Field RCT	7	0	0	13	0	20
Lab. experiment	6	0	0	0	0	6
<i>Developing country study</i>	1	0	1	22	2	26
Observational	0	0	1	15	2	18
Field RCT	0	0	0	6	0	6
Lab. experiment	1	0	0	1	0	2
Total	28	0	17	93	15	153

Source: Authors' own findings.

Figure 1. Aggregate Findings on Performance-related Pay



Source: Authors' own findings.

some type of effect of performance pay schemes, with experimental studies showing more positive findings than observational ones.

In drawing lessons, however, it is important to distinguish the findings more systematically by their research quality. Study quality was ranked in two different ways. First, each study was assessed for its “internal validity”, or strength of the causal arguments, using a five-point ranking (from weak to strong) as follows:

1. no empirical study or a faulty research design⁸
2. descriptive; small sample size⁹
3. secondary data analysis and/or descriptive data analysis; small sample size; some statistical analysis¹⁰
4. quasi-experimental design; reasonable sample size; conclusions based on statistical analysis¹¹
5. laboratory experiments; randomized controlled trial; large sample size; strong statistical analysis; strong conclusions.

Second, studies were evaluated on the dimension of “external validity”, or the extent to which the causal connections in the specific context of the study would remain valid if replicated in other contexts. For example, lab experiments and RCTs offer very strong evidence of causality (high internal validity), but in a specific context: they tell us the average impact of a particular intervention in a particular location with a particular sample at a particular point in time. They are often accused of being low in external validity because the study subjects (usually college students, in the case of laboratory experiments) are not representative of the general population or, in this case, the population of interest (civil servants) and the requirements of the experiment imply very particular conditions that may not approximate real-world settings.

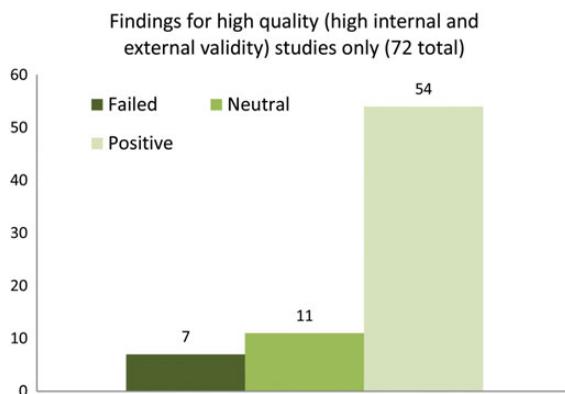
Applying these two quality filters to the 153 studies resulted in 72 high-quality studies (ranked 4 or 5 on the internal validity scale and “high” on the external validity scale).¹² Of these, 54 high quality studies found positive effects of PRP (Figure 2).

General Observational and Experimental Studies

Previous meta studies have attempted to aggregate evidence concerning performance-related pay from observational and experimental studies across different task types (Jenkins et al. 1998; Condly et al. 2003; Weibel et al. 2009). General literature reviews (Petersen et al. 2006; Eldridge and Palmer 2009) highlight the degree to which studies have focused on OECD country experiences.

These overviews are generally positive regarding the impact of performance pay schemes. Most meta studies confine themselves to a broad qualitative assessment of the overall impact of PRP. Only two claim to offer quantitative data on the change in output per worker. Based on an analysis of 47 studies, (Jenkins et al. 1998) concluded that these incentives resulted in a 12 percent improvement in performance quantity and a negligible effect on performance quality. Based on a meta-analysis of 64 field and laboratory experiments as well as observational studies, Condly et al. (2003) concluded that employees and other research participants who received performance incentives achieved an 22 percent average increase in work performance. In a meta-analysis of 46 high-quality empirical studies covering both simple and complex tasks and with both quantitative and qualitative outcome measurements, Weibel et al. (2009) found a statistically significant and positive effect of

Figure 2. Findings by Research Quality



Source: Authors' own findings.

PRP on performance. The findings were significantly positive for simple tasks and significantly negative, albeit smaller, for complex tasks. The authors argued that this negative effect was because of the reduction in intrinsic motivation brought about by the incentive scheme.

Observational studies on performance-related pay are particularly suggestive that PRP has a positive impact, but with many caveats. [Belfield and Marsden \(2003\)](#) use panel data from a large UK workplace survey, which include both piece-rate jobs and “knowledge work” jobs in which performance is based on the achievement of previously agreed goals. They find strong effects of individual pay-for-performance, but only conditionally on the monitoring regime. Another study uses the British Household Panel Survey to distinguish the productivity and sorting effects of performance pay, finding that jobs with performance-related pay attract workers of higher ability and induce workers to provide greater effort ([Booth and Frank 1999](#)).

In contrast, [Beer and Cannon \(2004\)](#) study the failure of 13 incentive plans at Hewlett Packard using interviews and internal documents. They find that managers abandoned the programs because of the perceived costs. In a worldwide survey of 205 top managers, [Beer and Katz \(2003\)](#) document the weak support of incentive schemes among management.

These findings are in distinct contrast to the conclusions from qualitative observational OECD reports and discussion papers that report a negative effect of PRP on staff motivation, such as perceptions of unfairness, the difficulties of implementing such schemes given the tendency in bureaucracies to rate most employees as performing satisfactorily, and the political and operational difficulties of introducing any major pay reform within the public service ([OECD 1993](#); [OECD 1996](#); [OECD 1997](#); [Burgess and Ratto 2003](#); [OECD 2004](#); [OECD 2005a](#); [OECD 2005b](#); [Perry et al. 2006](#); [Ketelaar et al. 2007](#); [Rexed et al. 2007](#); [OECD 2008](#); [OECD 2009](#); [Perry et al. 2009](#)). This negative finding is mirrored in the smaller set of observational meta studies concerning developing countries ([World Bank 1999](#); [Kiragu and Mukandala 2003](#); [Independent Evaluation Group 2008](#)).

The evidence from experimental studies is largely from OECD countries. In a field experiment from the private sector ([Bandiera et al. 2006](#)), some managers were treated with the introduction of a performance-pay system, and the productivity of lower-tier workers was used as an outcome measure. The study finds evidence of both an incentive and sorting effect: managers supported their high-productivity workers and fired the least qualified employees. An experimental treatment of monitoring efforts by management in a call center found that employees largely behaved according to a rational-cheater model of human behavior, highlighting the importance of performance measurement as well as a substantial portion of employees that remains unaffected by monitoring attempts ([Nagin et al. 2002](#)).

Laboratory experiments have enabled the exploration of specific aspects of performance pay, including the functional relationship between bonus sizes and

performance, the incentive and sorting effects of performance pay, the impact of different types of bonuses, and the possible tradeoffs between extrinsic and intrinsic motivation. [Ariely et al. \(2009\)](#) explore the effect of bonus size on performance in laboratory experiments using subjects in the US and India. Participants had to solve cognitive tasks under time pressure and were incentivized with bonuses that varied from small to large relative to their normal pay. The authors find evidence of an “inverse-U” relationship between bonus size and performance, with a “choking-under-pressure” effect in which bonuses at very high levels lead to a worsening of performance.

