

Public Disclosure Authorized  
August 2013

Public Disclosure Authorized  
The World Bank Research Observer

Public Disclosure Authorized  
Volume 28, Issue 2

Public Disclosure Authorized  
OXFORD

THE WORLD BANK

# Research Observer

Volume 28 • Number 2 • August 2013

OXFORD  
UNIVERSITY PRESS

## INSIDE

Ravallion on How Long Will It  
Take to Lift One Billion People  
Out of Poverty?

Gersovitz and Kriger on  
What Is a Civil War?

Servén and Nguyen on  
Global Imbalances

Ferreira and Lugo on  
Multidimensional  
Poverty Analysis

Acharya and others on  
The Impact of Health Insurance  
Schemes for the Informal  
Sector

Baez and others on Rural  
Households in a  
Changing Climate

ISSN 0257-3032 (PRINT)  
ISSN 1564-6971 (ONLINE)  
[www.wbro.oxfordjournals.org](http://www.wbro.oxfordjournals.org)

THE WORLD BANK

# Research Observer

---

EDITOR

Emmanuel Jimenez, World Bank

CO-EDITOR

Luis Servén, World Bank

EDITORIAL BOARD

Harold Alderman, World Bank

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, World Bank

*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](http://www.wbro.oxfordjournals.org), the World Bank at [www.worldbank.org](http://www.worldbank.org), and Oxford University Press at [www.oxfordjournals.org](http://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 28, 2 issues, 2013)

#### Academic libraries

Print edition and site-wide online access: US\$206/£137/€206

Print edition only: US\$189/£125/€189

Site-wide online access only: US\$172/£114/€172

#### Corporate

Print edition and site-wide online access: US\$312/£207/€312

Print edition only: US\$286/£190/€286

Site-wide online access only: US\$260/£172/€260

#### Personal

Print edition and individual online access: US\$58/£39/€58

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](http://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](mailto: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](mailto: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: [jnlororders@oup.com](mailto:jnlororders@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](mailto: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](http://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](http://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](mailto: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** © 2013 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

How Long Will It Take to Lift One Billion People Out of Poverty? Martin Ravallion	139
What Is a Civil War? A Critical Review of Its Definition and (Econometric) Consequences Mark Gersovitz and Norma Kriger	159
Global Imbalances: Origins and Prospects Luis Servén and Ha Nguyen	191
Multidimensional Poverty Analysis: Looking for a Middle Ground Francisco H. G. Ferreira and Maria Ana Lugo	220
The Impact of Health Insurance Schemes for the Informal Sector in Low- and Middle-Income Countries: A Systematic Review Arnab Acharya, Sukumar Vellakkal, Fiona Taylor, Edoardo Masset, Ambika Satija, Margaret Burke, and Shah Ebrahim	236
Rural Households in a Changing Climate Javier E. Baez, Dorothy Kronick, and Andrew D. Mason	267

---

# Editors' Note

---

Emmanuel Jimenez • Ana L. Revenga, Symposium Co-editors

---

The lives of girls and women in developing countries have improved dramatically over the past quarter of a century. Female life expectancy at birth has increased dramatically to reach 71 years globally in 2007. Literacy rates have risen and in a third of developing countries there are more girls in secondary school than boys. Most strikingly, in more than half the developing countries more women attend university than men. Women have entered the labor force in massive numbers and now account for 40% of the global work force. Moreover, change has come much more quickly than it did historically in the rich countries.

In some areas, however, progress toward gender equality has been limited or very slow. Too many girls and women are still dying in childhood and in the reproductive ages. Women still fall behind in earnings and productivity, and in the strength of their voices in society. These patterns of progress and stagnation matter in shaping development outcomes and were discussed in the World Bank's flagship report, *The World Development Report 2012: Gender Equality and Development (WDR2012)*.

Comprehensive as that report was, some of the rich material commissioned by or otherwise used to inform the report could not all be reflected in it. This issue of the *World Bank Research Observer* selects papers that present insights that go beyond what could be covered in the *WDR 2012*. Bandiera and Natraj critically review the evidence on the influence of gender inequality on economic growth. Croppenstedt, Goldstein and Rosas Raffo take a sectoral cut on the issue through the lens of rural development. Two papers reflect on how individual behavior might affect broader gender inequality: Farre's paper examines the role of men in shaping the advancement of women and Doss reviews the work on household bargaining in allocating resources. Finally, Buvinic and colleagues delve deeply into the interaction between gender equality and violent conflict.

These papers were all subjected to our usual review process. We think that they have not only been useful inputs into the WDR; they also stand on their own as contributions to a fascinating literature.

---

# How Long Will It Take to Lift One Billion People Out of Poverty?

---

Martin Ravallion<sup>1</sup>

---

*Alternative scenarios are considered for reducing by one billion the number of people surviving on less than \$1.25 a day. The low-case, “pessimistic” path to that goal envisages the developing world outside China returning to the slower pace of economic growth and poverty reduction of the 1980s and 1990s, but with China maintaining its progress. This path would take 50 years or more to lift one billion people out of poverty. A more optimistic path is identified that would maintain the developing world’s (impressive) progress against absolute poverty since the turn of the century. This path would lift one billion people out of poverty by 2025–30. The optimistic path is consistent with both linear projections of the time-series data and nonlinear simulations of inequality-neutral growth for the developing world as a whole. JEL codes: I32, O15*

At the time of this writing (in 2012), the available data indicate that 1.2 billion people in the world live in poverty, as judged by the frugal standards used for defining poverty in the world’s poorest countries—giving an international poverty line of \$1.25 a day (at 2005 purchasing power parity). This paper draws on recent research on global poverty to assess how long it might take to lift one billion people out of such extreme poverty. The paper identifies both “pessimistic” and “optimistic” paths to that goal.

International development institutions and almost all developing countries track poverty measures, and the results are keenly watched. The development literature has devoted less analytic attention to the important task of benchmarking performance.<sup>2</sup> When can we say that performance against poverty is “good” or “bad”?

Setting poverty-reduction goals can help motivate their achievement. The Millennium Development Goals (MDGs) were clearly conceived with that intention. As Hume (2009, p. 4) states, the MDGs aimed to “stretch ambitions and mobilize political commitment and public support.”<sup>3</sup> The first MDG was to halve the developing world’s 1990 “\$1-a-day” poverty rate by 2015. Using the \$1.25-a-day poverty line in 2005 prices, the first MDG was attained in 2010, a full five years ahead of the goal (Chen and Ravallion 2012). Even so, that important achievement leaves over one billion people living in extreme poverty.

To motivate extra effort, the goals must not be either too easy or too hard. Past experience can provide useful data for assessing proposed benchmarks for future performance. At the time that the first MDG was set, there were very few time-series observations to draw on. Fortunately, the situation has changed dramatically. Although the data are still far from ideal, we now have a firmer empirical basis for considering alternative goals that are within the range of experience.

The benchmarks proposed here rest on explicit scenarios for future economic growth and distributional change, informed by knowledge of recent past performance in reducing poverty and current expectations about economic growth prospects across the developing world. Unavoidably, there is a measure of conjecture in the establishment of any targets. But it is hoped that this paper’s analysis will help guide assessments of our performance toward eliminating extreme poverty over the coming decades and help mobilize future efforts toward that goal.

The first of the two main benchmark trajectories proposed here assumes that the faster pace of poverty reduction that we have observed in the developing world since the 1990s will not be maintained. Rather, a series of economic and policy reversals would require the developing world outside China to return to the slower pace of poverty reduction observed in the 1980s and 1990s, although with China staying on its track. On this trajectory, it would take 50 years to lift one billion people out of extreme poverty.

One can speculate on many more optimistic paths. At one extreme, one might imagine that poverty could be eliminated tomorrow, and for ever after, by using perfectly targeted transfers to close the aggregate poverty gap—bringing everyone to (say) \$1.25 a day. However, such perfect targeting has never happened, and it appears unlikely to ever happen (in part because of its likely incentive effects, which would create 100 percent marginal tax rates on poor people). Arguably, such a goal would be too optimistic to mobilize serious effort.

Instead, the focus here is on an optimistic trajectory that is within the range of experience rather than well outside that range. That trajectory assumes that the developing world’s recent success in reducing the incidence of extreme poverty will be maintained going forward. It assumes that China continues on its path toward eliminating extreme poverty and that the developing world outside China

remains on the encouraging new path of more rapid poverty reduction that we have observed since the turn of the century.

It is reasonable to expect that this optimistic trajectory will lift one billion people out of extreme poverty by 2025–30. That would require the developing world to be successful across multiple dimensions of poverty reduction, including fostering the conditions for continued, reasonably rapid, economic growth, avoiding major crises (financial and agro-climatic) and assuring that poor people are able to participate fully in that growth, which will in turn require that they have access to schooling, health care, labor-market opportunities and financial resources when needed.

After reviewing what we know about recent economic growth, distributional change, and poverty reduction in the developing world, the paper discusses the proposed benchmark paths for poverty reduction. Two main methodological approaches are used, one based on time-series evidence and one based on simulations. The final section concludes.

## Recent Trends in Growth, Redistribution, and Poverty Reduction

The measure of poverty obtained for a given distribution of consumption (or income) depends on the mean of that distribution (relative to the poverty line) and the extent of “inequality” in the distribution.<sup>4</sup> The following discussion will focus first on the mean and then turn to inequality.

Prior to about 2000, poor countries were not typically showing much sign that they would eventually catch up to rich countries in terms of mean income; the convergence process was weak or absent. Indeed, this was often seen as a “stylized fact” of economic development, and it motivated various theoretical explanations centered on the idea of a “poverty trap;” see, for example, [Azariadis \(2006\)](#). Under certain conditions, these models predict that positive economic growth at a low initial capital stock will not be sustainable; rather, dynamic economic, and possibly political, forces will pull the economy back to its initial (low) level in due course. An example of the (potentially many) ways that a poverty trap can arise is when low life expectancies in poor countries dissuade saving and, hence, investment, which in turn keeps life expectancy low. Under certain conditions, this type of “vicious cycle” entails that a small amount of extra investment will not bring a lasting output gain, but a sufficiently large injection of capital will do so—moving the economy out of its trap into a “virtuous cycle” of progress toward a sustainably higher long-run level of income. Such theories have prompted policy arguments in favor of a substantial increase in development aid to poor countries (as in [Sachs, 2005](#)).

Turning to recent evidence, we have seen a marked acceleration in the developing world's economic growth since the turn of the century. The developing world as a whole has been maintaining a growth rate for GDP of approximately 6 percent over most of the last decade, although it dipped substantially (and temporarily) in 2008–09 because of the global financial crisis. It is noteworthy that this rate is a full 2 percentage points higher than the average economic growth rate of approximately 4 percent from the 1960s through to the mid-1990s. This growth was fuelled (in part) by substantially greater investment in developing countries since 2000 (from both domestic and external sources), and this was not just in China and India (Lim, 2012).

Has this growth come with higher average household living standards? Macro-level economic growth, as measured in the national accounts (NAS), does not automatically translate into similar growth in average household living standards, as measured by the mean for the distribution of household consumption or income on which poverty measures are based. There are invariably gaps between NAS aggregates and the grossed-up consumption or income aggregates from the household surveys used to measure poverty. There are a number of reasons for these gaps. In practice, the way that the NAS are constructed means that there is nothing exactly corresponding to household consumption as measured in surveys, so full agreement should not be expected. Measurement errors in both sources also play a role, as do differences in accounting periods and sampling problems. A likely source of discrepancies between the two data sources is underreporting of incomes or consumptions in sample surveys or selective compliance in the randomized assignments that are used in implementing the surveys. The rich will undoubtedly have a stronger incentive than the poor to underreport their incomes or consumption and will be less likely to be available for interviews. Indeed, in one assessment (using data for the United States), selective compliance—in which the rich are less likely to participate in surveys than the poor—resulted in a sizeable underestimation of inequality using the unadjusted sample survey data but made little difference for poverty measures (Korinek et al. 2006).

Compiling evidence on how average living standards have been evolving over time in the developing world requires a large number of household surveys, which come irregularly over time. Thus, the calculations are slightly more complicated than when using annual NAS data. Drawing on some 900 household surveys for 125 countries, Chen and Ravallion (2012) provide estimates of poverty measures for “reference years” spanning 1981–2008 at three yearly intervals, which have been updated to include 2010 for this paper. Given the irregular and unsynchronized spacing of surveys, Chen and Ravallion (2010, 2012) use an interpolation method that employs growth rates in private consumption from the NAS for those dates when surveys are unavailable. The survey means are a mixture of consumption and income, depending on what is available in the

surveys. For the Chen-Ravallion poverty measures, two-thirds of the surveys use consumption, which is preferred to income as a welfare metric. (Income is only used when consumption is unavailable.)

On calculating the overall survey means for the Chen-Ravallion reference years, one finds that the trend rate of growth over 1981–2010 was 1.9 percent per annum with a standard error of 0.3 percent. (The trend rate of economic growth is defined here as the regression coefficient of the log mean on time.)

However, as was seen in the NAS data, there was a marked break in the trajectory around the turn of the century. The trend growth rate prior to 1999 was 0.9 percent per annum (standard error = 0.1), while it was 4.3 percent (0.2 percent) from 1999 onward.<sup>5</sup> (The difference in trends is statistically significant:  $t = 3.32$ ,  $p = 0.009$ .) It is especially notable that this break reflects the higher economic growth rate of the developing world outside China, which rose from 0.6 percent (0.1 percent) prior to 1999 to 3.8 percent (0.3 percent) from 1999 onward.

The key point to emerge from these observations—drawing on both NAS and surveys—is that the output and living standards of the developing world as a whole have been growing at faster rates than for the rich world since the turn of the century, reversing a prior pattern of little or no progress toward economic convergence. How this happened is not yet clear. Possibly, the (private and public) capital flows to developing countries since 2000 have been sufficient to get out of past poverty traps. Or the conditions required for the existence of poverty traps may not in fact hold in practice (as argued by [Kraay and Raddatz, 2007](#)), or they may only hold for certain countries or for pockets of poverty within countries, but not in the aggregate. Then, it was just a matter of time, aided by more supportive policy environments (including political stability), before the catching up process would be revealed. Less optimistically, given the uncertainties about the dynamics of the processes involved, one cannot yet rule out the possibility that the forces creating the earlier traps are still there and will reemerge in due course in the form of setbacks to continuing economic growth.

Economic forecasters have never seemed particularly good at predicting reversals or even significant slowdowns. It is possibly not surprising then that current expectations amongst mainstream forecasters are that this stronger new convergence process since the turn of the century will continue. At the time of this writing, the World Bank's economic growth projections assume that a full-blown Euro crisis will be avoided and that the annual growth rate of GDP for the developing world will continue to be 6 percent in the coming few years (having fallen slightly this year) ([World Bank 2012b](#)). Given current population projections, a 6 percent growth rate in the GDP over the coming decade would represent a 4.9 percent rate for GDP per capita. (The compound rate of population growth over 2012–17 and 2012–22 is 1.1 percent per annum.) Although the recent growth has not been even across all regions, the three regions that account for the bulk

of absolute poverty—East Asia, South Asia, and sub-Saharan Africa—have observed strong GDP growth rates in recent years: approximately 8 percent in East Asia, 7 percent in South Asia, and 5 percent in sub-Saharan Africa. Current expectations are that these rates will be maintained, although most forecasters acknowledge that serious risks persist that further crises emanating in the “rich world” will spill over significantly into economic growth and poverty reduction in the developing world; see, for example, the discussions in [World Bank \(2012a, b\)](#).

Let us turn now to the other key variable determining the extent of poverty, namely, inequality. The concept of “global inequality” that is relevant to the measurement of global poverty pools all residents of all countries and measures the inequality among them as if they were one country. Figure 1 plots one such measure of inequality, the mean log deviation. The mean log deviation is a theoretically sound measure with the useful property of exact decomposability by population subgroups ([Bourguignon 1979](#)). Thus, we can cleanly separate the “between-country” component of total inequality from the “within-country” component. Figure 1 provides this breakdown of total inequality in the developing world.

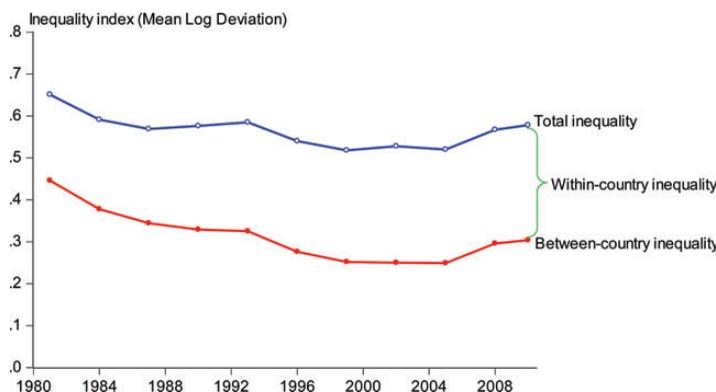
We see that there has been a trend decrease in total inequality; over the period as a whole, there is a small but statistically significant negative trend, at  $-0.002$  per year (standard error = 0.001), compared to a mean log deviation of 0.57. However, the bulk of the decline in overall inequality was in the period until the late 1990s. There is an indication of rising overall inequality since 2005.<sup>6</sup>

The future evolution of overall inequality will be crucial to the trajectories of overall poverty measures. It might be conjectured that higher rates of economic growth will (at least initially) put upward pressure on inequality within low- and

---

**Figure 1.** Inequality in the Developing World

---



Source: Update to [Ravallion and Chen \(2012\)](#).

---

possibly middle-income developing countries, as predicted by the famous Kuznets Hypothesis (Kuznets 1955). However, this conjecture is not consistent with past evidence, which indicates that inequality within growing developing countries falls about as often as it rises (Ravallion 2001; Ferreira and Ravallion 2009). Moreover, a number of high-inequality, economically growing, developing countries have succeeded in attenuating and even reducing inequality. The available evidence leads one to doubt that higher inequality is a necessary “price” of higher economic growth and lower absolute poverty (Ravallion 2005).

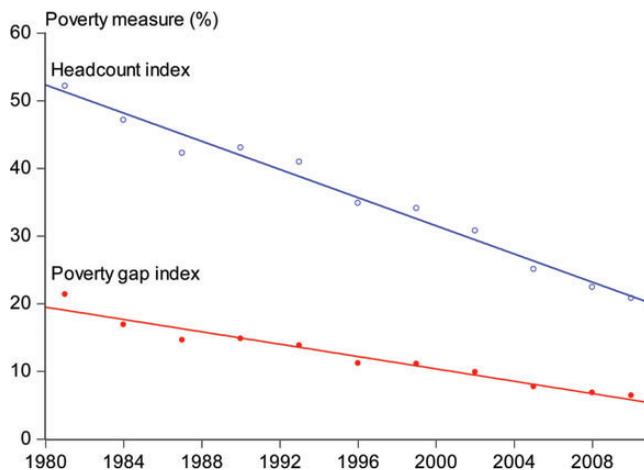
However, as is evident from figure 1, inequality between countries matters more to the evolution of total inequality. Recent economic growth in India and (especially) China has played an important role in the evolution of the between-country component. Given their initially low average incomes and high population weights, economic growth in China and India has been a strong force for global inequality reduction. This is likely to change when these countries reach the overall mean for the developing world. That has not yet happened but will soon happen in the case of China, where mean consumption in 2010 was 95 percent of the mean for the developing world as a whole (\$5.03 per person per day). Continuing improvement in sub-Saharan Africa’s economic growth performance would work in the opposite direction, putting downward pressure on overall inequality.

We have seen that substantial economic growth in the developing world as a whole over the 1980–2010 period has come with falling inequality, though with signs of rising inequality since 2005. What then has been the record for poverty reduction?

Following past practice, the incidence of poverty is measured here by the standards used to define poverty in the poorest countries of the world. On this basis, the World Bank’s current international poverty line is \$1.25 per person per day at 2005 purchasing power parity (Ravallion et al. 2009). The latest (survey-based) estimate of the proportion of the population of the developing world living below \$1.25 a day is 21 percent for 2010 (representing 1.2 billion people), down from 33 percent in 2000 and 43 percent in 1990. Figure 2 plots the full series over time.<sup>7</sup> The overall poverty rate (“headcount index”) has been falling at a robust 1 percentage point per year over 1981–2010, and this was maintained after 2005. The regression coefficient of the poverty rate on the year in figure 2 is  $-1.04$  (standard error = 0.05;  $n = 11$ ). The “\$1.25-a-day” poverty rate for the developing world as a whole in the year of writing (2012) is estimated to be 19 percent (standard error = 0.6 percent), representing 1.1 billion people.

The extent of the linearity in time is a striking feature of the series for the headcount index in figure 2. This linearity is clear from looking at the graph, but it is also confirmed statistically using a standard test for nonlinearity in the function form.<sup>8</sup> Such linearity is not what one would expect if the overall growth rate

**Figure 2.** Poverty Measures for the Developing World, 1981–2010



Note: Both measures use a poverty line of \$1.25 a day at 2005 purchasing power parity.  
Source: Chen and Ravallion (2012) and author's calculations.

in the mean was constant and the elasticity of the poverty rate to the mean also remained constant; then, one would expect the annual rate of decline in the poverty rate to fall over time as the poverty rate fell. However, these conditions have plainly not held. Instead, we have observed higher economic growth rates over time in developing countries and falling overall inequality for the period as a whole, which would tend to increase the elasticity of poverty to economic growth (Ravallion 1997). These forces have been strong enough to keep the annual pace of poverty reduction roughly constant. But that will not happen automatically in the future and rather will require sustained effort. The discussion will return to this point.

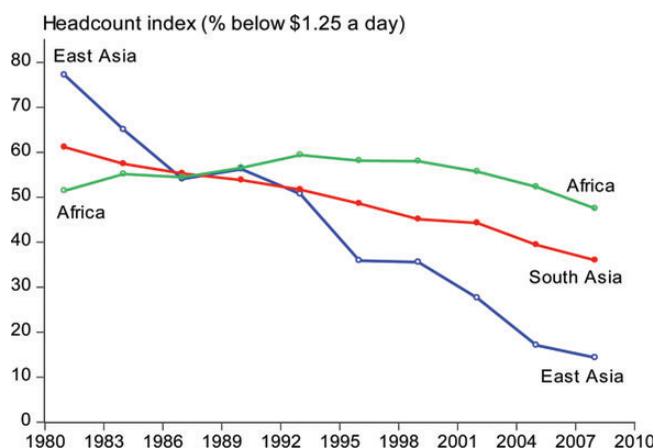
Figure 2 also provides the series for the poverty gap (PG) index, which is the mean gap below the poverty line as a proportion of the line, expressed as a percentage. Equivalently, the PG index is the product of the headcount index and the income gap ratio, given by the difference between the poverty line and the mean consumption or income of those living below the line, expressed as a percentage of the line. The PG index fell from 21 percent to 6 percent over the 1981–2010 period. The regression coefficient of the index on year is  $-0.46$  (standard error =  $0.04$ ;  $n = 11$ ). The income gap ratio also fell, from  $0.41$  to  $0.31$ .

Progress in poverty reduction has been uneven across regions. Table 1 shows the regional breakdown for selected years. Figure 3 shows the series for three regions accounting for 95 percent of those living below \$1.25 a day (in 2008). For East Asia, the rate of poverty reduction has been approximately double the

**Table 1.** Regional Breakdown of Poverty Incidence and Projections for 2015

	1990	1999	2008	2015
<b>Poverty rate</b> (% of the population living below \$1.25 a day)				
East Asia and Pacific	56.2	35.6	14.3	7.7
Europe and Central Asia	1.9	3.8	0.5	0.3
Latin America and Caribbean	12.2	11.9	6.5	5.5
Middle East and North Africa	5.8	5.0	2.7	2.7
South Asia	53.8	45.1	36.0	23.9
sub-Saharan Africa	56.5	58.0	47.5	41.2
Total	43.1	34.1	22.4	16.3
<b>Number of poor</b> (millions living below \$1.25 a day)				
East Asia and Pacific	926.4	655.6	284.4	159.3
Europe and Central Asia	8.9	17.8	2.2	1.4
Latin America and Caribbean	53.4	60.1	36.8	33.6
Middle East and North Africa	13.0	13.6	8.6	9.7
South Asia	617.3	619.5	570.9	418.7
sub-Saharan Africa	289.7	376.8	386.0	397.2
Total	1,908.6	1,743.4	1,289.0	1,019.9

Source: Chen and Ravallion (2010) and World Bank (2012a).

**Figure 3.** Poverty Rates for the Three Poorest Regions

Source: Chen and Ravallion (2012).

average, at slightly over 2 percentage points per year; for South Asia, it has been slightly less than 1 percent, and for sub-Saharan Africa, it has been about zero over the period as a whole. (The coefficients (standard error) are  $-2.24$  percent (0.14) for East Asia,  $-0.88$  percent (0.04) for South Asia and  $-0.09$  percent (0.14) for Africa.)

China's success in reducing absolute poverty has (understandably) attracted much attention. (For further discussion, including discussion of the reasons for this success, see [Ravallion and Chen 2007](#).) However, a less widely appreciated fact is that the developing world outside China has moved to a steeper trajectory of poverty reduction since 2000 (in keeping with the survey means). There have been clear signs of a positive trend emerging in Africa in the 2000s, similar to South Asia. Chen and [Ravallion \(2012\)](#) show that the 2000s have achieved progress in reducing poverty—with both a falling incidence and falling numbers of poor—in all six regions of the developing world (Eastern Europe and Central Asia, Latin America and the Caribbean, and the Middle East and North Africa, in addition to those in figure 3).

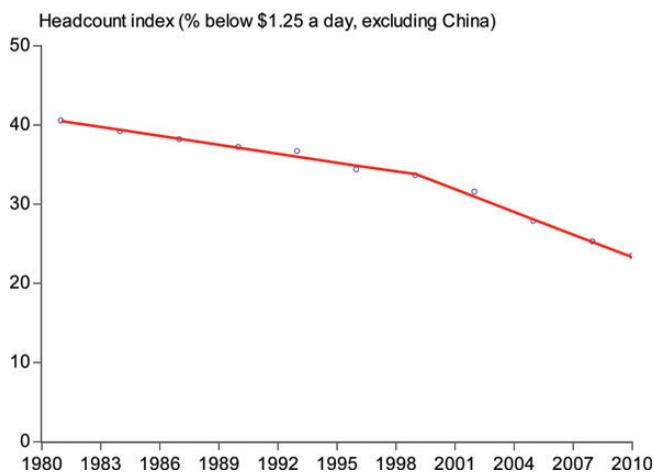
If we focus on the developing world outside China, the rate of poverty reduction (again using the regression coefficient on time and the headcount index for \$1.25 a day) over the 1981–2010 period was  $-0.6$  percentage points per year (standard error = 0.05;  $n = 11$ ). It was China's success that pulled the overall rate up to 1 percentage point per year.

However, it has not yet been widely appreciated that there was a marked positive change in the trajectory for the developing world outside China around the turn of the century, as figure 4 shows. For the period up to the late 1990s, the trend rate of poverty reduction for the developing world outside China was  $-0.4$  percentage points per year (standard error = 0.03;  $n = 6$ ). For the period from 1999 onward, the rate rose (in absolute value) to  $-1.0$  percentage points per year (standard error = 0.04;  $n = 5$ ).<sup>9</sup> The difference in trends is significant

---

**Figure 4.** Poverty Rates for the Developing World Outside China

---



Source: Author's calculations.

---

( $t = 10.01$ ,  $p < 0.00005$ ). If the pre-2000 trajectory had continued, then 29.6 percent of the population of the developing world outside China would have lived below \$1.25 a day in 2012 instead of 23.4 percent, representing an extra 280 million people who would otherwise have lived below \$1.25 a day.

As an aside, the clear signs of a trend break in about 2000 in figure 4 might be taken to suggest that the MDGs have helped because they were officially ratified that year at the Millennium Summit. It is also notable that we would expect the MDGs to have had an impact outside China, notably in Africa; this is what we see in the data. However, attribution to the MDGs must await a fuller consideration of other possible explanations for the break in trajectories at about 2000.

## Pessimistic and Optimistic Trajectories

In the following analysis, whether one is “optimistic” about future progress in poverty reduction depends on whether one thinks that the new path for the poverty rate in the developing world outside China since 2000 will be extended into the foreseeable future. Current population projections imply that lifting one billion people out of poverty, measured against the \$1.25-a-day standard, would require a poverty rate of only 3 percent (to the nearest integer).<sup>10</sup> In describing the two paths to that goal, the discussion will focus on three dates: 2022 (10 years from the time of this writing), 2030, and the date at which the 3 percent target is reached.

The low-case trajectory of poverty reduction implies that the developing world will not reach that 3 percent target by 2030 or even well beyond that date. By this trajectory, the developing world outside China will return to its pre-2000 pace of poverty reduction from 2012 onward, although China will remain on track. Projecting the series forward linearly from 2012 but at the pre-1999 rate of poverty reduction, this trajectory implies a poverty rate of 17.7 percent for the developing world outside China in 2022 and 14.7 percent by 2030.<sup>11</sup> China’s poverty rate would have fallen to zero. Given that current population projections indicate that China’s share of the developing world’s population will be 21 percent by 2022 (and 20 percent by 2030), the overall poverty rate would fall to 14.0 percent by 2022 and to 11.8 percent by 2030. The number of poor would fall from 1.1 billion in 2012 to 0.9 billion in 2022 and 0.8 billion in 2030. In this low-case trajectory of poverty reduction, it would not be until 2060 that one billion people would likely be lifted out of poverty.

The guiding principle for the optimistic benchmark is that the recent success against extreme poverty in the developing world as a whole will be maintained. Two main ways are used here to quantify such a benchmark. The first is based on time-series projections of past experience, while the second method (in this section and the next) uses simulations.

A simple linear projection of the series in figure 2 gives a poverty rate for 2022 of 8.6 percent (standard error = 1.0 percent), or 0.6 billion people, falling to 3.4 percent (standard error = 1.2 percent), or 0.2 billion people, by 2027. The 95 percent confidence interval for the 2027 poverty rate is (1.0 percent, 5.8 percent). The precise date at which the 3 percent target is expected to be reached is 2027.4 (standard error = 1.16 years); the 95 percent confidence interval for the year in which the 3 percent target is met is (2025, 2030).

The linear projection is an implausible specification for the PG index since it implies that the PG index goes to virtually zero by 2022. (The projected value is 0.3 percent, which is not significantly different from zero; standard error = 0.97 percent.) A log specification for projecting the income gap ratio clearly performs better and gives a projected PG index of 3.9 percent for 2017, 2.3 percent for 2022, and 0.9 percent for 2027.<sup>12</sup> Notice that these projections (allowing for nonlinearity) imply only a modest decline in the income gap ratio under the optimistic trajectory, from 0.31 in 2010 to 0.26 in 2027. The main driving force for the lower PG index is expected to be the decline in the headcount index.

Despite the strong linearity in reducing the headcount index over time (as evident in figure 3), the method of linear projection can be questioned as a means of setting the optimistic benchmark for future progress. The method may suffer from a form of “aggregation bias.” As we have seen, past progress has been quite uneven across regions (and countries). China has naturally had a high weight, and we can reasonably expect the virtual elimination of extreme poverty there by 2022.<sup>13</sup> Then, we should recalculate the projection setting China’s poverty rate to zero in 2022. For 2022, the linear projection for the developing world outside China is 18.5 percent living below \$1.25 a day. With a projected 79 percent of the developing world’s population living outside China by that date, we would see an overall poverty rate of 14.6 percent. However, this calculation ignores the clear acceleration that we have seen in the progress of absolute-poverty reduction in the developing world outside China since the mid-1990s. If instead we project forward from 1999 as the baseline, the poverty rate in the developing world outside China would fall to 11.9 percent by 2022, implying an aggregate rate of 9.4 percent—only slightly higher than the simple linear projection.

An alternative method of checking for aggregation bias is to use simulation methods that apply country-level growth projections to baseline distributions at country level and then aggregate up to the global level. This calculation is more complicated. (Note that the base-year distribution is projected forward for each country separately. A constant elasticity is not assumed, as this is unlikely to hold.) However, reassuringly, the calculation is in close accord with the prior calculation based on aggregate trends. Trajectories for poverty in 2015 by this method are reported in the 2012 Global Monitoring Report, drawing on the

Bank's country-specific economic growth projections. These projections give a poverty rate of 16.3 percent for 2015 (World Bank 2012a),<sup>14</sup> which is very close to that implied by the first method, namely, 15.9 percent. Table 1 includes the forecasts by region for 2015 from the Global Monitoring Report. As is clear from table 1, success in reducing poverty in South Asia and sub-Saharan Africa will be crucial to overall progress in poverty reduction. Consistent with the more optimistic scenario above, 80 percent of the projected extreme poverty count for 2015 is in these two regions (about equally).