[Cadsby et al. \(2007\)](#) find support for the incentive and sorting effect of performance pay and found that subjects with higher levels of risk aversion avoided pay-for-performance, suggesting important unintended side effects.

Addressing the problem of the multi-dimensionality of many tasks, [Fehr and Schmidt \(2004\)](#) conduct an experiment with university students to understand the effects of varying bonus schemes on effort provision on two distinct tasks, only one of which is contractible. They find that simple piece-rate contracts led to a focus on the contractible task, whereas bonus arrangements designed to be more encompassing and to explicitly address the multi-tasking problem also induce participants to spend time on the second task.

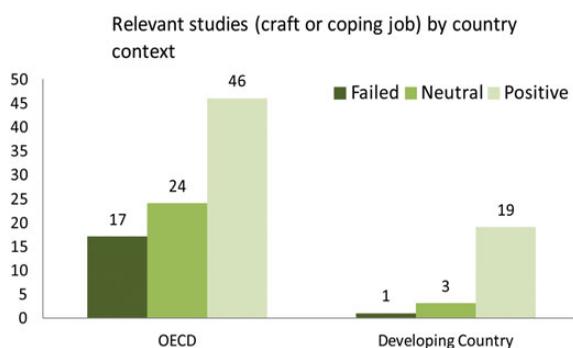
In a laboratory experiment by [Gneezy and Rustichini \(2000\)](#), high school and university students in Israel were offered different size of bonuses for specific tasks. The results suggest that subjects showed higher levels of productivity when offered large rewards but that small awards led to worse performance than did offering no monetary reward at all. This finding suggests the importance of framing of performance pay. If bonuses adequately communicate the importance of performing assigned tasks well compared to the overall goals of the organization, they can work; however, if bonuses trigger a change in the evaluation of the worker relationship, crowding out of intrinsic motivation can worsen productivity.

Studies on Jobs that are Representative of the General Public Sector

This paper examines the effect of PRP in craft and coping jobs because these jobs most closely resemble public sector organizational contexts. [Figure 3](#) presents the evidence for all of the reviewed studies for craft and coping jobs. Overall, 65 of the 110 studies found positive effects of PRP, with stronger evidence for the relatively few studies in developing countries.

Many of these studies are in the health sector. The British NHS introduced performance-pay elements into the remuneration of primary care physicians in 2004. Several studies ([Campbell et al. 2005](#); [Campbell et al. 2007](#); [Steel et al. 2007](#); [Vaghela et al. 2009](#); [Chalkley et al. 2010](#)) report positively on the impact of these incentives. In contrast, in the US, [Hillman et al. \(1991\)](#) report more mixed findings

Figure 3. Findings by Country Context



Source: Authors' own findings.

of similar schemes for physicians, and [Shen \(2003\)](#) and [Doran et al. \(2006\)](#) find evidence of “gaming”.

The few studies of PRP in the health sector in low- and middle-income countries have generally found positive results but illustrate data problems. [McNamara \(2005\)](#) discusses six cases of payment for quality in the health services sector across developed and developing countries, with cases in Nicaragua and Haiti having a positive effect. The Nicaraguan reform efforts combined the decentralization of decision-making authority and increased local accountability with explicit performance agreements. They were considered to have led to an overall improvement ([Jack 2003](#)), but it is difficult to disentangle the effects of each reform element. Similarly, in a recent study by [Witter et al. \(2011\)](#), a pay-for-performance arrangement in a NGO-led health project in the Battagram district of Pakistan was found to have improved general services provision, but with an unclear effect of the performance-based elements.

Other studies are clearer. [Meessen et al. \(2007\)](#) evaluate the performance of 15 health centers in Kabutare, Rwanda, and document a sharp increase in staff productivity after the introduction of output-based bonuses. [Soeters et al. \(2006\)](#) highlight the potential applicability of the Rwandan experience in sub-Saharan Africa more generally. In a similar vein, efforts to improve health services provision in Haiti using performance-based payment for NGOs in a USAID pilot project showed encouraging effects on immunization coverage and organizational behavior ([Eichler et al. 2001](#)).

Outside the health sector, revenue authorities have been studied extensively, most likely because they provide examples of craft jobs in which, although the methods of work are difficult to observe, the outputs (the number of audits conducted and tax fines collected) are more easily measurable. [Kahn et al. \(2001\)](#) examine a 1998

Brazil incentive scheme and found that it resulted in a 75 percent increase in fines per inspection. The [World Bank \(2001\)](#) concludes that “circumstantial evidence” suggests that bonus systems do seem to have an impact on organizational effectiveness in revenue administrations.

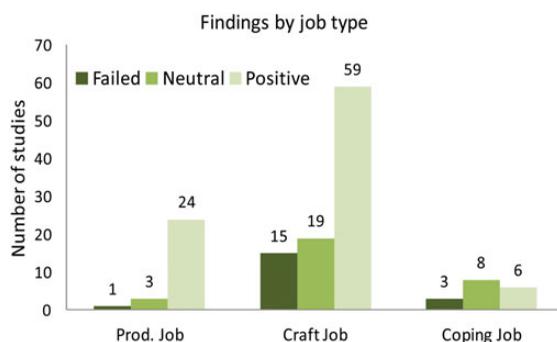
Overall, as Figure 4 illustrates, PRP for craft jobs is generally found to have a positive effect.

Narrowing by Quality

The higher-quality studies for craft and coping jobs that were reviewed were primarily in the education and health sectors. There is an extensive and growing body of literature on performance pay for teachers. In the US, most observational studies have primarily examined the impact of performance incentives on student test scores, although a few studies ([Clotfelter et al. 2004](#); [Clotfelter et al. 2007](#); [Clotfelter et al. 2008](#)) use detailed data from North Carolina schools to show that accountability and performance pay systems can also contribute positively to retaining quality teachers. With regard to student test scores, the evidence is mixed ([Eberts et al. 2002](#); [Dee and Keys 2004](#)). A number of studies identify problems of gaming, such as outright cheating ([Jacob and Levitt 2003](#); [Jacob 2005](#)) or, more subtly, the adjustment of the caloric content of school lunches to improve cognitive ability on test days ([Jacob and Levitt 2003](#); [Figlio and Winicki 2005](#); [Jacob 2005](#)).

Outside the US, an analysis by Atkinson et al. finds clear positive effects of performance pay for British schools ([Atkinson et al. 2004](#)). A set of observational studies ([Lavy 2008](#); [Lavy 2009](#)) use data from an Israeli policy experiment with tournament-based teacher competition for bonuses and find significant gains in student achievements.

Figure 4. Findings by Job Type

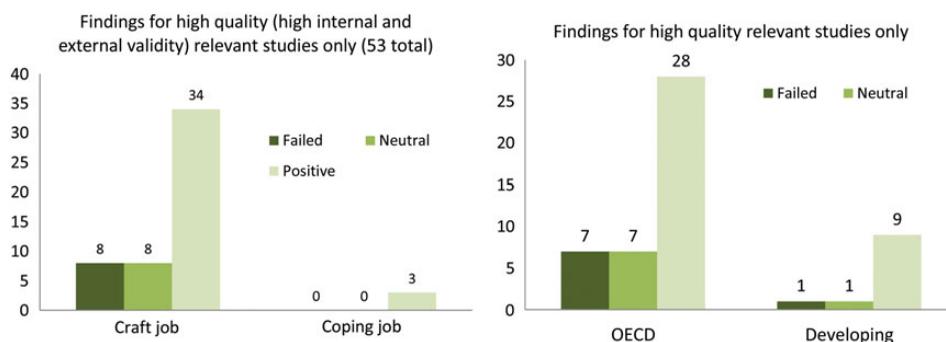


Source: Authors' own findings.

A number of field experiments have evaluated the impact of performance pay for teachers on reducing absenteeism and improving learning outcomes. The findings are generally mixed. [Duflo et al. \(2010\)](#) show that random assignment for monitoring and financial incentives for teachers in rural India led to a strong reduction of teacher absenteeism and increased students' test scores. [Kremer and Chen \(2001\)](#), in contrast, show that subjective monitoring arrangements by an individual in the institutional hierarchy (such as the headmaster of a school) may not work in developing country settings because the monitor might shirk, attempt to avoid confrontation, or collude with the workers. These studies suggest that impersonal, external monitoring by a camera coupled with a clear, credible, and automatic threat of punishment and promise of reward was the key design feature for program success.

A field experiment in 50 Kenyan schools linking teacher salaries to student test scores failed to find lasting effects ([Glewwe et al. 2010](#)). Teacher attendance did not improve, and teachers did not adjust their teaching methods or conduct more preparation sessions. Students in treated schools performed better during the program duration, but these gains did not extend beyond the study period. A field experiment conducted in NYC public schools also failed to find statistically significant effects of team incentives for teachers on student outcomes ([Fryer 2011](#)). A related study that assessed the effects of the NYC group incentive program on classroom activities and teacher turnover and qualification, apart from test scores and teacher effort, similarly found no effects ([Goodman and Turner 2010](#)). A three-year experimental evaluation of the Project on Incentives in Teaching in Metropolitan Nashville schools also found no significant effects of bonus incentives on student test scores ([Springer et al. 2010](#)).

Figure 5. Findings Concerning High-quality Craft and Coping Studies by Sector and Country Context



Source: Authors' own findings.

In contrast, a large-scale field experiment in a representative sample of 300 government-run rural primary schools in India found that bonus pay linked to the mean improvement of student test scores in an independent learning assessment led to a statistically significant and substantively meaningful improvement of student outcomes ([Muralidharan and Sundararaman 2009](#)).

In the health sector, a number of randomized-controlled trials have been implemented to determine the role of performance pay on health worker productivity, patient treatment, and outcomes. Similar to studies on health care relying on observational data, the majority of studies assess these questions in the context of OECD health care systems. [Kouides et al. \(1998\)](#) implemented a randomized-controlled trial, offering a randomly selected set of primary care physicians financial incentives based on influenza immunization rates of the elderly as part of a Medicare demonstration project. Doctors in the treatment group performed more immunizations. [Hillman et al. \(1998\)](#) and [Hillman et al. \(1999\)](#) used two RCT designs to incentivize cancer screenings for women aged 50 and above and pediatric immunizations, respectively. In both studies, the authors documented no significant difference between the treatment and control groups. Similarly, an RCT implemented by [Grady et al. \(1997\)](#) found no clear effects of financial incentives on mammography referrals by primary care physicians.

In contrast, a set of studies ([Fairbrother et al. 1999](#); [Fairbrother et al. 2001](#)), also focusing on pediatric immunizations, found that performance incentives increased immunizations rates by several percentage points compared to the control group. A randomized field trial at the clinic-level found that financial incentives improved the treatment of smoking cessation outcomes ([Roski et al. 2003](#)). Work on performance pay for cognitive services interventions by pharmacists has also demonstrated positive effects ([Christensen et al. 2000](#)).

To our knowledge, the only two available randomized-controlled trials on performance pay in health care in a low-income country are a study by [Basinga et al. \(2010\)](#) in Rwanda and a study by [Singh \(2010\)](#) in India. Basinga et al. used an RCT design to evaluate performance pay in Rwandan primary health care centers. The authors took advantage of a sequenced roll-out of the scheme across Rwandan health care facilities, collecting data on child preventive care and prenatal delivery. To isolate the performance-pay effect from a general increase in resources, comparison facilities received an equivalent increase in their budgets. The study used information from 166 facilities and 2158 households. The authors found large effects on all central outcome measures, with particularly striking effects for services with the highest payoffs and smallest necessary staff effort.

[Singh \(2010\)](#) treated three groups of mothers and the staff providing child care and nutritional advice to them in Chandigarh, India. In one group, the workers received performance pay; in a second group, the workers had no performance pay, but the women they worked with were separately given factual information about

nutrition; and the third group received both treatments. The study found that children's weights improved only in the third group compared to the control group.

It is noteworthy that nearly all of the identified studies on the health care sector focus on fairly narrow types of performance pay and specific, single outcome measures in preventative care, not necessarily overall multidimensional patient treatments and outcomes.

Outside these sectors, [Burgess et al. \(2010\)](#) used an RCT to examine the impact of a pilot team-based incentive scheme introduced in 2002 on the indirect tax assessment and collection agency of the UK government. The tax yield increased for both the treatment teams relative to the control group, with the increases occurring because more time was spent auditing, which resulted in the recovery of greater tax revenue. [Bertelli \(2006\)](#) found that in the Internal Revenue Service in the US, the incentive scheme crowded in intrinsic motivation at the lowest pay levels and crowded it out at the highest levels. A set of studies of performance incentives for agencies with responsibility for training and recruitment found considerable evidence of gaming among the agency staff in the choice of the termination date of the training for the participants ([Asch 1990](#); [Heckman et al. 1997](#); [Courty and Marschke 2004](#)).

Other interesting findings from this group of high quality studies of craft and coping jobs were from [Straberg \(2010\)](#), whose empirical study concerning the perceived impact of performance-related pay in Sweden showed that men were much more likely than women to see the arrangement as fair and reasonable. A laboratory experiment in India assessed teacher efforts when rewards were a function of average student test scores and revealed that poorly designed incentive plans led to the misallocation of teacher effort, which produced an unequal distribution of effort across student groups. However, properly designed incentives could mitigate such behavior ([Jain and Narayan 2011](#)).