A further concern is that a constant annual-percentage-point decline cannot continue indefinitely. The trajectory can be expected to slow, becoming nonlinear in time. What is much less clear is when the trajectory is likely to slow. The next section will use a version of the simulation method (as described above for addressing aggregation bias) to determine whether the linear trajectory for the overall poverty-reduction rate is consistent with the expected economic growth rates associated with the optimistic path under explicit assumptions about income distribution.

## An Alternative Approach Using Simulations

We have observed that the linear projection of the time series of poverty measures for the developing world suggests that one billion people could be lifted out of poverty by 2027. However, the above discussion has also pointed to some concerns about whether the linearity will hold that long. This section approaches the problem of quantifying an optimistic trajectory in a rather different way, without relying on the time-series evidence.

The key assumption driving the following simulations for the optimistic path is that the overall level of inequality does not increase going forward. In other words, all income levels across the developing world as a whole grow at the same rate, maintaining overall inequality at the same level. As discussed above, while inequality has declined over the period as a whole, new forces may well emerge to put upward pressure on inequality, and we have seen an indication of such upward pressure since 2005 (figure 1). For example, starting in the near future, China's economic growth will begin to put upward pressure on overall inequality in the developing world. Whether Africa maintains its higher average economic growth rates since 2000 or so will be key to whether the 3 percent target will be reached by 2025–30.

To quantify the implications of an inequality-neutral process of economic growth in the developing world as a whole, one can use the baseline distribution and project this forward with a higher mean until any given poverty rate is met and then determine what growth rate is required. This method automatically

accounts for the nonlinearity in how the cumulative distribution function of consumption varies with the poverty line relative to the mean. (Note that nonlinearity can hold, yet the path of poverty-reduction measures over time can still be a roughly linear trajectory.) Thus, we ask the following question: will a continuation of the higher economic growth rates seen in recent times in the developing world be sufficient to lift one billion people out of poverty without a change in the relative distribution of income?

The results of this exercise are found in table 2, column (1), which gives the growth rate in household consumption per capita needed to bring the \$1.25-a-day poverty rate down to various levels by 2027, assuming no increase (or decrease) in inequality within the developing world as a whole. The calculations in column (1) are anchored to the 2008 distribution (the latest reference year in *PovcalNet*). For each targeted poverty rate, one solves backward to find the growth rate needed to reach that target, holding constant the 2008 Lorenz curve.

This analysis indicates that without any increase in overall inequality, a growth rate of 4.5 percent in mean household consumption per capita (which is very close to the trend rate of growth in the mean of 4.3 percent since 1999) would get us close to the 3 percent poverty rate in 2027. This simulation suggests that the optimistic target of lifting approximately one billion people out of poverty by 2025–30 is attainable with a continuation of the relatively higher economic growth rates observed in the developing world in the last 15 years or so as long as this does not come with rising overall inequality.

**Table 2.** Growth Rates for the Developing World as a Whole Required for Various Target Poverty Rates at a Fixed Level of Overall Inequality within the Developing World

Target poverty rate for 2027 (\$1.25 a day)	(1) Required annual growth rate in household consumption per capita at 2008 level of inequality	(2) Required annual growth rate in household consumption per capita at 1999 level of inequality
1%	7.6%	6.1%
2%	5.5%	4.0%
3%	4.5%	3.4%
4%	3.7%	2.8%
5%	3.3%	2.3%
6%	2.9%	2.0%

*Note:* These are the required annual growth rates for survey means over the 2008–2027 period. The calculations in column (1) are based on the distribution of consumption for the developing world in 2008, whereas those in column (2) are anchored to the 2008 mean but use the Lorenz curve for 1999, which gives the lowest overall inequality during the 1981–2008 period. For each target poverty rate, the author solved backward to find the growth rate needed to reach that target, holding constant the Lorenz curve at either 2008 or 1999.

*Source:* Author's calculations using *PovcalNet*.

The goal could be comfortably achieved even if recent economic growth is not maintained as long as overall inequality falls, and the drop in overall inequality would be within the range of recent experience. To illustrate the magnitudes, column (2) of table 2 gives the economic growth rates needed to attain each poverty rate in 2027, using the relative distribution of 1999. The year 1999 was the year of lowest total inequality in the series in figure 1, with an inequality index of 0.52, compared to 0.57 in 2008. This lower level of inequality will entail lower poverty but will also foster a more poverty-reducing pattern of growth going forward. We now see that if mean consumption grew at 3.4 percent (with no further change in inequality), we would reach a 3 percent poverty rate by 2027.

Naturally, a higher economic growth rate would most likely achieve the poverty-rate target ahead of that date. For example, an extra 1.2 percentage points in the economic growth rate at the lower inequality path of column (2) would lift one billion people out of poverty five years ahead of time, by 2022. At the higher inequality path of column (1), this would require an extra 1.6 percentage points.

## Conclusions

The developing world as a whole has seen enormous progress against absolute poverty. While China has contributed greatly to that progress, it goes well beyond China's borders. The developing world outside China saw slow progress in reducing poverty until around the turn of the century, but that progress has accelerated since then. The acceleration in growth since 2000 casts doubt on the idea (which is still common in the development economics literature) that the typical, less developed economy is in some form of a poverty trap. Perhaps capital flows or enhanced productivity of capital through policy reforms, or both, has meant that the developing world as a whole is in the process of escaping whatever trap it found itself in (if indeed there was one). Or perhaps the countervailing (economic and political) forces creating the trap will soon reemerge to retard continued progress and even bring lasting setbacks.

In thinking about how we might judge future progress in reducing poverty, a natural choice for a low-case, "pessimistic," trajectory is to assume that the developing world outside China does indeed regress back to the relatively slow progress of the 1980s and 1990s. This trajectory implies that the proportion of the developing world's population living below \$1.25 a day would fall to 12 percent over the coming 15 years, from 19 percent in 2012. On this path, it would take another 50 years or so to lift one billion people out of poverty. This would surely be judged to be poor performance. This paper has attempted to suggest what good performance might look like.

Goals that can motivate extra effort to do much better than this pessimistic trajectory must represent real progress in reducing the incidence of extreme poverty below that trajectory, but they cannot be so far outside the range of experience as to be deemed impossible. This paper's proposed optimistic benchmark for the \$1.25 poverty rate in 10 years' time is 9 percent. This "staying-on-the-path" scenario clearly represents very good economic performance, maintaining an impressive recent trajectory of poverty reduction. In the 20 years from 1990 to 2010, the developing world halved its overall poverty rate from 43 percent to 21 percent. On this trajectory, it would be halved again in a mere 10 years.

The optimistic trajectory suggests that we can be confident that the goal of lifting one billion people out of poverty (relative to the count of the number of poor in 2010) would be reached by 2025–30, with 2027 as the most likely date. However, this assumes that the robust linear path that we have observed for the reduction of the poverty rate over time will be maintained. That challenge will not be easy to meet. Instead, it might be conjectured that the pace of poverty reduction will begin to decline at low levels—below, say, 10 percent—thus making it more difficult to reach the goal. From what we know, we cannot be confident about when such a slowdown might be expected.

The paper has also provided simulations to determine what combinations of economic growth and distributional changes might attain the optimistic trajectory. The simulations suggest that a continuation of the economic growth performance of the developing world as a whole since 2000 can be expected to lift one billion people out of poverty by 2027 provided that there is no further deterioration in overall inequality. Although overall inequality in the developing world has been fairly stable since the 1990s, there have been signs of its recent rise. If this continues, then higher economic growth rates than those seen since 2000 will be needed to reach the proposed poverty reduction target. By contrast, a reduction in overall inequality would enable one billion people to be lifted out of poverty with lower economic growth than we have observed in recent times; lower inequality will both directly reduce poverty and make subsequent economic growth more poverty reducing.

In choosing among the multiple solutions for lifting one billion people out of poverty, the sustainability of poverty-reduction efforts is clearly important. We do not want to reach the poverty-reduction target only to fall back in subsequent years. On an encouraging note, recent research has suggested that lower initial levels of absolute poverty at a given mean consumption foster higher subsequent rates of growth in average living standards in developing countries and help to ensure that economic growth itself is poverty reducing (Ravallion 2012). Thus, a "virtuous cycle" can be anticipated that would help to ensure the sustainability of the reduction in poverty.

Environmental sustainability is also important. An economic-growth path that consumes all of a country's natural resources in 15 years may attain the proposed poverty-reduction target, but the poverty rate would likely bounce back. However, existing measures of poverty reduction do not tell us whether the changes observed are actually sustainable; they relate only to the current time period, which might be quite short. Thus, we would need a separate check on the sustainability of observed poverty reduction. The best data that we currently have for that check appear to be the "adjusted net savings" rate based on [Hamilton and Clemens \(1999\)](#).<sup>15</sup> The sustainability of a reduction in poverty that came with reduced (or negative) adjusted net savings would clearly be questionable.

The best sustainable route will naturally vary from country to country. The economic growth projections underlying these benchmark trajectories are grounded in the economic realities of both the particular countries concerned and the global economy. However, the policy challenges of ensuring that poor households share sufficiently and sustainably in that economic growth at country level still need to be addressed.

The bulk of the work needed to reach this global poverty-reduction target will need to be done at the country level. A similar benchmarking exercise for individual countries would be desirable. Naturally, this work would need to reflect the specific conditions and resource base of each country. One should be cautious in attempting to use cross-country comparisons to infer what any one country can accomplish by a specific year, even when such comparisons control for observable differences; there are invariably idiosyncratic factors at country level that cannot be addressed by such cross-country comparisons. This work should ideally accompany a reasonably clear plan of how the lower poverty-reduction targets would be achieved. Various tools for economic analysis exist that can inform such plans, ranging from computable general equilibrium models to microsimulations tools, all of which have strengths and weaknesses.<sup>16</sup>

Monitoring performance against these benchmarks poses a number of serious data challenges. There has been huge progress in collecting primary household survey data. When the World Bank's current global poverty-monitoring effort began in 1990, the estimates used 22 surveys for 22 countries ([Ravallion et al. 1991](#)). Today, approximately 900 surveys are used by the Bank's researchers, spanning 125 countries, with more than six per country. The latest estimates use a "global" sample of 2.1 million households. However, many problems remain. There are persistent lags and uneven coverage. (The surveys used here cover 90 percent of the population of the developing world as a whole in 2008, but this varies from 94 percent in East Asia to only 50 percent in the Middle East and North Africa.) Understandably, there are continuing concerns about the comparability of the surveys over time and across countries. Furthermore, there are concerns about underreporting and selective compliance in household surveys; the rich are difficult to interview, and this task is not

becoming any easier. Moreover, the weak integration of macro and micro data is a long-standing concern that warrants more attention than it has received. Better data will help to determine how close—or far—we are from reaching our development goals, including poverty reduction, and in assessing the efficacy of alternative policies for attaining those goals.

## Notes

1. The author is the Edmond D. Villani Professor of Economics at Georgetown University; his email address is [mr1185@georgetown.edu](mailto:mr1185@georgetown.edu). At the time of this writing, the author was Director of the World Bank's research department. The author is grateful to Jim Yong Kim for encouraging him to think about this issue and for helpful discussions with Kaushik Basu, Francois Bourguignon, Shaohua Chen, Stefan Dercon, Shanta Devarajan, Francisco Ferreira, Hiroki Uematsu, Sri Mulyani Indrawati, Peter Lanjouw, Ernesto May, Johan Mistiaen, Mahmoud Mohieldin, David Rosenblatt, Jaime Saavedra, Warrick Smith, Hans Timmer, Jos Verbeek, Nobuo Yoshida, Dominique van de Walle and seminar participants at the Overseas Development Institute, London, the UK Department of International Development, London, and the World Bank.

2. The only careful attempt to benchmark performance at the country level appears to be [Newman et al. \(2010\)](#), who draw on time-series evidence across developing countries. They use the empirical distribution of absolute changes in the poverty rate as data to inform the establishment of performance benchmarks for countries in Latin America and the Caribbean.

3. The MDGs were not the first such effort at goal setting; Hume identifies a number of antecedents, including the UN Declaration of Human Rights.

4. Inequality is in quotation marks here because this is not “inequality” as normally defined but rather a more complex function of relative distribution, as discussed in [Datt and Ravallion \(1992\)](#). However, for the present purposes, it is defensible to refer to this simply as “inequality.”

5. Two further tests were used. First, the calculation was repeated dropping the 2010 observation because weaker survey coverage means that the estimate is more affected by NAS consumption growth rates used for interpolation. However, the growth rate for the recent period was almost identical (4.3 percent with a standard error of 0.4 percent). Second, these calculations do not constrain the predicted values of the log mean to be identical in 1999. If one prefers to impose that constraint, then the trend growth rates for the two subperiods are slightly lower, at 0.7 percent per annum (0.1 percent) and 4.1 percent (0.2 percent), respectively.

6. Note that these are aggregate indices for the developing world as a whole. Different regions have observed different patterns in how average inequality across countries has been evolving; [Ravallion and Chen \(2012\)](#) provide details by regions.

7. The estimates of absolute poverty measures used here are from Chen and [Ravallion \(2012\)](#), who discuss data and methods. (The update for 2010 is 20.8 percent below \$1.25 a day, representing 1.227 billion people.) A fuller discussion of the absolute measures (with various tests of robustness) can be found in [Chen and Ravallion \(2010\)](#).

8. The Ramsey RESET test using the squared fitted values indicated that one could not reject the null hypothesis of linearity; the  $t$  statistic for the squared fitted values was  $-0.83$ . Moreover, the null could not be rejected on adding a cubed term.

9. If one constrains the predicted values of the poverty rate to be identical in 1999, then the trend-based annualized rates of change become  $-0.4$  percent per annum (0.02 percent) and  $-0.9$  percent (0.04 percent).

10. The required poverty rate falls from 3.3 percent in 2012 to 2.8 percent in 2030. The Bank's current population projections in millions for the developing world, as defined in 2012, are 5.884

in 2012, 6,230 in 2017, 6,561 in 2022 and 7,048 by 2030. Using the 1990 classification of developing countries (to be consistent with *PovcalNet*), the numbers are 6,009, 6,355, 6,687 and 7,173, respectively. This difference does not change the poverty counts reported in this paper at the stated number of statistically significant figures.

11. The projected poverty rate for the developing world outside China in 2012 (based on the trajectory since 1999) is 21.5 percent (standard error = 0.4 percent), and the rate of poverty reduction based on the pre-1999 trajectory is  $-0.38$  percentage points per year.

12. Regressing the log of the income gap ratio (PG index divided by headcount index) on the year, one obtains projections of 0.282 for 2017, 0.271 for 2022, and 0.261 for 2027. The projected PG indices are then obtained by applying these projections to the prior linear projections for the headcount index.

13. Note that China's current official poverty line is \$1.80 a day (at 2005 purchasing power parity), above the international line used here of \$1.25.

14. These forecasts factor in historical gaps between growth rates in the survey means and growth rates in private consumption per capita in the NAS. For most countries, approximately 90 percent of the NAS growth rate is passed onto the survey means, but for India, it was only approximately half, consistent with the larger gap between the two growth rates for India. In addition, an allowance is made for rising inequality in both China and India. For further details, see [World Bank \(2008\)](#).

15. At the time of this writing, the latest estimates by country are in [World Bank \(2012c\)](#); see table 4.11.

16. A useful compendium of the tools available can be found in [Bourguignon et al. \(2008\)](#). On microeconomic simulation methods, see [Ferreira and Leite \(2003\)](#).

## References

- Azariadis, C. 2006. "The Theory of Poverty Traps. What Have we Learned?" In S. Bowles, S. Durlauf, and K. Hoff, eds., *Poverty Traps*. New York: Russell Sage Foundation.
- Bourguignon, F. 1979. "Decomposable Income Inequality Measures." *Econometrica* 47 (4): 901–20.
- Bourguignon, F., M. Bussolo, and L. Pereira da Silva, eds. 2008. *The Impact of Macroeconomic Policies on Poverty and Income Distribution. Macro-Micro Evaluation Techniques and Tools*. Washington, DC: Palgrave Macmillan and the World Bank.
- Chen, S., and M. Ravallion. 2010. "The Developing World is Poorer than We Thought, but No Less Successful in the Fight Against Poverty." *Quarterly Journal of Economics* 125 (4): 1577–625.
- . 2012. "More Relatively-Poor People in a Less Absolutely-Poor World." Policy Research Working Paper 6114. World Bank, Washington, DC.
- Datt, G., and M. Ravallion. 1992. "Growth and Redistribution Components of Changes in Poverty: A Decomposition with Application to Brazil and India." *Journal of Development Economics* 38: 275–95.
- Ferreira, F., and P. Leite. 2003. "Meeting the Millennium Development Goals in Brazil: Can Microeconomic Simulations Help?" *Economica* 3 (2): 235–79.
- Ferreira, F., and M. Ravallion. 2009. "Poverty and Inequality: The Global Context." In W. Salverda, B. Nolan, and T. Smeeding, eds., *The Oxford Handbook of Economic Inequality*, 599–638. Oxford: Oxford University Press.
- Hamilton, K., and M. Clemens. 1999. "Genuine Savings Rates in Developing Countries." *World Bank Economic Review* 13 (2): 333–56.