Table 3. Findings of High-quality craft and Coping Studies by Sector and Country Context

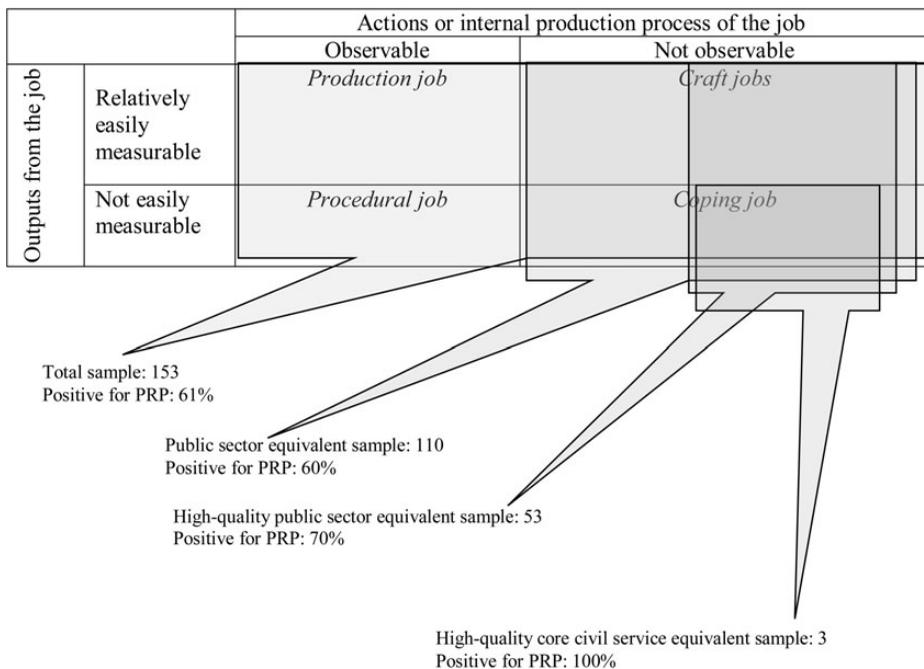
	<i>Craft jobs</i>				<i>Coping jobs</i>	
	<i>Education</i>	<i>Health</i>	<i>Tax</i>	<i>Other</i>	<i>Public</i>	<i>Private</i>
<i>OECD</i>	19	13	2	5	1	2
<i>Positive</i>	12	8	1	4	1	2
<i>Negative or neutral</i>	7	5	1	1	0	0
<i>Developing country</i>	6	4	1	0	0	0
<i>Positive</i>	4	4	1	0		
<i>Negative or neutral</i>	2	0	0	0		
Total	25	17	3	5	1	2

Source: Authors' own findings.

Overall, raising the quality bar still produces a positive result for craft and coping jobs (Figure 5). Thirty-seven of the 53 high-quality studies (4 or 5 on the measure of internal validity and “high” on the measure of external validity) show positive findings. However, it must be emphasized that the number of studies for developing countries is low and that, as noted, the majority of the studies are in the health and education sectors (Table 3).

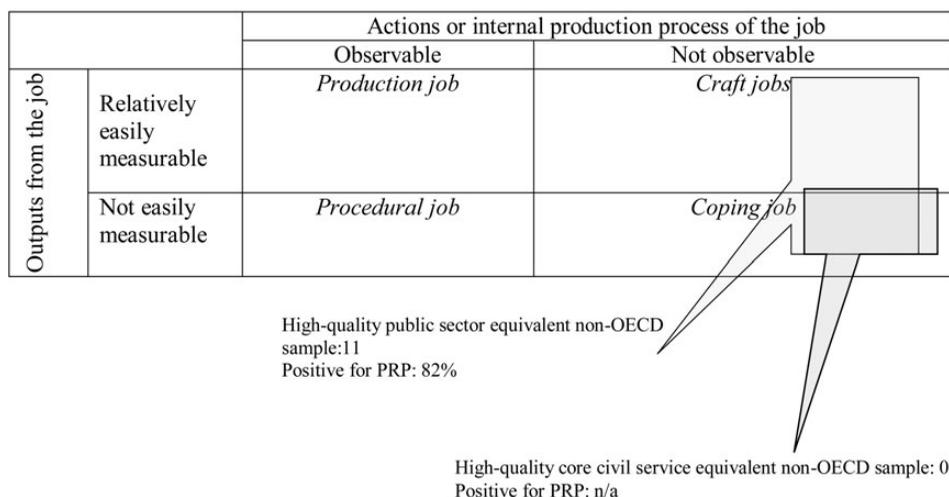
Narrowing further – high-quality studies in relation to coping jobs. Coping jobs are our proxy for jobs within the core civil service. We found very few high-quality studies that addressed these jobs, and none within developing countries (Table 3). Of the three studies that were reviewed, two were on performance pay for managerial positions in the private sector, and only one was on core administrative jobs in the public sector. All of these studies showed positive effects of the performance incentive. Hochberg and Lindsey (2010) reviewed the impact of stock options on company rank-and-file on firm performance (as opposed to the impact of options on top executives, on which there is a large body of literature), finding a positive effect on firm performance. However, although Aboody et al. (2010) similarly showed

Figure 6. Numbers of Studies and their Findings, Filtered by Job Type and Study Quality



Source: Authors' own findings.

Figure 7. Numbers of High-quality Public Sector Equivalent Studies in Non-OECD Settings



Source: Authors' own findings.

that firms that repriced their options had a larger increase in operating income and cash flows compared to non-repricers, they also found that this impact was entirely because of executive stock options. In examining performance-related pay for managers in the UK National Health Service, [Dowling and Richardson \(1997\)](#) found higher staff views on the performance-related pay system in these jobs.

Summary and Implications for Developing Countries

Figure 6 summarizes the major findings of this review with the pool of studies filtered based on job type and quality of the empirical study. Figure 7 summarizes the public sector equivalent studies by country type.

The overall body of evidence paints a generally supportive picture of performance pay in most jobs and in craft and coping jobs that reflect the tasks found in the public sector. This conclusion is strengthened when the sample is narrowed to exclude lower-quality studies and is surprisingly even stronger when the sample is limited to high-quality public sector equivalent jobs in developing countries. However, when focusing narrowly on coping jobs, which more closely resemble those found in the core civil service, the number of studies becomes trivial, and there are none that assist in understanding the relevance of PRP to such jobs outside of the OECD.