- Hume, D. 2009. "The Millennium Development Goals (MDGs): A Short History of the World's Biggest Promise." BWPI Working Paper 100. Brooks Institute, University of Manchester, Manchester, UK.
- Korinek, A., J. Mistiaen, and M. Ravallion. 2006. "Survey Nonresponse and the Distribution of Income." *Journal of Economic Inequality* 4 (2): 33–55.
- Kraay, A., and C. Raddatz. 2007. "Poverty Traps, Aid and Growth." *Journal of Development Economics* 82 (2): 315–347.
- Kuznets, S. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45:1–28.
- Lim, J. 2012. "The Emerging Pattern of Global Investment." Prospects for Development Blog, World Bank. <http://blogs.worldbank.org/prospects/the-emerging-pattern-of-global-investment>.
- Newman, J. L., J. P. Azevedo, J. Savedra, and E. Molina. 2010. "The Real Bottom Line: Benchmarking Performance in Poverty Reduction in Latin America and the Caribbean." Washington, DC: World Bank.
- Ravallion, M. 1997. "Can High Inequality Developing Countries Escape Absolute Poverty?" *Economics Letters* 56:51–57.
- . 2001. "Growth, Inequality and Poverty: Looking Beyond Averages." *World Development* 29 (11): 1803–15.
- . 2005. "A Poverty-Inequality Trade-Off?" *Journal of Economic Inequality* 3 (2): 169–82.
- . 2012. "Why Don't We See Poverty Convergence?" *American Economic Review* 102 (1): 504–23.
- Ravallion, M., and S. Chen. 2007. "China's (Uneven) Progress Against Poverty." *Journal of Development Economics* 82 (1): 1–42.
- . 2012. "Monitoring Inequality." Blog Post on Let's Talk Development, World Bank. <http://blogs.worldbank.org/developmenttalk/monitoring-inequality>.
- Ravallion, M., S. Chen, and P. Sangraula. 2009. "Dollar a Day Revisited." *World Bank Economic Review* 23 (2): 163–84.
- Ravallion, M., G. Datt, and D. van de Walle. 1991. "Quantifying Absolute Poverty in the Developing World." *Review of Income and Wealth* 37 (4): 345–61.
- Sachs, J. 2005. *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals*. New York: Millennium Project, United Nations.
- World Bank. 2008. "Global Economic Prospects." *Development Economics*. Washington, DC: World Bank.
- . 2012a. "Global Monitoring Report 2012." *Development Economics*. Washington, DC: World Bank, in collaboration with the International Monetary Fund.
- . 2012b. "Global Economic Prospects June 2012." *Development Economics*. Washington, DC: World Bank.
- . 2012c. *World Development Indicators 2012*. Washington, DC: World Bank.

---

# What Is a Civil War? A Critical Review of Its Definition and (Econometric) Consequences

---

Mark Gersovitz • Norma Kriger

---

*We argue that the academic literature, both qualitative and quantitative, has mislabeled most episodes of large-scale violence in Africa as civil wars; these episodes better fit our concept of regional war complexes. Our paper seeks to highlight the fundamental flaws in the conception of civil war in the econometric literature and their implications for econometric specification and estimation, problems that this literature is inherently incapable of rectifying. We advocate the comparative study of regional war complexes in Africa based on historical narratives. JEL codes: D74, H56, O10*

## Preliminaries

Civil wars and similar large-scale violence bring death, injury, disease, displacement, and misery. The impairment of the state diminishes its provision of protection for people and property and of basic social services, such as health care and education. Private economic activity becomes difficult or impossible. People want to understand these episodes and how they can be avoided.

This essay focuses on two broad fatal flaws in the econometric literature on civil war, a large and rapidly growing body of literature that has received considerable attention (for surveys, see [Collier and Hoeffler 2007](#); [Dixon 2009](#); [Blattman and Miguel 2010](#)). First, rather than providing a definition of civil war, econometric studies use a coding rule based on battle deaths (or, sometimes, any deaths) associated with civil war. Second, these studies apply econometric techniques that cannot address the spatial and temporal aspects of large-scale violence. (We leave aside questions of bidirectional causality between the dependent and so-called independent variables that plague all econometric studies using country-level data.)

A good definition of civil war is an essential starting point for analysis. The definition of civil war informs the universe of potential cases that can be compared with respect to their causes, dynamics, consequences, termination, and any other concerns, such as postwar reconstruction and rehabilitation. Moreover, the definition of civil war is critical for identifying trends in the nature and location of different types of warfare. After defining civil war, we critique the coding rules based on the number of deaths used by econometric studies and the properties of the data produced by such coding rules. We pay particular attention to the temporal and spatial properties of these data and their implications for econometric specification and estimation. We also develop our critique of the econometric literature on civil war through the use of narratives of historical experiences of large-scale violence. This qualitative approach leads us to understand these episodes of large-scale violence as regional war complexes. We consider both the historical comparative method and the concept of a regional war complex to be promising ways forward in the study of large-scale violence. We limit our examples of civil wars and how to analyze them to Africa only because this region is our area of primary specialization and we have first-hand experience with this region. Africa is a region that suffers from large-scale violence.

## A Definition of Civil War

No abstract definition of civil war can provide a set of necessary and sufficient conditions that can be applied mechanically. Even with a good abstract definition of civil war, the process of classifying which countries have experienced civil wars and when inevitably involves judgments based on close reading of the historical narratives of country experiences with large-scale violence. Not all countries will share all dimensions in the definition of civil war.

We are reassured by our position's affinity to the classic discussion of the problems of definition illustrated by Wittgenstein (1958, §66) in his examination of the definition of "games." He concluded: ". . . we see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail." Therefore, the best that can be done in providing a definition of "games" is a list of activities that one believes should be classified as games and some general commentary on the similarities that some or others of them share. We believe that the concept of civil war shares some attributes of such a family-resemblance concept.

We define a civil war as a politically organized, large-scale, sustained, physically violent conflict that occurs within a country principally among large/numerically important groups of its inhabitants or citizens over the monopoly of physical force

within the country. Civil wars usually have incumbent governments that control the state and have a monopoly of force before the civil war and challengers—people who have not effectively challenged the monopoly of others before the outbreak of the civil war but whose challenge initiates the outbreak of the civil war. The challengers may begin as a small group, but for the episode to rise to the level of a civil war, they must become numerically important. The challengers may seek to replace the incumbents in control of the monopoly of force within the extant territory of the state, or they may seek the secession of part of the original territory. The achievement of their goals must be plausible. A civil war is not just a sufficiently large group of people over a sufficiently long time who make trouble for the government and undermine security, such as large drug cartels, because they do not aim to possess the monopoly of force. Civil wars must entail large-scale and sustained internal political violence to distinguish them from intense but limited episodes of political violence that contest the monopoly of force, such as political assassinations, mutinies, or coups. Civil war violence may involve external actors, but the violence occurs within the boundaries of a country and predominantly involves internal actors. This last characteristic of our definition is perhaps the most problematic because almost all violence that is called a civil war has some external dimension. We return to this issue in the section on regional wars. Although our conception is narrow, we believe that the elements of other violent internal political conflict are critical to understanding the causes of civil war. In short, we prefer a narrow definition and a broad analysis.

Our definition, like many others, derives, in part, from Max Weber's classic characterization of the state: ". . . the state is that human community which within a defined territory successfully claims for itself the *monopoly of legitimate physical force*; and 'territory' . . . is a characteristic of the state . . . . The state is the sole source for the 'right' to exercise violence" (Weber, 2004, 131–32). Thus, for us, at a minimum, civil war is one form of the impairment of the state. Unlike Weber, we do not mention legitimacy; if a tyrant can maintain the uncontested monopoly of force, there is no civil war. We also draw selectively on various authors on civil wars, most of whom provide requisites of civil war rather than abstract definitions (Fearon 2007; Kalyvas 2006; Edmonds 1972; Keegan and Bull 2006). These authors differ on numerous issues, including how broadly they conceive of wars (e.g., the inclusion of guerrilla insurgencies or only conventional military warfare), the importance of mass participation in violent conflict, and the importance of whether the war is fought exclusively by nationals or whether only most combatants must be nationals. On one issue, there is broad consensus: civil war violence takes place within the country.

Not all authors subscribe to the merits of defining civil war. Some authors (e.g., Cramer 2007, especially ch.2; Boås and Dunn 2007, 4) believe it may be dangerous to isolate civil war as a discrete phenomenon. For example, Cramer (2007, 51) fears

that focusing on civil wars “produces a misleading idea that violent conflict is a relatively rare phenomenon.” Cramer and those who share his conceptual critique of civil war prefer to conceive of violence as a continuum. We take from these critics the important caution that the absence of civil war is not necessarily peace, as it is in the quantitative literature based on death counts, and that one should expect to find prior to a civil war the violent origins of the war itself. However, the notion of a continuum of violence shares a fundamental weakness with the 1960s concept of internal war, which included almost all forms of violence against authority: assassinations, coups, rebellions, revolutions, and civil wars (Eckstein 1964 and 1965; Rosenau 1964; Luard 1972). There are too many disparate phenomena under simultaneous consideration. We therefore remain committed to the a priori value of a definition of civil war. Baev (2007, 248), writing on post-Soviet conflicts, challenges the rationale “for drawing firm analytical borders by exploring several ‘gray areas’ where civil war becomes barely distinguishable from other forms of violence.” This position is shared by Salehyan (2009), who considers the distinction between civil wars and international wars to be arbitrary (p. 15), although he persists in using both terms in his study of transnational rebels. We are committed to a rigorous definition of civil war, but we share Baev’s (2007, 247) belief that “only very nuanced examination rather than application of rigid criteria could help in distinguishing civil wars from other crises.”

We now discuss how to implement an abstract definition of the type that we have provided as well as the pitfalls inherent in the application of the coding rules used in the econometric literature. We focus on two dimensions of our definition of civil war: “large-scale, sustained, physically violent conflict” (the temporal dimension) and conflict “among large/numerically important groups of its inhabitants or citizens” (the spatial dimension).

## Large-Scale, Sustained, Physically Violent Conflict

Now, we focus on the part of our definition that refers to large-scale, sustained violent conflict. One measure of physical violence is the number of people killed as a result of a violent conflict over the monopoly of force. This general intuition is the basis for the coding rule in the quantitative literature. In this section, our goal is to explain and assess the rules that econometric studies use to code civil wars and three time-series properties of the data that they produce: the number of countries reported as having civil wars, the extent to which these wars are coded as being sustained (as we believe from the narratives that they are), and the correlation over time of whether a country is reported as having a civil war. In econometric terms, the first two properties of the data correspond to errors in variables, and the last property corresponds to autocorrelation in the dependent variable.

The most recent set of data used widely by econometric researchers is the Armed Conflict Dataset of the International Peace Research Institute of Oslo and the Department of Peace and Conflict Research, Uppsala University (hereinafter PRIO/Uppsala), discussed in Gleditsch et al. (2002) and documented in the codebook of Strand et al. (2005). This data set is available in various versions. We do not use the most recent one to remain as consistent as possible with the published studies on which we comment. The codebook defines four types of conflict in which a country can be involved (Strand et al. 2005, 10). Two types of conflict are relevant to our discussion because they are used in econometric studies of civil war: “[type] 3. *Internal armed conflict* occurs between the government of a state and internal opposition groups without intervention from other states. [type] 4. *Internationalized internal armed conflict* occurs between the government of a state and internal opposition groups with intervention from other states.” A type 4 conflict is qualified by the variable “location,” which identifies whether the country that is coded as involved in a type 4 conflict experienced violence in its territory. The data set codes each type of conflict into three categories of intensity, of which the first two are “not ordinal”: “1. *Minor*: At least 25 battle-related deaths per year for every year in the period. 2. *Intermediate*: More than 25 battle-related deaths per year and a total conflict history of more than 1000 battle-related deaths, but fewer than 1000 per year. 3. *War*: At least 1000 battle-related deaths per year” (Strand et al. 2005, 10).

Table 1 presents, in compact form, the entire PRIO/Uppsala data (Version 3) for the countries on the African continent for the years 1960–2004. In the original data set, the categories type 3 and type 4 are not mutually exclusive. Table 1, however, presents the data in categories that are mutually exclusive and exhaustive so that all the information in any cell (country-year) can be represented by only one symbol. For countries with two rows of codes, in the top row (row 1) our coding is as follows: An X denotes conflicts that are type 3 (“internal”) at intensity 3 (“war”) if PRIO/Uppsala does not also designate a type 4 conflict at intensity 3 in that year. An I denotes a type 4 (“internationalized internal”) conflict at intensity 3 in which the “location” of the conflict was within the country if PRIO/Uppsala does not also designate a type 3 conflict at intensity 3 in that year. An O denotes years that PRIO/Uppsala designates as both type 3 and type 4 (and the location of the type 4 conflict is within the country), each at intensity 3 (i.e., there is overlap according to the PRIO/Uppsala data). An F denotes type 4 conflicts of intensity 3 in which a country was involved but for which the location was outside that country. There are no countries that have years in which PRIO/Uppsala designates a type 3 conflict and a type 4 conflict with a location outside the country. A blank denotes a year of no conflict of either type 3 or type 4 at intensity 3, in effect a year of peace in the coding used by the econometric studies if they use intensity 3 data. In the bottom row (or the only row for countries with only one row; row 2), X, I, O, F, and blank







have the same meanings as just discussed except that the intensity is 1–3 (“minor,” “intermediate,” or “war”) rather than intensity 3 alone.

Among the econometric studies that have recoded the PRIO/Uppsala data to study aspects of civil war, we provide three examples: Miguel et al. (2004), Gleditsch (2007), and Buhaug and Gleditsch (2008) (hereinafter MGB). Miguel et al. (2004) study whether countries in sub-Saharan Africa had a civil war for each of the years 1981 to 1999. Because they used a slightly earlier version of the data, their table C1 is not identical to our table 1 when the two tables overlap, but they seem close and consistent in their method of construction. Miguel et al. (2004, 730–31, and table C1) seem to use primarily the union of the X, I, O, and F data at intensities 1–3 (corresponding to the concepts in row 2 of table 1), but they also present some results based on intensity 3 (row 1 of table 1). In a study of the onset of civil war, Gleditsch (2007, 300–01) seems to do the same but considers only states in which conflict has erupted to have had an onset of civil war, seemingly excluding our F coding. Buhaug and Gleditsch’s (2008, 223) study of civil war onset also seems to use the union of X, I, and O data for intensities 1–3. Earlier quantitative studies using different data also code a civil war by a battle-death count, usually of at least 1,000 per year (Licklider 1995), or more than 1,000 battle-related deaths in any major conflict (Wallenstein and Sollenberg 1999, 595). Thus, the most recent studies seem to favor the lower threshold of row 2, but the earlier studies favored a higher threshold, similar to that of row 1. As argued below, both approaches have serious shortcomings.

Deaths related to a conflict are difficult to estimate, a point well made by Cramer (2007, table 1). Furthermore, “battle” is a strange modifier of deaths when most of the violence that is considered part of civil wars in these data sets does not involve set-piece battles, making it difficult to know what a battlefield death is. This modifier is especially problematic because these episodes often involve civilian deaths. These deaths hardly ever seem to occur on a battlefield, and certainly not in a case like Sierra Leone, where terrorizing civilians in their villages and homes was a prevalent strategy during the conflict. Similar situations occur when civilians in camps for refugees or internally displaced persons are targets, as in Darfur, Congo-Kinshasa, Burundi, or Rwanda. Some researchers employ a count of any death associated with large-scale violence (whether civilian or battle deaths) to identify a civil war. Doyle and Sambanis’s (2000b, 3) coding of wars “uses the 1,000 deaths threshold for the entire war as long as the war caused 1,000 deaths in any single year [sic].” To code the onset of a civil war, Fearon and Laitin (2003a, 76) require at least 1,000 deaths over its course, with a yearly average of at least 100 deaths.

Ultimately, the battle deaths are coded as 0 or 1 depending on whether the threshold of 1,000 deaths was met. It is not possible to use most of these data sets to obtain the actual number of deaths. Large countries tend to be more likely to meet the threshold and to be coded as having had civil wars. This may be why at least

some of these data sets ignore many violent episodes in the smaller countries of Africa that are widely thought to be civil wars, as discussed below. There are other problems with a definition of civil war based on the numbers and years of deaths (Sambanis 2004; Cramer 2007). Perhaps most telling is Sambanis's (2004, 835) conclusion from his correlation analysis of the coding of civil wars in the data sets that "there is still considerable disagreement about which armed conflicts should be classified as civil wars. Many wars are coded in only one out of a dozen data sets." This disagreement among data sets is illustrated in both tables 1 and 2. Furthermore, Sambanis shows that the choices in assembling these data on civil war-related deaths affect econometric and other inferences.

The whole focus on deaths relies on a concept of active violence rather than one of potential violence. A contest for the monopoly of force in a country can be under way without involving large numbers of deaths, such as if some areas of a country become no-go zones for the government, as evidenced by anecdotal accounts or the inability to raise revenues from an area. Côte d'Ivoire did not meet a 1,000-death threshold, so there is no row 1 reported for it in table 1. Nonetheless, there were many years in the 2000s when few people would have wanted to drive from Abidjan to Bouaké. Foreign forces were barely keeping the two sides apart, and the country was experiencing a de facto partition, the very essence of a contestation of the monopoly of force and civil war. For this and other reasons, Côte d'Ivoire is an important case, to which we return later.

When does one war end and a new one begin? Our criterion is that the contest over the monopoly of force has ended for the foreseeable future, not that there has been a period that is free of a certain type and number of deaths, as in the econometric literature. Ironically, to determine a war's end, many researchers who pursue an econometric approach rely on their judgment rather than making exclusive use of their own mechanical coding rules. For example, Licklider identifies a civil war as having ended when there is an end to concerns about living together, an end to multiple sovereignty or fewer than 1,000 battle deaths in each of five years (Licklider 1995, 682, 685). Doyle and Sambanis (2000a, 783) determine that a war has ended if "an end to the war and to residual lower-level violence and uncontested sovereignty" has occurred for two years. Doyle and Sambanis (2000b, 3) provide somewhat more detailed "rules of thumb" for determining a war's end on the basis of, among other criteria, peace settlements, a dramatic change in the parties, or case study literature that determines that there was a single war. We are sympathetic with such a multidimensional approach.

What happens when one attempts to determine whether one war has ended and a new one has begun or whether there has been one continuous conflict in a particular case, say the case of Zimbabwe? Both Licklider (1995) and Doyle and Sambanis (2000a, 2006) identify Zimbabwe as having had two separate wars (table 2), but they date the wars slightly differently. For Licklider, the first war was from 1972 to 1980; for Doyle and Sambanis (2006), it was from 1972 to 1979.

**Table 2.** Classification of Civil Wars: Some Alternatives in the Literature

Country		60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04		
Chad	A						I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	X									I	I					F			
	B						I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	X	X	X	X	X	X			X	O	O	X	X	X	X	F			
	C						I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	1,2	1,2	1,2	1,2	1,2	1,2	1	-	-	-	-	-		
	D						I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	2	2	2	2	2	2	2	2	2	2	2	2	2	2,3	3	3	3					-	-	-		
	E																						X	X	X	X	X	X	X	X									-	-	-	-	-	-	-	-		
Congo-Brazzaville	A																																															
	B																																															
	C																																															
	D																																															
	E																																															
Congo-Kinshasa	A					X	X											F																														
	B	X	X	X		X	X		X									F	X	X																												
	C	1	1	1	1	1	1												2	2																												
	D	1	1	1	1	1	1		2										3	3																												
	E	1	1	1	1	1	1																																									
Zimbabwe	A	-	-	-	-	-												X	X	X	X								F	F	F	F	F															
	B	-	-	-	-	-							X	X	X	X	X	X	X	X									F	F	F	F	F															
	C																																															
	D																																															
	E																																															

Notes: An empty cell denotes no civil war coded by the source, and a “-” denotes that the source ended before this date. For rows labeled A and B, coding is the same as that for rows 1 and 2 in table 1; these data do not distinguish one civil war from another for the same country. For rows labeled C, D, and E, an X denotes a civil war if the source coded only one war; otherwise, a number indicates a year of civil war and which one.

Source: A, table 1, row 1; B, table 1, row 2; C, Fearon and Laitin (2003b); D, Doyle and Sambanis (2006); E, Licklider (1995).

According to Licklider, the second war was from 1983 to 1984; for Doyle and Sambanis (2000a), the second war occurred in 1984. In subsequent work, Doyle and Sambanis (2006) date the second war from 1983 to 1987. Atlas and Licklider (1999) depict the first war as a “black-white civil war” in which the coalition of two guerrilla forces maintained cohesion against a common white enemy and the second war as “a black-on-black armed conflict between former civil war allies” who “had fought to bring Ian Smith’s Rhodesia to an end” (Atlas and Licklider 1999, 40). However, Atlas and Licklider also note that the relationship between the two guerrilla parties and their leaders was fraught even during the first war. Atlas and Licklider’s efforts to justify separating the two wars they identify for Zimbabwe actually point to the close connections and inseparability of the two wars. More important, the preoccupation with death counts comes at the expense of understanding the history of how the two violent episodes are connected (Kriger 2003).

A number of general points can be made based on further examination of the PRIO/Uppsala data presented in table 1. Points 1–4 relate to nonrandom errors in the classification of the dependent variable, and point 5 addresses the question of autocorrelation.

1. The strictest rule for coding a civil war that seems applicable to the PRIO/Uppsala data is “internal armed conflict” (type 3, denoted by X and O) and “war” intensity (the threshold based on 1,000 battle deaths in any one year, reported in row 1 for each country that has two rows). Countries with at least one civil war with this criterion are Algeria, Angola, Burundi, Chad, Congo-Kinshasa, Ethiopia, Liberia, Morocco, Mozambique, Nigeria, Rwanda, Sierra Leone, Somalia, South Africa, Sudan, Uganda, and Zimbabwe. These instances of civil war are far fewer than suggested by a cursory reading of the narrative literature on Africa. Cameroon, Central African Republic, Congo-Brazzaville, Côte d’Ivoire, Kenya (Shifta War, 1963 to 1967), Mali, Niger, and Senegal are all plausible cases of civil war, some more than others (however, as we argue in the next section, almost all countries in both groups better fit the definition of participants in a regional war complex).

2. If the coding of civil war is expanded to include “internationalized internal armed conflict” (type 4, denoted by I when there is no overlap with type 3 denoted by O) and the intensity continues to be “war,” some countries that were not coded as having a civil war now are: Congo-Brazzaville, Egypt, Guinea, Libya, Guinea-Bissau, and Senegal. In addition, the conflicts of some countries are denoted by F; these countries intervene in violent conflicts in other countries but do not experience associated violence within their own boundaries, which is hardly anyone’s idea of a civil war. Mauritania, Namibia, and Tanzania would not even appear in table 1 except for these types of conflicts. Such a conflict seems to be the only reason that Namibia appears in table C1 of Miguel et al. (2004), and such coding is material for their econometric inferences in some cases (Jensen and Gleditsch

2009). Tanzania seems to appear as a result of its invasion of Uganda after Amin attempted to annex part of Tanzania, so these dates for both countries seem closer to an interstate war than either type of internalized conflict. Ugandan exiles in Tanzania were also involved but probably could not have undertaken hostile action against Amin alone. Libya appears almost only for its interventions, except that in 1987, it was burned in its intervention in Chad, leading to the recording of an I; Libya also reports an X in its row 2 for 2002.<sup>1</sup>

3. Even those countries that would be coded by battle death–based rules as having had civil wars under the “war” intensity (1,000 battle-related deaths) in the PRIO/Uppsala data often show relatively few years of warfare (row 1 of table 1). Similarly, there are often oddly isolated years in what would seem, from narrative accounts, to have been sustained periods of conflict over the monopoly of force. Liberia is such an example, which experienced violent contestations of the monopoly of force following the Doe coup in 1980, beginning with the Dokie raid of 1983, intensifying with the reentry of Taylor in 1989, and continuing at least until the 2003 ouster of Taylor. However, row 1 of table 1 reports only three isolated years of civil war, 1990, 1992, and 2003. (Furthermore, all of these years for Liberia are reported as type 3 conflicts despite the involvement of other African countries in the dynamic of this violence; see the next section on the regional question.)

4. This problem of spotty coverage is mitigated by a more lenient coding rule that is presented in row 2 of table 1 based on fewer deaths (PRIO/Uppsala intensities 1–3) and is the mainstay of the MGB studies. However, a new problem appears. There are country-years that are reported as having instances of so-called civil war under the MGB rules but are unknown to us: Burkina Faso (1987), Burundi (1965), Cameroon (1984), Equatorial Guinea (1979), Ethiopia (1960), Gabon (1964), Gambia (1981), Ghana (1966, 1981, and 1983), Guinea (1970), Kenya (1982), Lesotho (1998), Liberia (1980), Nigeria (2004), Somalia (1978), Togo (1986 and 1991), Tunisia (1980), and Uganda (1971, 1972, and 1977). Sixteen of these 22 instances are denoted as “military faction” in the PRIO/Uppsala documentation “Armed Conflicts 1946–2004,” and 17 instances appear in the list of 188 coups by McGowan (2003, appendix A). An 18th, Tunisia, is outside of McGowan’s sub-Saharan database. The data thus seem to report other types of isolated violent episodes. Moreover, anything termed a civil war with a duration of only one year seems suspicious. For example, the violent episode in northern Nigeria that produced the 2004 entry was one of many in Nigeria’s violent post-1970 history. This conflict appears in the data as an intensity 1–3 event (25 or more deaths) and thereby assumes the same civil war status in an MGB-type coding as the years when Biafra attempted to secede from the Federation. Other instances in northern Nigeria are the Maitatsine risings in the early 1980s. None of these many violent episodes threatened the effective monopoly of force of the Nigerian state, which repressed them brutally.

5. Casual inspection of the data in table 1 shows that there is clustering of the years of civil war produced by the death-based coding rule. This property is true regardless of whether the intensity is “war” (row 1), with its spotty coverage, or the merger of intensities 1–3 (row 2), although it is more pronounced in the latter case. This observation does not require the support of formal statistical analysis. This feature of the data is expected because almost any classification of country-years into civil war or its absence must include the sustained nature of these conflicts. Following the third point above, if country narratives establish that the gaps between years of civil war in table 1 are also years of civil war, the autocorrelation of these series would increase because additional years of civil war would be followed by civil war, and likewise for years of peace. Trite as these observations may be, econometric studies of civil war have ignored their implications.

Proper regard for the statistical properties of the errors in the relationship between the dependent variable and the explanatory variables is key to reliable econometric inferences. The autocorrelation of the dependent variable, whether a country had a civil war in any given year, which is evident in table 1, is not a sufficient condition for the errors for each country in an econometric equation explaining whether a country had a civil war to be correlated across years. It is, however, highly suggestive of autocorrelation in the errors and requires that econometric analysis pay special attention to this possibility. However, econometric studies of whether countries have civil wars have not tested for this property of the error terms or corrected for autocorrelation. Two problems ensue if autocorrelation in the error terms exists but is ignored. First, levels of statistical significance reported in econometric studies are likely to be exaggerated, making it seem that inferences are reliable when they are actually the result of chance. Second, if the specification of the equation explaining civil war includes a variable defined as whether the country had a civil war in the preceding year, inferences will be biased if error terms are autocorrelated (the case of autocorrelated errors and a lagged dependent variable). Furthermore, we do not see autocorrelation as arising from some unmeasurable factor that is common to the civil war experience of all countries, such as the end of the Soviet Union’s role in Africa. Rather, we see persistence in the dependent variable, whether there is a civil war, as arising from slow-moving, unmeasurable factors specific to each country’s experience. Consequently, the problems posed by autocorrelation cannot be overcome by a set of controls common to all countries (year dummies, in econometric terminology), as would be justified in some econometric applications.

## Among Large Groups of Inhabitants or Citizens of a Country

So far, we have referred to episodes of large-scale violence as civil wars, in accordance with econometric studies. Now, we want to distinguish between civil wars and regional war complexes. The reasons for this distinction are twofold. First, the concept of a regional war complex emphasizes the interdependence of the large-scale violence experienced by African countries. This interdependence is the source of the presumed correlation across countries in the dependent variable and associated error terms in the econometric models and is a severe problem for this research strategy. Second, the regional war complex is an essential building block for further exploration of large-scale violence in Africa through narratives. Basically, it provides a way of distinguishing the experience of Nigeria from 1967 to 1970 as a civil war from the experience of the countries in the five main regional war complexes in Africa that we identify and discuss in this section.

Central to our distinction between civil wars and regional war complexes is the extent to which “foreigners,” who are neither inhabitants nor citizens of a country, are key participants (whether as providers of fighters, arms, bases, or other forms of assistance) in the violent conflict. A pivotal element in our definition of civil war—and one that is shared, as noted earlier, by otherwise disparate definitions—is that the principal participants are inhabitants or citizens of the country experiencing the violence. That is, “domestic” participation (of inhabitants or citizens) is high, and “foreign” participation is low. In contrast, a regional war complex has high foreign participation, and domestic participation inside at least one of the countries involved in the violent conflict must be high enough to challenge the government’s monopoly of force in that country. Regional war complexes must also be distinguished from international or interstate wars. Again, the critical dimension is the degree of foreign versus domestic participation. In an interstate war, foreign participation in the country experiencing the violence is high, and domestic participation is low. If the inhabitants or citizens are weak proxies or auxiliaries or are followers on the coattails of others, then the phenomenon is an international war. Distinguishing what constitutes high and low foreign versus domestic participation in a country experiencing violence requires judgments based on the historical narratives of large-scale violence. The challenge of delimiting the boundary between intra- and interstate wars (see [Modelski 1964](#), 42; [Cramer 2007](#), 70), which is considered by some to be a recent phenomenon ([Levy 2007](#)), inheres in our concept of a regional war complex.

We offer a tentative definition of a regional war complex that parallels our definition of a civil war: A regional war complex is a politically organized, large-scale, sustained, physically violent conflict among large groups of inhabitants or citizens of more than one country, and in at least one of these countries, at least two groups of inhabitants or citizens must be bona fide contenders for the monopoly of physical

force. One way to think of the threshold of bona fide contenders is that each of these groups must be mounting efforts comparable to what would be expected of participants in a civil war. Groups that engage in violent conflict with each other but do not contest the state, which interest some researchers (e.g., [Williams 2011](#)), fall outside our purview. Most often, violence occurs in more than one country, and these multiple violent conflicts overlap or interlock and thus mutually reinforce, in practice, a distinctive attribute of a regional war complex.

The concept of regional war has been addressed explicitly, but much less frequently than the concept of civil war ([Gleditsch 2007](#)). We highlight a few key differences between two of these conceptions and our own definition. [Wallenstein and Sollenberg \(1998\)](#) find that more than half of the armed conflicts in the world between 1989 and 1997 were linked to conflicts in neighboring states. Merely on the basis of this simple count, these authors suggested the importance of “regional war complexes.” They do not require, as we do, that at least one of the linked conflicts have two groups of citizens or inhabitants contending for the monopoly of force. Recently, [Rubin et al. \(2001\)](#) advocated the concept of a regional conflict formation to understand the conflict in the Great Lakes region and to contribute to conflict management in this region. [Rubin et al. \(2001, 3\)](#) define a regional conflict formation as “a set of transnational conflicts that form mutually reinforcing linkages with each other across state borders.” These authors consider regional conflict formations to be armed conflicts that are not local, civil, or intrastate and that are contemporary phenomena. Our notion of a regional war complex differs from that of Rubin et al. in at least two ways. First, we require at least one country in the complex to have two groups of its citizens or inhabitants contending over the monopoly of force. Second, as discussed below, we believe that regional war complexes are not only contemporary but also historical phenomena.