At the same time, several observational studies identify problems with unintended consequences, generically subsumed under “gaming” the incentive scheme, which can run counter to the original intentions of the reforms. With the current evidence, however, it remains unclear whether incidents of gaming have a net negative effect in the presence of increased productivity. Furthermore, although explicit incentive schemes certainly increase the opportunity for gaming, standard civil service arrangements have their own unintended incentive effects (i.e., employees will engage in behavior that increases the chances of easy work assignments or promotions). It is simply unknown whether the existing forms of gaming are worse than similar behavior under performance pay. In addition, important cultural differences might exist in the prevalence of gaming performance standards in the public sector between developed and developing countries. Although, to our knowledge, no explicit research on this question exists, work on the prevalence of corruption, behavioral norms, and the effectiveness of anti-corruption efforts suggests that gaming might be more problematic in highly corrupt bureaucracies.

Moving to studies that attempt to fulfill the gold standard of experimental design, the evidence overall again supports the potential utility of performance pay for craft jobs. Comparing various laboratory experiments, the results suggest that explicit performance incentives can work, but the studies employ easily measurable performance indicators and use fairly unrepresentative subject pools. Both concerns should caution policy makers against accepting the results independently of other research. In contrast, similar results have been found across a varied set of experimental settings, test locations, and subject pools, and the overall findings resonate with the observational literature, improving the overall credibility and external validity.

The strongest form of evidence comes from field experimental studies for craft jobs that neatly address concerns of internal and external validity. Here, the evidence is somewhat more mixed. Several studies of teacher incentive programs have found no or transient effects of bonus pay systems in the context of US schools, but in the developing world, the evidence has been more positive. The discrepancy between teacher incentives in the developed and developing world could stem, on the one hand, from the relative magnitude of incentives compared to normal salary or, on the other hand, from higher marginal effects in the education production function in developing countries. Many factors enter the production of education, all of which are likely lacking in many developing country schools. Improving one input aspect (e.g., teacher presence and effort) could have conceivably larger marginal effects than the same input improvement in a developed country school.

What can be concluded with some degree of confidence from this evidence is that if policy-makers are sensitive to design and vigilant about the risks of gaming, then PRP can incentivize workers in both OECD and developing countries to increase effort in craft jobs in which the outputs are readily observable, such as teaching,

health care, and revenue administration. The evidence confounds, at least in the short term, the behavioral economics concern about the crowding out of intrinsic incentives. The key focus for policy-makers should be on the design and implementation modalities of the PRP scheme, such as the size of the incentive, whether it should be an individual or group-based award, the nature of the performance evaluation, and the monitoring regime.

For the core civil service, it is more difficult to draw conclusions for three reasons. First, there is very little research on PRP in these organizational contexts. Second, although some studies have shown that PRP can work even in the most dysfunctional bureaucracies in developing countries, there are too few cases to draw firm conclusions concerning its effectiveness outside of OECD settings for coping jobs. Most glaringly, the role of politicized bureaucracies has not been addressed properly. Finally, although studies do not show that problems with unintended consequences and gaming of incentive schemes result in an overall decline in productivity compared to the counter-factual, few studies follow up performance-related pay effects over a long period of time, leaving the possibility that the positive findings may occur because of Hawthorne Effects (i.e., only as a result of the subjects knowing that they are part of a study and not because of the incentive itself) and that gaming behavior may increase over time as employees become more familiar with the scheme and learn to manipulate it.

Given this problem of the measurability of outputs, PRP for core civil service jobs will largely need to draw on the subjective performance evaluations of individuals. It would be reasonable to expect that performance pay schemes in the core administration are likely to be more successful if they foster better performance dialogue between staff and their managers, based on individual-specific results agreements. Thus, prior investments in improving this performance dialogue and individual performance assessments may increase the likelihood of PRP's positive effects on worker motivation and effort. This dialogue is also more likely to function better under decentralized arrangements for human resource management in which the management of staff and decisions over the performance bonus are the responsibility of line managers and not central human resource agencies or oversight bodies.

Similarly, it can be hypothesized that PRP may have a role in coping jobs through complementing other management reforms, such as results-based management or performance budgeting, which attempt to inculcate a goal orientation in an organization. While the measurability of outputs will always be a challenge by the very nature of these jobs, even these jobs can develop some proxy measures for outputs, such as the level of satisfaction of the users of the services of these jobs, which in the case of policy and regulatory jobs, will be other the government agencies impacted by these policies and regulations. To the extent that reasonable proxy measures are developed, then the positive results from the studies from craft jobs may be more generalizable.

PRP's contribution to agency level productivity through the sorting channel — attracting higher-quality workers who are likely to do better in this scheme — has already been noted, and there is some fairly robust evidence of this in developed country contexts (Booth and Frank 1999; Bandiera et al. 2006; Cadsby et al. 2007). Consequently, it is possible that PRP may have a clearer effect on individual incentives if the sorting effect had longer to work its way through the system. To evaluate this effect, future studies need a longer time frame to assess changes in recruitment and workplace behavior.

Finally, this literature review focuses almost entirely on the individual incentive effects of PRP because this has been the emphasis of the bulk of the studies. It should be noted that there are, at least in the policy literature, potential agency-level and public sector-wide effects of PRP, such as a change in performance culture or fiscal sustainability, that should be explored in future studies.

Notes

1. This paper is a revised version of World Bank Policy Research Working Paper 6043 (Hasnain et al., 2012). The paper benefited from several suggestions and comments, and the authors would like to particularly thank Mike Stevens, Willy McCourt, Mariano Lafuente, Gary Reid, and Svetlana Proskurovska. The views expressed are those of the authors and do not necessarily reflect those of the World Bank or its affiliated organizations.

2. World Bank (zhasnain@worldbank.org).

3. World Bank (nmanning@worldbank.org).

4. The Ohio State University (pierskalla.4@osu.edu).

5. The list of studies reviewed and classified by type of job and quality is available from the authors upon request.

6. Wilson originally used this framework to classify organizations and not jobs, with the implicit assumption that organizations were homogenous in the tasks that they performed.

7. Inevitably, there is some subjectivity in the classification of studies. Studies were rated as *positive* if there was general evidence of the basic functionality of incentive schemes, even if the additional results qualified the effect; for example, studies on the crowding out of intrinsic motivation generally still find positive effects of explicit incentives.

8. Interpreted as purely theoretical papers or studies with a weak research design (e.g., selection on the dependent variable only, no meaningful variation, and no explicit consideration of counter-factuals).

9. Studies that mostly describe reforms implemented in a small number of cases without comparing to cases without performance pay.

10. Studies based on a small number of cases but with at least an implicit consideration of a counter-factual and the use of some minimal data analysis.

11. Studies with an explicit counter-factual analysis, representative sample of cases with and without treatment, and often explicit use of statistical techniques to limit threats to causal inference.

12. We believe that a minimal level of internal and external validity is necessary to draw reliable conclusions from the evidence presented in a study, especially when policy recommendations are concerned. For that reason, we opted to classify studies as “high” quality only if their analysis was based on quasi-experimental methods and the analyzed sample was somewhat representative of the theoretical population under study.