Although the concept of regional wars has not been widely used, some of the dynamics that characterize such wars have been analyzed for many decades. At least since the 1960s, scholars have examined the role of external influences, including neighboring countries, on internal political violence ([Rosenau 1964](#); [Midlarsky 1992](#); Zartman originally published 1966, reissued 1989). [Scott \(1964\)](#) described the dynamics of how poor states that lack conventional capabilities, such as resources, population, and military power, can foment internal violence in other countries without great expense, thus destabilizing international politics. [Zartman \(1987, 1989 and 1992\)](#) described the high incidence of state propaganda and financial and military support of opposition groups intended to weaken the target state and the broader regional relations of neighbors and more distant states. In studies of civil wars, many case studies discuss the theme of regional influences on a particular civil war but do not delve into the dynamics that lead to these regional influences, which lie outside the scope of a single case as defined by these authors. For instance, studies of Chad ([Azevedo 1998](#); [Nolutshungu 1996](#)) refer to Libya’s

role but do not consider the situation from the Libyan side. There have been many efforts to examine systematically the role of neighboring countries in intrastate wars. [Brown \(1996\)](#) argues that most internal violent conflicts are caused by the actions of internal elites. On the few occasions when external forces trigger violent internal conflicts, the “bad” actions of neighboring states play a more prominent role than mass-level or spillover phenomena. [Woodwell \(2004\)](#) discusses some aspects of ethnic commonalities among neighboring countries and their probabilities of experiencing civil wars. Since the 1960s, scholars of regional subsystems have highlighted the importance of geography in defining patterns of conflict and cooperation that characterize contiguous states ([Khadiagala 2006](#)). The more recent literature on regional security systems discusses how even noncontiguous states may be members of a regional system ([Lake and Morgan 1997](#)). [Salehyan \(2009\)](#) is interested in analyzing endogenous regional conflict clusters formed by the intrinsic links between civil wars and international conflicts. The bridge between the two types of mutually constituted and mutually reinforcing wars is provided by transnational rebel sanctuaries, which generate international hostilities and empower rebels in their struggles against their own state ([Salehyan 2009](#), 50).

Our distinction between a civil war and a regional war complex is superficially similar to that made by PRIO/Uppsala between “internal armed conflict” and “internationalized internal armed conflict.” Our concepts and those of PRIO/Uppsala share a recognition of the importance of external involvement as a factor that has the capacity to alter thinking about these experiences of large-scale violence. Our concept differs in practice from the concept implemented by PRIO/Uppsala because we consider much more of the African violence to be part of regional war complexes, as argued below in our discussion of table 1. By contrast, users of these data, such as MGB, lump both of these categories of conflict together under the heading “civil war.”

In Africa, the episodes that are popularly termed civil war and that are coded as such in the econometric studies are often not predominantly internal conflicts. The Nigerian Civil War over Biafra’s attempted secession from the Federation is an exception to most so-called civil wars in Africa in many ways and thereby provides an instructive case (coded entirely and appropriately by Xs in table 1). Its inception was an entirely domestic affair. Following political troubles in the Western Region, two bloody coups and pogroms against Ibos outside the Eastern Region and their flight home, the Ibo-dominated Eastern Region attempted to secede as Biafra from the spring of 1967. The war involved fighting by massed armies commanded by officers, many of whom had been former colleagues in the federal army, just as the opposing officers had in the American Civil War, and with every other attribute of structured combat. Biafra sustained itself for over three years of organized combat in the form of offensive and defensive campaigns and with plausible prospects of success at different junctures. Both belligerents maintained essentially all the de

jure and de facto attributes of states, a capital, a civil administration, international representation, a currency, and postage stamps. The Nigerian war ended with a formal military surrender, similar to Appomattox, on January 12, 1970.

Foreign intervention was not absent from the conflict. Although combat was almost exclusively by nationals of what had been the Federation of Nigeria, arms supplied by non-African countries to each side were certainly important. The United Kingdom and Soviet Union supplied the Federalists, and France supplied Biafra, although Biafra's own purchases from arms suppliers were probably more important. Furthermore, Portugal, Côte d'Ivoire, and Gabon assisted in the airlift of arms to Biafra and other countries, including Tanzania, and Zambia recognized Biafra (De St. Jorre 1972; Stremmlau 1977).

Nigeria is the most compelling (and, arguably, the only) civil war case in post-independence sub-Saharan Africa. The war began because of domestic conflict, and two large groups of Nigerians contested the integrity of the country. The main foreign participants provided arms (and relief) to domestic actors, which affected the dynamic of the conflict. Ultimately, however, the war remained primarily a contest between domestic actors and was resolved by internal factions. In short, foreign participation was low, and domestic participation was high. Moreover, foreign participation did not result in an interlocking regional conflict. Indeed, the most involved foreign actors were not even neighbors of Nigeria.

The exact identification of the regional war complexes, their membership, and timing is a demanding research effort because it involves piecing together many narrative accounts and making judgments on the importance of outsiders' involvement. However, we believe that a tentative list of the major war complexes and their members is feasible by utilizing commonly available knowledge. We identify the following complexes in Africa and their member countries within Africa: (1) Mano River (Burkina Faso, Côte d'Ivoire, Guinea, Liberia, Libya, Sierra Leone, and Senegal); (2) Chad (Central African Republic, Chad, Libya, and Sudan); (3) The Horn (Djibouti, Eritrea, Ethiopia, Kenya, Somalia, now in its several parts, and Sudan) (4) Great Lakes (Angola, Burundi, Central African Republic, Chad, Congo-Brazzaville, Congo-Kinshasa, Libya, Namibia, Rwanda, Sudan, Tanzania, Uganda, Zambia, and Zimbabwe); (5) South (Angola, Botswana, Lesotho, Malawi, Mozambique, Namibia, South Africa, Swaziland, Tanzania, Zambia, and Zimbabwe).

Regional war complexes and their participants are diverse. From our identification of the members of these complexes, it is evident that a country (e.g., Sudan) can participate in more than one complex. In some cases, a complex may begin as a civil war or two or more independent civil wars and then merge into a regional war sustained by important interactions among the participants. It may even begin as an international war and become a regional war complex. In general, ending some conflicts in a complex may disrupt the links that began or fueled wars and thereby may initiate a regional peace process. This dynamic is one measure of the

interdependence of these episodes of violence and a practical consequence of thinking about them in this way because it can help to design strategies for establishing peace.<sup>2</sup>

Some members of these complexes contribute importantly to violence in other African countries without experiencing violence as a consequence. Thus, Qadhafi's Libya played an extensive role in the violence in Chad, in the Mano River complex, and even (highly unsuccessfully) in the Ugandan/Tanzanian war that overthrew Amin (Lemarchand 1988; Huliaras 2001; Solomon and Swart 2005). Qadhafi's Libya is notable for the ambition of its reach into countries with which it is not contiguous. Except for one I in table 1 when it was active in combat in Chad, its activities are coded with an E, but this information does not identify where Libya was intervening. Burkina Faso's involvement in the Mano River complex is another example of an important player that has thus far avoided internal violence. Except for 1987, when it experienced a coup that seems to be the source of the X, Burkina Faso does not appear in table 1. According to the coding rules, it is at peace as much as any country in the sample. In addition to African countries, other countries that play this role are outside of Africa, including Cuba, France, the United States, and countries acting under the aegis of the United Nations, which have sent ground troops to participate in African conflicts.

Most of the countries that participate in these complexes are involved in reciprocal interactions, however, in which one country supports antagonists in a neighboring country and the neighbor responds in kind. Such pairs include Sudan and Uganda, Angola and Congo-Kinshasa, and Liberia and Côte d'Ivoire, among many others. Yet, the PRIO/Uppsala data record few if any of the international activities of Burkina Faso, Liberia, Libya, Sierra Leone, Côte d'Ivoire, or Somalia, among others. This lack of information means that econometric studies based on these data cannot incorporate such intercountry relationships even if they use specifications based on neighborhood effects, as in the spatial econometric literature (see below).

Côte d'Ivoire has been a long-time member of the Mano River complex, influencing the dynamics of the complex well before it experienced associated violence and, in particular, providing a launching place for Taylor's invasion of Liberia in 1989. From the coup attempt of September 19, 2002, until at least 2011, it has experienced a violent contest over the monopoly of force within its own borders. This contest is tied to the dynamics of the Mano River war complex (International Crisis Group 2003, 2004, 2007; Banégas and Marshall-Fratani 2007, especially 99).

The coup plotters in 2002 were Ivorian soldiers from the north of Côte d'Ivoire who were given shelter in Burkina Faso, from where they launched their coup. In September 2002, a French (Licorne) force separated the north from the south, preserving the Gbagbo government and thwarting the coup leaders, who were left controlling the north of the country. In November 2002, revolt broke out in the west of the country to the south of the French-maintained line of (temporary) partition.

The situation in the west involved the participation of Ivorians and Liberians in each other's conflicts. The United Nations provided peacekeeping forces in February 2004. In November 2004, the French forces destroyed the Ivorian air force of the Gbagbo government in retaliation for the killing of peacekeepers and an aid worker by the air force. Finally, in 2011, French and United Nations forces intervened in the postelection violence consequent on Gbagbo's refusal to accept the election results of November, 2010. Thus, foreign forces from noncontiguous countries played important roles in this conflict.

These events have their origins long ago, however, in the colonial and postcolonial economic strategies of the Ivorian economic miracle with its focus on labor movement from outside and from within Côte d'Ivoire to its tree-crop producing areas. In the receiving areas, especially the southwest, this movement led to potential conflict between the established people and the newcomers, especially over land rights. There are two regional consequences. First, any dynamics within Côte d'Ivoire that threatened immigrants, especially the numerous Burkinabés, would have implications for the neighboring governments in protecting their citizens abroad and in dealing with potentially large return migrations of refugees. Second, dissatisfaction and unrest between the groups of established and recently arrived citizens within Côte d'Ivoire would provide opportunities for intervention from neighboring Liberia, especially because ethnic conflicts in Liberia mirrored divisions in Côte d'Ivoire. Because recent events proceed from these deeply structural factors that operate on a regional level and because the particulars of the conflict that began in Côte d'Ivoire in 2002 involved specific violent interventions by neighboring countries in each other's affairs, we designate the Ivorian experience as part of the Mano River complex.

PRIO/Uppsala, however, do not code for the many interactions among countries evident in this history. Therefore, an MGB-type coding also misses them. PRIO/Uppsala only code Côte d'Ivoire as experiencing conflict at the "minor" intensity level and only beginning in 2002, thereby ignoring its participation in the violence of the Mano River complex before 2002. This coding also neglects the important roles of outsiders in the violent Ivorian experience beginning in 2002: there is no recognition in the coding of the role of Burkina Faso in hosting soldiers from northern Côte d'Ivoire who engaged in a coup and the subsequent partition of the country. The distant French and the contiguous neighbor Liberia were implicated in the post-2002 Ivorian conflict, but this involvement is not coded either.

We consider Mozambique's conflict between Renamo and the Frelimo government (Vines 1991; Finnegan 1992; Hume 1994) part of a regional war complex rather than a civil war, as MGB's econometric studies do (table 1).<sup>3</sup> To begin to understand this post-independence conflict, one must acknowledge its inextricable links to the internal and external politics of the wars to decolonize Mozambique and Rhodesia. During the struggle for independence in Mozambique, Frelimo invited

Mugabe's ZANLA army to establish military bases and refugee camps in Frelimo-occupied territory. Rhodesians retaliated by attacking the camps and bases, thus regionalizing the conflict. After Frelimo became the government in Mozambique, the Rhodesian government continued to attack ZANLA bases and attempted to destabilize the Frelimo government by providing financial support to Renamo rebels. Renamo itself drew in part upon people who had been marginalized in political conflicts within Frelimo during and after the struggle for independence as well as upon those who opposed Frelimo's policies. After Mugabe's party formed the government of Zimbabwe, the South African government took over the training, financing, and arming of Renamo rebels. The Frelimo government provided bases to African National Congress guerrillas who infiltrated South Africa through southern Mozambique. Mugabe provided military support to the Frelimo government, his former liberation war ally. Renamo retaliated against Zimbabwe's intervention by attacking villagers living close to the Mozambican border. The war between Renamo and Frelimo ended in an externally mediated peace agreement in 1992 that was possible, in part, because South Africa and Zimbabwe ceased to intervene on behalf of their respective allies. Thus, regional dimensions are crucial to understanding the origins, dynamics, and ending of large-scale violence in Mozambique. Keller (1997) correctly views the domestic conflict in Mozambique as having contributed to regional insecurity in southern Africa. However, as we have shown, the "domestic" conflict in Mozambique was itself an inextricable product of regional interventions. This conflict is best viewed as a component of a regional war complex rather than a civil war or a domestic conflict with regional dimensions.

Labeling the war between Renamo and Frelimo in Mozambique a "civil war," as MGB do, is therefore a misnomer. Foreign participation in sponsoring, training, and equipping Renamo, first by the Rhodesian government and then by the South African government, played a vital role in making Renamo a serious contestant for the monopoly of force. However, Renamo was not a mere proxy of its foreign benefactors because it had substantial domestic support. South African support for Renamo and Frelimo support for the African National Congress represent examples of "interlocking" conflict, as did Zimbabwean support for Frelimo and Renamo attacks on Zimbabwean civilians on the eastern border. Violence was not contained inside Mozambique; it also occurred within Zimbabwe and South Africa. Finally, the disruption of regionalized links—South African and Zimbabwean support for their respective allies—facilitated the peace.

Another example of interdependent violent conflict and its consequences for peacemaking comes from the Sudan–Uganda interaction. Uganda supported southern opponents of the Khartoum government after Khartoum began to support opponents of the Uganda government in northern Uganda in 1986 and subsequently within Sudan (Prunier 2004). With the Comprehensive Peace Agreement of January 2005 that ended the war between southern Sudan and the Khartoum

government, security problems in northern Uganda subsided, although the Lord's Resistance Army opponents of the Kampala government remain active in parts of Congo-Kinshasa, the Central African Republic, and southern Sudan. From an econometric perspective, to the extent that conflicts end because regional conflicts wind down, there is correlation in the outcomes and, presumably, the error terms among the countries in samples used to study the duration and ending of civil wars.

Congo-Kinshasa (former Zaire and henceforth Congo, when the context is clear) is a major participant in another regional war complex that we call the Great Lakes complex. Table 2 illustrates how four datasets used by econometricians all classify Congo in the 1990s as having experienced civil war. The MGB studies code a civil war based on the concept of "internationalized armed conflict" from 1997 to 2000 as intensity 3 and that from 1996 to 2001 as intensity 1–3. Their coding rule does not indicate whether the years coded as civil war constitute one or more such wars. The other two datasets code two civil wars each. [Fearon and Laitin \(2003b\)](#) code one from 1996 to 1997 and another from 1998 to 1999. [Doyle and Sambanis \(2006\)](#) also code one civil war from 1996 to 1997 and another from 1998 to 2001. We prefer to view the conflict in Congo during the 1990s and into the 21st century as a regional war complex with two major phases rather than as one or two civil wars. We only hint below at the multiple layers of actors, alliances, and interests involved in this complex. We believe that this complex has existed since many of its member countries attained political independence, but we cut into its dynamics only from the mid-1990s.

The war in 1996 to 1997 was fought primarily to remove Mobutu Sese Seko, dictator from 1965 to 1997. Internal opposition forces were too fragmented to remove Mobutu, but he was widely disliked inside the country. Many African leaders had considered joint military intervention to remove Mobutu, whom they viewed as an embarrassment and an imperialist stooge. Rwanda's postgenocide government unilaterally took the initiative, followed soon thereafter by Uganda. These were the major players in the violent conflict that finally removed Mobutu and installed Laurent Kabila in May 1997.

Rwanda's interests were chiefly security driven. The 1994 Rwandan genocide, in which extremist Hutus killed Tutsis and moderate Hutus, was followed by the taking of power in Rwanda by Tutsis, mostly former refugees who had been in Uganda. In turn, many Hutus, including extremists, fled as refugees into the eastern parts of Congo abutting Rwanda. The extremists took control of the Hutu refugees in Congo. These extremists, who first wanted to retake and then merely to destabilize Rwanda, launched attacks from Congo. The Rwandan government wanted to defeat them and end this threat on its borders.

Uganda joined Rwanda's effort to oust Mobutu primarily because Mobutu had allowed Ugandan rebels to operate against President Museveni's government from the northeastern corner of Congo. These rebels were aided by Museveni's

longstanding enemy, the government of Sudan, which hoped to use the Ugandan rebels not only to destabilize Uganda but also to Islamize Ugandans. Uganda had close ties to the rebels in southern Sudan who fought the Sudanese government.

Uganda and Rwanda joined forces with a weak Congolese rebel force led by Kabila. Some analysts believe that the military role of the Congolese rebel force was nominal, and its external formation primarily served to conceal the Rwandan and Ugandan invasion (Turner 2007; Lemarchand 2009; Nzongola-Ntalaja 2002; Kennes 2005; Prunier 2009). Other analysts credit the military role, leadership, and initiative of the Congolese rebels and refer to military support from Kabila's Rwandan, Ugandan, and other allies (Ndikumana and Emizet 2005; Zartman 2005; Baregu 2006; Adebajo 2006) or seem unconcerned by the nationality of "ethnic Rwandans" (presumably meaning Tutsi) who "heavily dominated" the Congolese rebel force (Reed, 1998). Whether the Congolese rebels are portrayed more as "clients" or "proxies" of their external patrons or more as genuine partners or allies, together they drove Mobutu from power with little resistance. Kabila became president and initially kept Rwandan officers in charge of his new army.

The war continued—or, as most analysts claim, another war began in 1998—when President Kabila ordered the Rwandan troops in Congo to return to Rwanda. The Rwandan army, the Ugandans, and Congolese rebel forces responded by acting to remove Kabila. Rwanda and Uganda no longer trusted Kabila to serve their security interests, and the otherwise disparate Congolese rebels were united solely by a desire to seize power themselves. Analysts again are divided between those who see the Congolese rebel organizations as "proxies" or "clients" of the Rwandans and Ugandans who created, financed, equipped, provided personnel for, and largely controlled the organizations (including the appointment of their leaders) that served as a cover for their own interests (Lemarchand 2009; Prunier 2009; Kennes 2005; Nzongola-Ntalaja 2002; Ndikumana and Emizet 2005; Tull 2003, 2007) and others who consider the Congolese rebel organizations to have merely enjoyed external support from their foreign allies (Zartman 2005; Baregu 2006). Jean Bemba's organization is typically considered to have developed autonomy from his Ugandan sponsor.

Analysts debate why Rwanda and Uganda invaded Congo in 1998. This discussion of motives highlights the regional dimensions and mutually interlocking dynamics in the Great Lakes complex. Some give primacy to their continued security interests (Tull 2007), others emphasize their aggressive quest for access to and control of Congo's rich resources (Baregu 2006), and still others consider the security concerns expressed by Rwanda and Uganda to be mere pretext for access to Congo's natural resources (Nzongola-Ntalaja 2002). Turner (2007) argues that Rwanda, Uganda, and Burundi intervened in 1998 to secure their respective western borders against rebel movements based in Congo *and* to gain access to Congo's resources. Many observers argue that security concerns were the initial

primary driving force behind Rwanda's 1998 invasion, but economic exploitation of Congo's resources became the main (albeit unacknowledged) reason (Reyntjens 2006; Lemarchand 2009).

For some (Kennes 2005; Zartman 2005), the fundamental stakes in both wars were the struggle for regional hegemony in the power vacuum created by Mobutu's loss of control. Kabila's key military allies were Zimbabwe, Angola, and Namibia. Chad was a reluctant participant, responding to the request of its ally Sudan to provide troops that were transported to Congo on Libyan planes (Prunier 2004). Some emphasize the pan-Africanism of Kabila's three key military allies (Baregu 2006), and others emphasize their different interests. Some stress the economic motives of Zimbabwe (Prunier 2006; Nzongola-Ntalaja 2002; Lemarchand 2009), whereas others point to Mugabe's rivalry with South Africa for regional hegemony (Kennes 2005). Most analysts highlight Angola's security concerns to prevent its rebel opponent, UNITA, from reestablishing bases in Congo (Nzongola-Ntalaja 2002; Lemarchand 2009; Prunier 2009), whereas Namibia is seen as either following its patron, Angola (Prunier 2006 and 2009), or as having an interest in the water and hydroelectric resources in Congo and security concerns to prevent UNITA from destabilizing Angola (Nzongola-Ntalaja 2002).

The discussion above merely sketches the most important players in the violence that engulfed Congo beginning in 1996. It omits how states such as the Central African Republic, Congo-Brazzaville, Burundi, Tanzania, Zambia, Kenya, Gabon, and South Africa directly or indirectly played a role in the war(s) in Congo, as did many armed and unarmed nonstate actors. Several members in the complex are not contiguous with Congo. South Africa and Zambia played a major role in brokering peace in Congo, beginning with accords in 1999. However, eastern Congo continued to be beset with violence even after peace had been restored elsewhere in the country under Joseph Kabila, who became president after his father's assassination in 2001.

Unlike a number of scholars who use an econometric approach (table 2), analysts using a narrative approach are divided on whether the war or wars in Congo (1996–2002) had a civil war component. Those commentators, such as Kennes (2005), Turner (2007), Prunier (2009), and Lemarchand (2009), who deny the Congo war or wars the status of civil war, consider the conflict to have been fought primarily by foreigners hiding behind weak internal proxy forces that lacked the capability to sustain a civil war. The foreigners, as discussed, were both neighboring states and more distant states as well as nonstate actors who were in Congo to wage war against their home governments.

In fact, if the years from 1996 and afterward are to be coded as anyone's civil war, it might make more sense to code them as a continuation of Rwanda's rather than Congo's (Reyntjens 2006, 30). The conflict between certain Tutsi and Hutu groups within Rwanda spilled into Congo. Similarly, Angola's experience was more

of a civil war than Congo's, and it too spilled over into Congo. Thus, both of these countries' experiences meet the requirement of our definition of a regional war complex, with at least one country having a contention for the monopoly of force between domestic groups that proceeds at high intensity. At the same time, the experience of the Great Lakes complex as a whole is one of interlocking and reinforcing conflicts within and among many countries, which is why we consider it a regional war complex.

The Congo conflict underscores the ongoing importance of regional conflicts involving state and nonstate armed actors in understanding mass violent upheaval in Africa. The Congo conflagration also sharply raises the questions of how one defines a civil war and how one distinguishes between civil wars fought with regional support and foreign invasions fought with foreign-controlled proxies. We do not count the violent conflicts between Congolese forces that occurred in Congo (1996–2002) as civil wars because of the preponderant role of foreign forces fighting in Congo and the associated violence outside Congo's borders. Importantly, the interconnected episodes of violence in 1996–1997, 1998–1999, and thereafter demonstrate the difficulty of determining when an old war ends and a new war begins.

Our narratives establish important interdependencies within groups of countries in these contestations over the monopoly of force. In turn, such interdependencies pose problems for econometric research on civil war. Current research on civil wars does not correct or adequately test for these possibilities, creating a set of potentially serious misspecifications that may lead to an overstatement of statistical significance of, and even bias in, parameter estimates. These interdependencies among countries regarding whether a country has a civil war can enter econometric specifications in several ways.

The first (and most simple) way is if variables based on the characteristics of other countries enter the list of explanatory variables, such as information on the religions or ethnicities in neighboring countries. Such variables pose no additional econometric problems relative to a variable based on a country's own characteristics.

The second way is if there is correlation among the errors in explaining whether countries have civil wars or if the list of explanatory variables includes whether other countries, perhaps contiguous ones, are experiencing civil war, or both. These formulations pose estimation problems. For instance, if country A's outcome (whether it has a civil war) partially determines B's outcome, then the error term for A influences the outcome for B. However, B's outcome also influences A's outcome. In other words, some of the explanatory variables (the other countries' outcomes, say, B's) are correlated with the error in the equation for the country under consideration (say, A). Special estimation techniques can avoid biased estimators if the intercountry relationship is simple, such as spillovers between contiguous neighbors (LeSage and Pace 2009). The narratives, however, suggest that

influences of other countries on outcomes in a particular country are more complicated than easily permitted by econometric specifications. For instance, Libya was highly active in fomenting violence but had experienced almost no violence of the sort recorded in Table 1 prior to its own civil war of 2011, and it is not a contiguous neighbor of the countries that experienced violence in two of the complexes in which Libya has played a role. Beck et al. (2006) point out that problems do not arise if the explanatory variable is whether another country had a civil war in the preceding rather than the current year, but their point is only relevant if the errors are not correlated over time. As argued in the preceding section, however, there is reason to expect that errors are autocorrelated when data look like they do in table 1.

The third way is if there are both correlations over time and across countries, as we have argued. For instance, the value of the lagged dependent variable for the country in question, whether it had a civil war in the preceding year, and the contemporaneous values of the dependent variable for other countries, whether they had civil wars, may be among the variables that explain the dependent variable of each country. In addition, the errors in the explanatory equations for each country may be correlated over time and across countries.<sup>4</sup>

Among the limited number of studies that have attempted some type of spatial analysis of what they consider civil war are Fearon and Laitin (2003a), Gleditsch (2007), and Buhaug and Gleditsch (2008). The latter two are highly ambitious in their econometric procedures, although they do not adopt a full-fledged correction for correlation over time and across countries. Even if one persisted and somehow estimated a model that was flexible enough to encompass (nest) the pervasive correlation over time and space that must be presumed to characterize these countries' experiences, it is implausible that precise findings would emerge regarding the processes at work. The implied likelihood functions would have too many parameters, and the numbers of observations would be too few. Available results from incomplete implementations of the more encompassing specifications establish an econometric presumption that spatial interdependencies cannot be ignored. This conclusion is supported by Buhaug and Gleditsch (2008, 227), who study the onset of civil war and find that their fixed-effect estimator yields mostly insignificant results, nonetheless an indicator of whether there is conflict in neighboring countries retains significance. Similarly, in their reanalysis of Miguel et al. (2004), Jensen and Gleditsch (2009, 368) demonstrate that the key finding of the former study on the role of rainfall in predicting civil wars (through its effect on income) is attenuated, in some cases to the point of statistical insignificance, by the inclusion of spatial effects in the explanatory equation. This result is consistent with a conclusion that rainfall is just a proxy for regional effects. The presumed importance of such regional effects, supported by the Buhaug-Gleditsch and Jensen-Gleditsch studies, is exactly the conclusion of our narratives.

These spatial interdependencies also have implications for authors such as [Goldstein and Pinker \(2011\)](#), who attempt to conclude that there is a decline in the prevalence of wars. Because many of these wars form complexes, the countries that are involved may tend to return to a peaceful status at nearly the same time. If it were assumed that the end of violence in a number of countries represented independent events, it might make sense to infer a trend toward peace and to look for its causes. However, if the end of violence in a number of countries is not independent, then there is not a sufficient number of observations to make confident inferences. For this reason among others, the type of inferences that [Angell \(1910\)](#) made about the prospects for peace are always risky.

Fundamentally, the interdependencies among countries raise the question of whether the unit of observation is the individual country or the regional war complex. Econometric studies seem to struggle to incorporate these interdependencies. By contrast, an analysis based on the narratives can easily focus on the regional complex as the unit of observation.

## Conclusions

As we initially stated, we believe that a list of country-years classified into civil war or membership in a regional war complex or the absence of these situations requires a detailed compilation of country histories based on the case-study literature. This task is feasible for the five regional war complexes we have identified as well as others elsewhere in the world. We believe that the outcome will be the documentation of regional war complexes rather than properly defined civil wars. This task is large, and this paper begins to clear the way by attempting to craft definitions that distinguish civil wars from regional war complexes and to show how to use these definitions. The periodization of civil wars and regional war complexes must depend on many judgments and their justifications. Important aspects of these justifications will be the level and nature of violence and whether violence spills across borders, the extent of participation, the organization of the parties to the conflict, the purposes of the conflict (both stated by the participants and inferred), the plausibility of contention over the monopoly of force, the duration of the conflict, the freedom of movement in the country, the ability of a government to fulfill functions such as revenue collection, and the role of one country in the affairs of others. These are some of the relevant components of Wittgenstein's complicated network of definitional similarities for large-scale violent political conflict, whether a civil war or a regional war complex, that will set the context for the analysis of cause and consequence.

Regardless of the relative difficulties of producing narratives for a complex, we have stated why the econometric literature fails to provide a convincing analysis of

these episodes of violence in its own terms, so there is no solution here. For these reasons, we believe that the energy of researchers should turn to a comparative analysis of regional war complexes based on narratives extracted from the case-study literature.

## Notes

Mark Gersovitz is Professor in the Department of Economics, The Johns Hopkins University; email address: gerso@att.net. Norma Kriger is an Independent Consultant currently on contract at the Federal Research Division of the Library of Congress. We thank Timothy Besley, Henry S. Bienen, Richard H. Spady, and Tiemen M. Woutersen for their valuable comments.

1. Of course, an updated version of the data set would record a civil war in Libya for 2011, in which Qadhafi fell, albeit with considerable, perhaps determinative, help from NATO.

2. King (2007, 126) writing about civil wars understood as large-scale violence within the boundaries of a single state, notes that peace agreements “that build in a regional dimension – seeing the organizations of violence, in effect, as extending well beyond the borders of the state concerned – have now become a critical component of successful war termination.” Building regional dimensions into conflicts conceptualized as “civil wars” risks underemphasizing how fused the experiences of groups of these countries have been.

3. Table 1 codes the years of war in Mozambique as either an internal armed conflict or an internationalized armed conflict. An MGB coding rule would term them, collectively, years of civil war.

4. Driscoll and Kraay (1998) and Pesaran and Tosetti (2011) provide methods for dealing with errors that are correlated across both countries and years without placing restrictions on the form of the correlations. Driscoll and Kraay (1998) show the possibility of a large overstatement of statistical significance if such methods are not adopted. However, the methods of these authors do not directly apply to the case in which temporal and spatial lags of the dependent variable (whether there are civil wars in other countries or in earlier years in the country under consideration) may appear as explanatory variables.

## References

- Adebajo, A. 2006. “The United Nations.” In *Security Dynamics in Africa’s Great Lakes Region*, ed. G. Khadiagala, 235–68. Boulder, CO, and London: Lynne Rienner Publishers.
- Angell, N. 1910. *The Great Illusion: A Study of the Relation of Military Power in Nations to Their Economic and Social Advantage*. New York: Putnam’s.
- Atlas, P. M., and R. Lickliger. 1999. “Conflict Among Former Allies After Civil War Settlement: Sudan, Zimbabwe, Chad, and Lebanon.” *Journal of Peace Research* 36:35–54.
- Azevedo, M. J. 1998. *Roots of Violence: A History of War in Chad*. Amsterdam, The Netherlands: Overseas Publishers Association.
- Baev, P. K. 2007. “Defining Civil War by Examining Post-Soviet Conflicts.” *Terrorism and Political Violence* 19:247–68.
- Banégas, R., and R. Marshall-Fratani. 2007. “Côte d’Ivoire: Negotiating Identity and Citizenship” In *African Guerrillas: Raging against the Machine*, eds. M. Bøås, and K. C. Dunn, 81–112. Boulder, CO: Rienner.

- Baregu, M. 2006. "Congo in the Great Lakes Conflict." In *Security Dynamics in Africa's Great Lakes Region*, ed. G. Khadiagala, 98–131. Boulder, CO, and London: Lynne Rienner Publishers.
- Beck, N., K.S. Gleditsch, and K. Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50:27–44.
- Blattman, C., and E. Miguel. 2010. "Civil War." *Journal of Economic Literature* 48:3–57.
- Boås, M., and K. C. Dunn. 2007. "Introduction." In *African Guerrillas: Raging Against the Machine*, eds. M. Boås, and K. C. Dunn, 1–8. Boulder, CO: Lynne Rienner.
- Brown, M. E. 1996. "Introduction," and "The Causes and Regional Dimensions of Internal Conflict." In *The International Dimensions of Internal Conflict*, ed. M.E. Brown, 1–31 and 571–601. Cambridge, MA: MIT University Press.
- Buhaug, H., and K. S. Gleditsch. 2008 "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly* 52:215–33.
- Collier, P., and A. Hoeffler. 2007. "Civil War." In *Handbook of Defense Economics*, eds. T. Sandler and K. Hartley, 711–39. Amsterdam, The Netherlands: Elsevier.
- Cramer, C. 2007. *Violence in Developing Countries: War, Memory, Progress*. Bloomington, IN: Indiana University Press.
- De St. Jorre, J. 1972. *The Nigerian Civil War*. London: Hodder and Stoughton.
- Dixon, J. 2009. "What Causes Civil Wars? Integrating Quantitative Research Findings." *International Studies Review* 11:707–35.
- Doyle, M. W., and N. Sambanis. 2000a. "International Peace Building: A Theoretical and Quantitative Analysis." *American Political Science Review* 94:778–801.
- \_\_\_\_\_. 2000b. *Data Set Notes*, November 6, 2000. <http://www.worldbank.org/research/conflict/papers/peacebuilding>.
- \_\_\_\_\_. 2006. *Making War & Building Peace*. Princeton, NJ: Princeton University Press.
- Driscoll, J. C., and A. C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80:549–60.
- Eckstein, H. ed. 1964. *Internal War: Problems and Approaches*. New York: The Free Press.
- \_\_\_\_\_. 1965. "On the Etiology of Internal Wars." *History and Theory* 2:133–63.
- Edmonds, M. 1972. "Civil War, Internal War, and Intrasocietal Conflict." In *Civil Wars in the Twentieth Century*, ed. R. Higham, 11–26. Lexington, KY: University of Kentucky Press.
- Fearon, J. D. 2007. "Iraq's Civil War." *Foreign Affairs* 86:2–15.
- Fearon, J. D., and D. D. Laitin. 2003a. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97:75–90.
- Fearon, J. D., and D. D. Laitin. 2003b. "Additional Tables for 'Ethnicity, Insurgency, and Civil War.'" Stanford, CA: Stanford University.
- Finnegan, W. 1992. *A Complicated War: The Harrowing of Mozambique*. Berkeley, CA: University of California Press.
- Gleditsch, K. S. 2007. "Transnational Dimensions of Civil War." *Journal of Peace Research* 44:293–309.
- Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39:615–37.
- Goldstein, J. S., and S. Pinker. 2011, December 17. "War Really is Going out of Style." *New York Times*.
- Huliaras, A. 2001. "Qadhafi's Comeback: Libya and Sub-Saharan Africa in the 1990's." *African Affairs* 100:5–25.
- Hume, C. R. 1994. *Ending Mozambique's War: The Role of Mediation and Good Offices*. Washington, DC: United States Institute of Peace.