References

- Aboody, D., N. Johnson, and R. Kasznik. 2010. "Employee Stock Options and Future Firm Performance: Evidence from Option Repricings." *Journal of Accounting and Economics* 50 (1): 74–92.
- Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar. 2009. "Large stakes and big mistakes." *Review of Economic Studies* 76 (2): 451–69.
- Asch, B. J. 1990. *Navy Recruiter Productivity and the Freeman Plan*. Santa Monica: RAND Corporation.
- Atkinson, A., S. Burgess, B. Croxson, and P. Gregg. 2004. "Evaluating the Impact of Performance-related Pay for Teachers in England." Working Paper No.04/113. Bristol, UK, Centre for Market and Public Organisation.
- Bandiera, O., I. Barankay, and I. Rasul. 2006. "Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment." Discussion Paper No. 2062. Bonn, Germany, Institute for the Study of Labor
- Banuri, S., and P. Keefer. 2013. "Intrinsic Motivation, Effort and the Call to Public Service." Policy Research Working Paper 6729. Washington DC. World Bank,
- Barlevy, G., and D. Neal. 2011. "Pay for Percentile" Working Paper 17194. Cambridge, MA, National Bureau for Economic Research.
- Basinga, P., P. J. Gertler, A. Binagwaho, A. L. B. Soucat, J. R. Sturdy, and C. Vermeersch. 2010. "Paying Primary Health Care Centers for Performance in Rwanda." Policy Research Working Paper 5190. World Bank, Washington DC.
- Beer, M., and M. D. Cannon. 2004. "Promise and Peril in Implementing Pay-for-Performance." *Human Resource Management* 43 (1): 3–48.
- Beer, M., and N. Katz. 2003. "Do Incentives Work? The Perceptions of a Worldwide Sample of Senior Executives." *People and Strategy* 26 (3): 30–44.
- Belfield, R., and D. Marsden. 2003. "Performance pay, monitoring environments, and establishment performance." *International Journal of Manpower* 24 (4): 452–89.
- Benabou, R., and J. Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70 (3): 489–520.
- . 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–78.
- Bertelli, A. M. 2006. "Motivation crowding and the federal civil servant: Evidence from the U.S. Internal Revenue Service." *International Public Management Journal* 9 (1): 3–23.
- Booth, A. L., and J. Frank. 1999. "Earnings, Productivity, and Performance-Related Pay." *Journal of Labor Economics* 17 (3): 447–63.
- Burgess, S., C. Propper, M. Ratto, S. von Hinke Kessler Scholder, and E. Tominey. 2010. "Smarter Task Assignment or Greater Effort: What Makes the Difference on Team Performance?" *The Economic Journal* 120 (547): 968–89.
- Burgess, S., and M. Ratto. 2003. "The Role of Incentives in the Public Sector: Issues and Evidence." Working Paper. Bristol, UK, Centre for Market and Public Organisation.
- Cadsby, C. B., F. Song, and F. Tapon. 2007. "Sorting and Incentive Effects of Pay-for-Performance: An Experimental Investigation." *Academy of Management Journal* 50 (2): 387–405.
- Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics and Philosophy* 111 (2): 407–41.
- Campbell, S. M., D. Reeves, E. Kontopantelis, E. Middleton, B. Sibbald, and M. Roland. 2007. "Quality of Primary Care in England with the Introduction of Pay for Performance." *New England Journal of Medicine* 357:181–90.

- Campbell, S. M., M. Roland, E. Middleton, and D. Reeves. 2005. "Improvements in the Quality of Clinical Care in English General Practice: Longitudinal Observational Study." *British Medical Journal* 331: 1121.
- Chalkley, M., C. Tilley, L. Young, D. Bonetti, and J. Clarkson. 2010. "Incentives for Dentists in Public Service: Evidence from a Natural Experiment." *Journal of Public Administration Research and Theory* 20 (suppl 2): 207–23.
- Chirkov, V. I., R. M. Ryan, Y. Kim, and U. Kaplan. 2003. "Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and well-being." *Journal of Personality and Social Psychology* 84 (1): 97–110.
- Christensen, D. B., N. Neil, W. E. Fassett, D. H. Smith, G. Holmes, and A. Stergachis. 2000. "Frequency and characteristics of cognitive services provided in response to a financial incentive." *Journal of the American Pharmaceutical Association* 40 (5): 609–17.
- Clotfelter, C., H. E. Ladd, J. L. Vigdor, and R. A. Diaz. 2004. "Do School Accountability Systems Make It More Difficult for Low-Performing Schools to Attract and Retain High-Quality Teachers?" *Journal of Policy Analysis and Management* 23 (2): 251–71.
- Clotfelter, C., H. E. Ladd, and J. L. Vigdor. 2007. "How and Why Do Teacher Credentials Matter for Student Achievement?" Working Paper 2. Washington DC, National Center for Analysis of Longitudinal Data in Educational Research.
- Clotfelter, C., E. Glennie, H. Ladd, and J. Vigdor. 2008. "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina." *Journal of Public Economics* 92 (5-6): 1352–70.
- Condly, S., R. Clark, and H. D. Stolovitch. 2003. "The Effects of Incentives on Workplace Performance: A Meta-analytic Review of Research Studies." *Performance Improvement Quarterly* 16 (3): 46–63.
- Courty, P., C. Heinrich, and G. Marschke. 2005. "Setting the Standard in Performance Measurement Systems." *International Public Management Journal* 8 (3): 1–27.
- Courty, P., and G. Marschke. 2003. "Dynamics of Performance-Measurement Systems." *Oxford Review of Economic Policy* 19 (2): 268–84.
- _____. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics* 22 (1): 23–56.
- Dee, T. S., and B. J. Keys. 2004. "Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment." *Journal of Policy Analysis and Management* 23 (3): 471–88.
- Delfgaauw, J., and R. Dur. 2008. "Incentives and Worker's Motivation in the Public Sector." *The Economic Journal* 118 (525): 171–91.
- Dixit, A. 1999. "Incentives and Organization in the Public Sector. An Interpretative Review." *The Journal of Human Resources* 34 (4): 696–727.
- Doran, T., C. Fullwood, H. Gravelle, D. Reeves, E. Kontopantelis, U. Hiroeh, and M. Roland. 2006. "Pay-for-Performance Programs in Family Practices in the United Kingdom." *New England Journal of Medicine* 355:375–84.
- Dowling, B., and R. Richardson. 1997. "Evaluating Performance-related Pay For Managers in the National Health Service." *The International Journal of Human Resource Management* 8 (3): 348–66.
- Duflo, E., R. Hanna, and S. P. Ryan. 2010. "Incentives Work: Getting Teachers to Come to School." Cambridge, Mass., MIT (Department of Economics and J-PAL) and the Kennedy School of Government.
- Eberts, R., K. Hollenbeck, and J. Stone. 2002. "Teacher Performance Incentives and Student Outcomes." *Journal of Human Resources* 37 (4): 913–27.
- Eichler, R., P. Auxila, and J. Pollock. 2001. "Performance-Based Payment to Improve the Impact of Health Services: Evidence from Haiti." *World Bank Institute Online Journal* (April 2001).

- Eldridge, C., and N. Palmer. 2009. "Performance-based payment: some reflections on the discourse, evidence and unanswered questions." *Health Policy and Planning* 24 (3): 160–6.
- Fairbrother, G., K. L. Hanson, S. Friedman, and G. C. Butts. 1999. "The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates." *American Journal of Public Health* 89 (2): 171–5.
- Fairbrother, G., M. J. Siegel, S. Friedman, P. D. Kory, and G. C. Butts. 2001. "Impact of Financial Incentives on Documented Immunization Rates in the Inner City: Results of a Randomized Controlled Trial." *Ambulatory Pediatrics* 1 (4): 206–12.
- Fehr, E., and K. M. Schmidt. 2004. "Fairness and Incentive in a Multi-Task Principal-Agent Model." *Scandinavian Journal of Economics* 106 (3): 453–74.
- Figlio, D. N., and J. Winicki. 2005. "Food for thought: the effects of school accountability plans on school nutrition." *Journal of Public Economics* 89 (2-3): 381–94.
- Frey, B. S., and M. Osterloh. 1999. *Pay for Performance - Immer Empfehlenswert?* Münster, Germany, Zeitschrift für Führung und Organisation.
- Fryer, R. G. 2011. "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools." NBER Working Paper 16850. Washington DC, National Bureau for Economic Research.
- Glewwe, P., N. Ilias, and M. Kremer. 2010. "Teacher Incentives." *American Economic Journal: Applied Economics* 2 (3): 205–27.
- Gneezy, U., and A. Rustichini. 2000. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics* 115 (3): 791–810.
- Goodman, S., and L. Turner. 2010. "Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program." Working Paper. PEPG Conference "Merit Pay: Will It Work? Is It Politically Viable?". Harvard Kennedy School, June 3-4.
- Grady, K., J. Lemkau, N. R. Lee, and C. Caddell. 1997. "Enhancing Mammography Referral in Primary Care." *Preventive Medicine* 26 (6): 791–800.
- Hasnain, Z., N. Manning, and J. H. Pierskalla. 2012. "Performance-related Pay in the Public Sector: A Review of Theory and Evidence." World Bank Policy Research Working Paper 6043. Washington DC, World Bank.
- Heckman, J., C. Heinrich, and J. Smith. 1997. "Assessing the Performance of Performance Standards in Public Bureaucracies." *The American Economic Review* 87 (2): 389–95.
- Hillman, A., M. Pauly, K. Kerman, and C. R. Martinek. 1991. "HMO manager's views on financial incentives and quality." *Health Affairs* 10 (4): 207–19.
- Hillman, A., K. Ripley, N. Goldfarb, I. Nuamah, J. Weiner, and E. Lusk. 1998. "Physician Financial Incentives and Feedback: Failure to Increase Cancer Screening in Medicaid Managed Care." *American Journal of Public Health* 88 (11): 1699–701.
- Hillman, A., K. Ripley, N. Goldfarb, J. Weiner, I. Nuamah, and E. Lusk. 1999. "The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care." *Pediatrics* 104 (4 pt 1): 931–5.
- Hochberg, Y. V., and L. Lindsey. 2010. "Incentives, Targeting and Firm Performance: An Analysis of Non-Executive Stock Options." *Review of Financial Studies* 23 (11): 4148–86.
- Holmstrom, B., and P. Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics & Organization* 7:24–52.
- Evaluation Group, Independent. 2008. *Public Sector Reform: What works and Why?* Washington DC: World Bank.
- Jack, W. 2003. "Contracting for health services: an evaluation of recent reforms in Nicaragua." *Health Policy and Planning* 18 (2): 195–204.

- Jacob, B. A. 2005. "Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89 (5-6): 761–96.
- Jacob, B. A., and S. D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118 (3): 843–77.
- Jain, T., and T. Narayan. 2011. "Incentive to discriminate? An experimental investigation of teacher incentives in India." Working Paper, Indian School of Business.
- Jenkins, G. D., A. Mitra, N. Gupta, and J. D. Shaw. 1998. "Are Financial Incentives Related to Performance? A Meta-analytic Review of Empirical Research." *Journal of Applied Psychology* 83 (5): 777–87.
- Kahn, C. M., E. C. De Silva, and J. P. Ziliak. 2001. "Performance-Based Wages in Tax Collection: The Brazilian Tax Collection Reform and Its Effects." *The Economic Journal* 111 (468): 188–205.
- Kellough, E. J., and H. Lu. 1993. "The Paradox of Merit Pay in the Public Sector: Persistence of a Problematic Procedure." *Review of Public Personnel Administration* 13 (2): 45–64.
- Kerr, S. 1975. "On the Folly of Rewarding A, While Hoping for B." *The Academy of Management Journal* 18 (4): 769–83.
- Ketelaar, A., N. Manning, and E. Turkisch. 2007. "Performance-based arrangements for senior civil servants OECD and other country experiences." OECD Governance Working Paper, Paris.
- Kiragu, K., and R. Mukandala. 2003. "Public Sector pay reform - Tactics Sequencing And Politics In Developing Countries: Lessons From Sub-Saharan Africa." Dar es Salaam: Pricewaterhousecoopers and University of Dar es Salaam.
- Kouides, R., N. Bennett, B. Lewis, J. Cappuccio, W. Barker, and M. LaForce. 1998. "Performance-Based Physician Reimbursement and Influenza Rates in the Elderly." *American Journal of Preventive Medicine* 14 (2): 89–95.
- Kremer, M., and D. Chen. 2001. "An Interim Report on a Teacher Attendance Incentive Program in Kenya." Mimeo. Cambridge, MA: Harvard University.
- Kreps, D. M. 1997. "Intrinsic Motivation and Extrinsic Incentives." *The American Economic Review* 87 (2): 359–64.
- Lavy, V. 2008. "Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-Based Pay Tournaments Among Teachers." NBER Working Paper No. 14338. Washington DC, National Bureau for Economic Research.
- _____. 2009. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics." *American Economic Review* 99 (5): 1979–2011.
- Luthans, F. 1973. *Organizational Behavior*. New York: McGraw-Hill.
- Marsden, D. 2004. "The role of performance-related pay in renegotiating the "effort bargain": the case of the British public service." *Industrial and Labor Relations Review* 57 (3): 350–70.
- _____. 2009. "The paradox of performance related pay systems: why do we keep adopting them in the face of evidence that they fail to motivate?" London, Centre for Economic Performance, London School of Economics.
- McNamara, P. 2005. "Quality-based payment: six case examples." *International Journal for Quality in Health Care* 17 (4): 357–63.
- Meessen, B., J. Kashala, and L. Musango. 2007. "Output-based payment to boost staff productivity in public health centres: contracting in Kabutare district, Rwanda." *Bulletin of the World Health Organization* 85 (2): 108–15.
- Muralidharan, K., and V. Sundararaman. 2009. "Teacher Performance Pay: Experimental Evidence From India." NBER Working Paper 15323. Washington DC, National Bureau for Economic Research.

- Murnane, R. J., and D. K. Cohen. 1986. "Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and Few Survive." *Harvard Educational Review* 56 (1): 1–17.
- Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review* 92 (4): 850–73.
- Neal, D. 2011. "The Design of Performance Pay in Education." NBER Working Paper 16710. Washington DC, National Bureau for Economic Research.
- Niemiec, C. P., R. M. Ryan, and E. L. Deci. 2009. "The Path Taken: Consequences of Attaining Intrinsic and Extrinsic Aspirations in Post-College Life." *Journal of Research in Personality* 73 (3): 291–306.
- OECD. 1993. *Pay Flexibility in the Public Sector*. Paris: OECD.
- . 1996. *Pay Reform in the Public Service: Initial impact on pay dispersion in Australia, Sweden, and the United Kingdom*. Paris: OECD PUMA.
- . 1997. *Trends in Public Sector Pay in OECD countries*. Paris: OECD/PUMA.
- . 2004. "Trends in Human Resources Management Policies in OECD Countries. An Analysis of the Results of the OECD Survey on Strategic Human Resources." Paper presented to the Human Resources Management Working Party. Paris: OECD.
- . 2005a. *Modernising Government: The Way Forward*. Paris: OECD.
- . 2005b. *Performance-related Pay Policies for Government Employees*. Paris: OECD.
- . 2008. *The state of the public service*. Paris: OECD.
- . 2009. *Government at a Glance*. Paris: OECD.
- . 2011. *Government at a Glance*. Paris: OECD.
- Perry, J. L., T. A. Engbers, and S. Y. Jun. 2009. "Back to the Future? Performance-Related Pay, Empirical Research and the Perils of Persistence." *Public Administration Review* 69 (1): 39–51.
- Perry, J. L., and A. Hondeghem, eds. 2008. *Motivation In Public Management: The Call Of Public Service*. Oxford: Oxford University Press.
- Perry, J. L., D. Mesch, and L. Paarlberg. 2006. "Motivating employees in a new governance era: the performance paradigm revisited." *Public Administration Review* 66 (4): 505–14.
- Petersen, L. A., L. D. Woodard, T. Urech, C. Daw, and S. Sookanan. 2006. "Does pay-for-performance improve the quality of health care?" *Annals of Internal Medicine* 145 (4): 265–72.
- Pink, D. H. 2009. *Drive: the surprising truth about what motivates us*. New York: Riverhead.
- Porter, L. W., and , III, E. Lawler. 1968. *Managerial Attitudes and Performance*. Homewood, IL: Dorsey Press.
- Prendergast, C. 1998. "What Happens within Firms? A Survey of Empirical Evidence on Compensation Policies." NBER Working Paper. Washington DC, National Bureau for Economic Research.
- Prendergast, C. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63.
- Propper, C., and D. Wilson. 2003. "The Use and Usefulness of Performance Measures in the Public Sector." CMPO Working Paper Series No. 03/073. Bristol, UK, The Centre For Market And Public Organisation.
- Rexed, K., C. Moll, N. Manning, and J. Allain. 2007. "Governance Of Decentralised Pay Setting In Selected Oecd Countries." OECD Working Papers on Public Governance, 2007/3. Paris, OECD.
- Roski, J., R. Jeddelloh, L. An, H. Lando, P. Hannan, C. Hall, and S. H. Zu. 2003. "The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines." *Preventive Medicine* 36 (3): 291–9.
- Ryan, R. M., and E. L. Deci. 2000. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." *American Psychologist* 55 (1): 68–78.

- Sauermann, H., and W. M. Cohen. 2008. "What Makes Them Tick? Employee Motives and Firm Innovation." NBER Working Paper No. 14443. Cambridge, MA, NBER.
- Shen, Y. 2003. "Selection Incentives in a Performance-Based Contracting System." *Health Services Research* 38 (2): 535–52.
- Singh, P. 2010. *Performance pay and information: reducing child malnutrition in urban slums*. London: London School of Economics.
- Skinner, B. F. 1969. *Contingencies of Reinforcement*. New York: Appleton-Century-Crofts.
- Soeters, R., C. Habineza, and P. B. Peerenboom. 2006. "Performance-based financing and changing the district health system: experience from Rwanda." *Bulletin of the World Health Organization* 84 (11): 884–9.
- Springer, M. G., D. Ballou, L. Hamilton, V. Le, J. R. Lockwood, D. McCaffrey, M. Pepper, and B. Stecher. 2010. "Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching." Working Paper. Nashville, TE, National Center on Performance Incentives at Vanderbilt University.
- Stajkovic, A. D., and F. Luthans. 2003. "Behavioral management and task performance in organizations: conceptual background, meta - analysis, and test of alternative models." *Personnel Psychology* 56 (1): 155–94.
- Steel, N., S. Maisey, A. Clark, R. Fleetcroft, and A. Howe. 2007. "Quality of clinical primary care and targeted incentive payments: an observational study." *British Journal of General Practice* 57 (539): 449–54.
- Straberg, T. 2010. "Employee perspectives on individualised pay: attitudes and fairness perceptions." PhD dissertation, University of Stockholm.
- Vaghela, P., M. Ashworth, P. Schofield, and M. C. Gulliford. 2009. "Population intermediate outcomes of diabetes under pay-for-performance incentives in England from 2004 to 2008." *Diabetes Care* 32 (3): 427–9.
- Vroom, V. H. 1964. *Work and Motivation*. Hoboken: Wiley.
- Weibel, A., K. Rost, and M. Osterloh. 2009. "Pay for Performance in the Public Sector - Benefits and (Hidden) Costs." *Journal of Public Administration Research and Theory* 20 (2): 387–412.
- Wilson, J. Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- Witter, S., T. Zulfiqur, S. Javeed, A. Khan, and A. Bari. 2011. "Paying health workers for performance in Battagram district." *Human Resources for Health* 9: 23.
- World Bank. 1999. *Civil Service Reform: a Review of World Bank Assistance: Report No. 19211*. Washington DC: OED, World Bank.
- . 2001. *Salary Supplements and Bonuses in Revenue Departments (Final report)*. Washington DC: World Bank.

THE WORLD BANK

1818 H Street NW
Washington, DC 20433, USA
World Wide Web: <http://www.worldbank.org/>
E-mail: researchobserver@worldbank.org

OXFORD  OPEN



SFI[®]

Certified Sourcing

www.sfiprogram.org
SFI-00453

ISBN 978-0-19-871357-9



9 780198 713579