- International Crisis Group. 2003. "Côte d'Ivoire: The War is Not Yet Over." Freetown/Brussels: ICG.
- . 2004. "Côte d'Ivoire: No Peace in Sight." Dakar/Brussels: ICG.
- . 2007. "Côte d'Ivoire: Can the Ouagadougou Agreement Bring Peace." Freetown/Brussels: ICG.
- Jensen, P. S., and K. S. Gleditsch. 2009. "Rain, Growth, and Civil War: The Importance of Location." *Defence and Peace Economics* 20:359–72.
- Kalyvas, S. N. 2006. *The Logic of Violence in Civil War*. Cambridge, MA: Cambridge University Press.
- Keegan, J., and B. Bull. 2006. "What is a Civil War?" *Prospect Magazine*, 129.
- Keller, E. 1997. "Rethinking African Regional Security." In *Regional Orders: Building Security in a New World*, eds. D.A. Lake and P. Morgan, 296–317. University Park, PA: The Pennsylvania State University Press.
- Kennes, E. 2005. "The Democratic Republic of the Congo: Structures of Greed, Networks of Need." In *Rethinking the Economics of War: The Intersection of Need, Creed, and Greed*, eds. C. I. Arnson, and I. W. Zartman, 140–77. Washington, DC, and Baltimore, MD: Woodrow Wilson Center Press and Johns Hopkins University Press.
- Khadiagala, G. 2006. "Toward Peace, Security and Governance in the Great Lakes Region" In *Security Dynamics in Africa's Great Lakes Region*, ed. G. Khadiagala, 8–25. Boulder, CO, and London: Lynne Rienner Publishers.
- King, C. 2007. "Power, Social Violence, and Civil Wars" In *Leashing the Dogs of War: Conflict Management in a Divided World*, eds. C.A. Crocker, F.O. Hampson, and P. Aall, 115–30. Washington, DC: United States Institute of Peace Press.
- Kruger, N. 2003. *Guerrilla Veterans in Post-War Zimbabwe: Symbolic and Violent Politics 1980–1987*. Cambridge, MA: Cambridge University.
- Lake, D. A., and P. M. Morgan, eds. 1997. *Regional Orders: Building Security in a New World*. University Park, PA: The Pennsylvania State University Press.
- Lemarchand, R. 1988. *The Green and the Black*. Bloomington, IN: Indiana University Press.
- . 2009. *The Dynamics of Violence in Central Africa*. Philadelphia, PA: University of Pennsylvania Press.
- LeSage, J., and R. K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman & Hall.
- Levy, J. S. 2007. "International Sources of Interstate and Intrastate War" In *Leashing the Dogs of War: Conflict Management in a Divided World*, eds. C.A. Crocker, F. O. Hampson, and P. Aall, 17–38. Washington, DC: United States Institute of Peace Press.
- Licklider, R. 1995. "The Consequences of Negotiated Settlements in Civil Wars, 1945–1993." *American Political Science Review* 89:681–90.
- Luard, E., ed., 1972. *The International Regulation of Civil Wars*. New York: New York University Press.
- McGowan, P. J. 2003. "African Military Coups d'Etat, 1956–2001: Frequency, Trends and Distribution." *Journal of Modern African Studies* 41:339–70.
- Midlarsky, M. I. 1992. *The Internationalization of Communal Strife*. London: Routledge.
- Miguel, E., S. Satyanath, and E. Sergenti (2004). "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112:725–53.
- Modelski, G. 1964. "The International Relations of Internal War." In *International Aspects of Civil Strife*, ed. J. N. Rosenau, 122–53. Princeton, NJ: Princeton University Press.
- Ndikumana, L., and K. F. Emizet. 2005. "The Economics of Civil War: The Case of the Democratic Republic of Congo." In *Understanding Civil War: Evidence and Analysis. Vol.1. Africa*, eds. P. Collier and N. Sambanis, 63–88. Washington, DC: The World Bank.

- Nolutshungu, S. C. 1996. *Limits of Anarchy: Intervention and State Formation in Chad*. Charlottesville, VA: University of Virginia Press.
- Nzongola-Ntalaja, G. 2002. *The Congo: From Leopold to Kabila*. New York and London: Zed Books.
- Pesaran, M. H., and E. Tosetti. 2011. "Large Panels with Common Factors and Spatial Correlation." *Journal of Econometrics* 161:182–202.
- Prunier, G. 2004. "Rebel Movements and Proxy Warfare: Uganda, Sudan and the Congo (1986–99)." *African Affairs* 103: 359–83.
- . 2006. "The Economic Dimensions of Conflict in the Region." In *Security Dynamics in Africa's Great Lakes Region*, ed. G. Khadiagala, 168–202. Boulder, CO, and London: Lynne Rienner Publishers.
- . 2009. *Africa's World War: Congo, The Rwandan Genocide, and the Making of a Continental Catastrophe*. Oxford: Oxford University Press.
- Reed, W.C. 1998. "Guerrillas in the Midst: The Former Government of Rwanda (FGOR) & the Alliance of Democratic Forces for the Liberation of Congo-Zaire (AFDL) in Eastern Zaire." In *African Guerrillas*, ed. C. Clapham, 134–54. Oxford: James Currey Ltd.
- Reyntjens, F. 2006. "Governance and Security in Rwanda." In *Security Dynamics in Africa's Great Lakes Region*, ed. G. Khadiagala, 27–68. Boulder, CO, and London: Lynne Rienner Publishers.
- Rosenau, J. N., ed. 1964. *International Aspects of Civil Strife*. Princeton, NJ: Princeton University Press.
- Rubin, B. R., A. Armstrong, and G. R. Ntegeye, eds. 2001. *Regional Conflict Formation in the Great Lakes Region of Africa: Structure, Dynamics and Challenges for Policy*. Center on International Cooperation. 11–13 November 2001, Safari Park Hotel, [http://www.cic.nyu.edu/peacebuilding/oldpdfs/RCF\\_NAIROBI.pdf](http://www.cic.nyu.edu/peacebuilding/oldpdfs/RCF_NAIROBI.pdf) (accessed August 25, 2010).
- Salehyan, I. 2009. *Rebels without Borders: Transnational Insurgencies in World Politics*. Ithaca, NY: Cornell University Press.
- Sambanis, N. 2004. "What is Civil War?: Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48:814–58.
- Scott, A. 1964. "Internal Violence as an Instrument of Cold Warfare." In *International Aspects of Civil Strife*, ed. J. N. Rosenau, 154–69. Princeton, NJ: Princeton University.
- Solomon, H., and G. Swart 2005. "Libya' Foreign Policy in Flux." *African Affairs* 104:469–92.
- Strand, H., J. Carlsen, N. P. Gleditsch, H. Hegre, C. Ormhaug, and L. Wilhelmsen. 2005. *Armed Conflict Dataset Codebook: Version 3-2005*. Oslo, Norway: International Peace Research Institute. ([www.prio.no/cscw/armedconflict](http://www.prio.no/cscw/armedconflict))
- Stremblau, J. J. 1977. *The International Politics of the Nigerian Civil War, 1967–1970*. Princeton, NJ: Princeton University Press.
- Tull, D. M. 2003. "A Reconfiguration of Political Order? The State of the State in North Kivu (DRC)." *African Affairs* 102: 429–46.
- . 2007. "The Democratic Republic of Congo: Militarized Politics in a 'Failed State.'" In *African Guerrillas: Raging against the Machine*, eds. M. Boås, and K.C. Dunn, 113–30. Boulder, CO, and London: Lynne Rienner Publishers.
- Turner, T. 2007. *The Congo Wars: Conflict, Myth & Reality*. New York and London: Zed Books.
- Vines, Alex. 1991. *Renamo: Terrorism in Mozambique*. Bloomington, IN: Indiana University Press.
- Wallensteen, P., and M. Sollenberg. 1998. "Armed Conflict and Regional Conflict Complexes, 1989–97." *Journal of Peace Research* 35:621–34.
- . 1999. "Armed Conflict, 1989–98." *Journal of Peace Research* 36:593–606.
- Weber, M. 2004. "Politics and the State." In *The Essential Weber: A Reader*, ed. S. Whimster, 131–32. New York and London: Routledge.

- Williams, Paul D. 2011. *War & Conflict in Africa*. Cambridge, UK: Polity.
- Wittgenstein, L. 1958. *Philosophical Investigations*. Upper Saddle River, NJ: Prentice-Hall.
- Woodwell, D. 2004. "Unwelcome Neighbors: Shared Ethnicity and International Conflict during the Cold War." *International Studies Quarterly* 48:197–223.
- Zartman, I.W. 1987. *International Relations in the New Africa*. Millburn, NJ: University Press of America.
- . 1989. *Ripe for Resolution*. Oxford: Oxford University Press.
- . 1992. "Internationalization of Communal Strife: Temptations and Opportunities of Triangulation." In *The Internationalization of Communal Strife*, ed. M. I. Midlarsky, 27–42. London: Routledge.
- . 2005. *Cowardly Lions: Missed Opportunities to Prevent Deadly Conflict and State Collapse*. Boulder, CO, and London: Lynne Rienner Publishers.

---

# Global Imbalances: Origins and Prospects

---

Luis Servén • Ha Nguyen

---

*This paper surveys the academic and policy debate on the origins of global imbalances, their prospects after the global crisis, and their policy implications. A conventional view of global imbalances considers them to primarily result from macroeconomic policies and cyclical forces that cause demand for goods to outstrip supply in the United States and other rich countries and that have the opposite effect in major emerging markets. An alternative view holds that global imbalances are the result of structural distortions and slow-changing factors that primarily affect assets markets. This paper reviews the analytical underpinnings of these two perspectives and the empirical evidence of their respective merits. The paper then assesses the outlook for global imbalances after the crisis, particularly in terms of policy action to reduce their magnitude. Policy intervention is warranted to the extent that the imbalances are driven by welfare-reducing distortions, but in this case, the primary target of policy intervention should be the distortions rather than the imbalances. Finally, the paper examines various forms of international spillovers that may call for multilateral action to limit global imbalances.*

JEL codes: F32, F36, F42

## Introduction

Since the outbreak of the world financial crisis in 2007, global imbalances—understood as large current account deficits and surpluses of major countries—have taken center stage in the debate over the international economic outlook. Academics have pondered the nature of the imbalances, offering contrasting views about the role of such imbalances at the beginning of the crisis, their potential threat to future global economic stability, and the need for policy measures to “rebalance” the global economy.

Global imbalances are not a new feature of the world economy. Episodes of large and persistent external imbalances in major economies were relatively frequent

in the 19th and 20th centuries.<sup>1</sup> The large U.S. current account deficits of the 1980s represent the most recent precedent of the situation observed in the 2000s. The main distinguishing feature of the current episode concerns the actors and their respective roles. In recent years, emerging economies have exhibited some of the largest surpluses, and advanced economies (notably, the United States) have had the largest deficits. Recent global imbalances involve what has been called an “uphill” flow of capital from poor countries to rich countries. This situation represents a departure from previous episodes and conflicts with the conventional economic theory prediction that developing countries should be net capital importers.

This paper offers a selective survey of the recent literature on global imbalances to shed light on their causes, their likely prospects, and the implications for domestic and multilateral policy action. From an analytical perspective, the paper argues that the various views advanced in the literature regarding the roots of global imbalances encompass two basic perspectives: one that emphasizes cyclical forces and macroeconomic policies affecting global goods markets and another that highlights structural distortions and slow-changing factors affecting global asset markets. The two perspectives are complementary, and both find support in the empirical evidence. However, the recent patterns of global capital flows seem to stress the relevance and the persistence of asset-market forces behind global imbalances.

This paper also reviews the outlook for global imbalances after the crisis, particularly the case for policy action to reduce the magnitude of these imbalances. Theory does not offer clear guidance on what constitutes an “excessive” external imbalance from the single-country perspective, but corrective action may be justified on welfare grounds to the extent that such an action seeks to address the structural and policy-induced distortions reflected in current account gaps.

Proposals for coordinated policy action to contain global imbalances have been advanced by recent academics and multilateral bodies with the rationale that the large current account gaps of individual countries impose costs on the global economy. The paper examines various mechanisms through which such spillovers may accrue, thereby justifying multilateral intervention.

The paper is organized as follows. Section 2 provides a brief retrospective on global imbalances. Section 3 summarizes various views of their causes. Section 4 discusses the outlook for global imbalances after the crisis and the case for national and international policy action to achieve rebalancing. Section 5 concludes the paper.

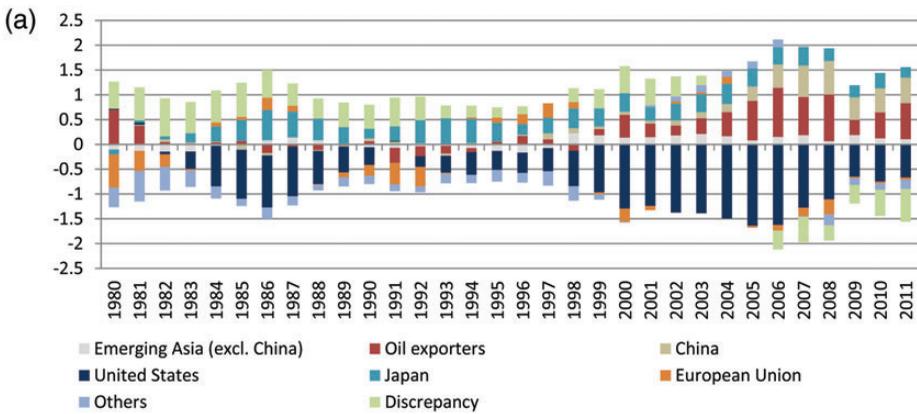
## Global Imbalances in Retrospect

Despite their recent rise to prominence in the policy debate, global imbalances have been a recurrent phenomenon in the world economy. Under the Gold Standard, for

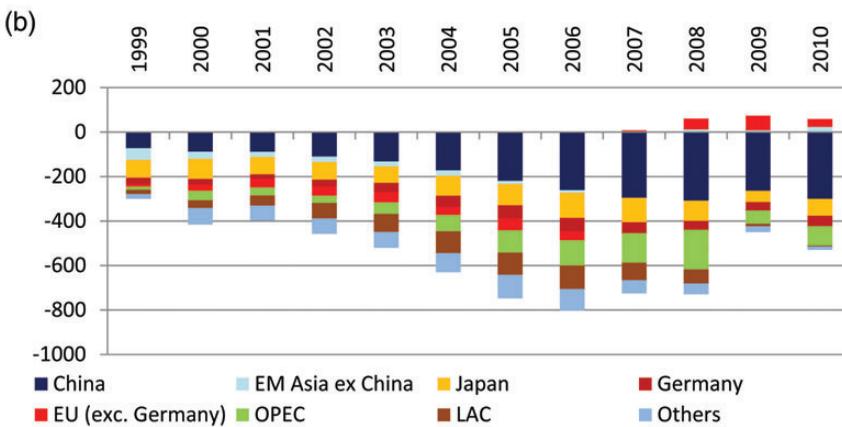
example, the United Kingdom consistently funded large deficits of the “emerging markets” of the time—the United States, Canada, Australia, and India. During the 1970s, oil-exporting economies had large surpluses that financed large deficits in developing countries, notably Latin America. In the 1980s, the United States incurred persistent deficits funded by major advanced economies.

The top panel of figure 1 provides a summary of current global account imbalances over the last three decades. It depicts the current account balances of the United States, the European Union (EU), Japan, China, and emerging Asia as well as oil-exporting countries, all measured as percentages of the world gross domestic product (GDP). Two episodes of large imbalances are clearly visible: one in the mid-1980s (already mentioned) and the latest one, of longer duration and larger

**Figure 1.** Current Account Balances (a) and U.S. Bilateral Current Account Balances (b)



Source: World Economic Outlook  
Data for 2010 are preliminary, data for 2011 are projected



Source: Bureau of Economic Analysis

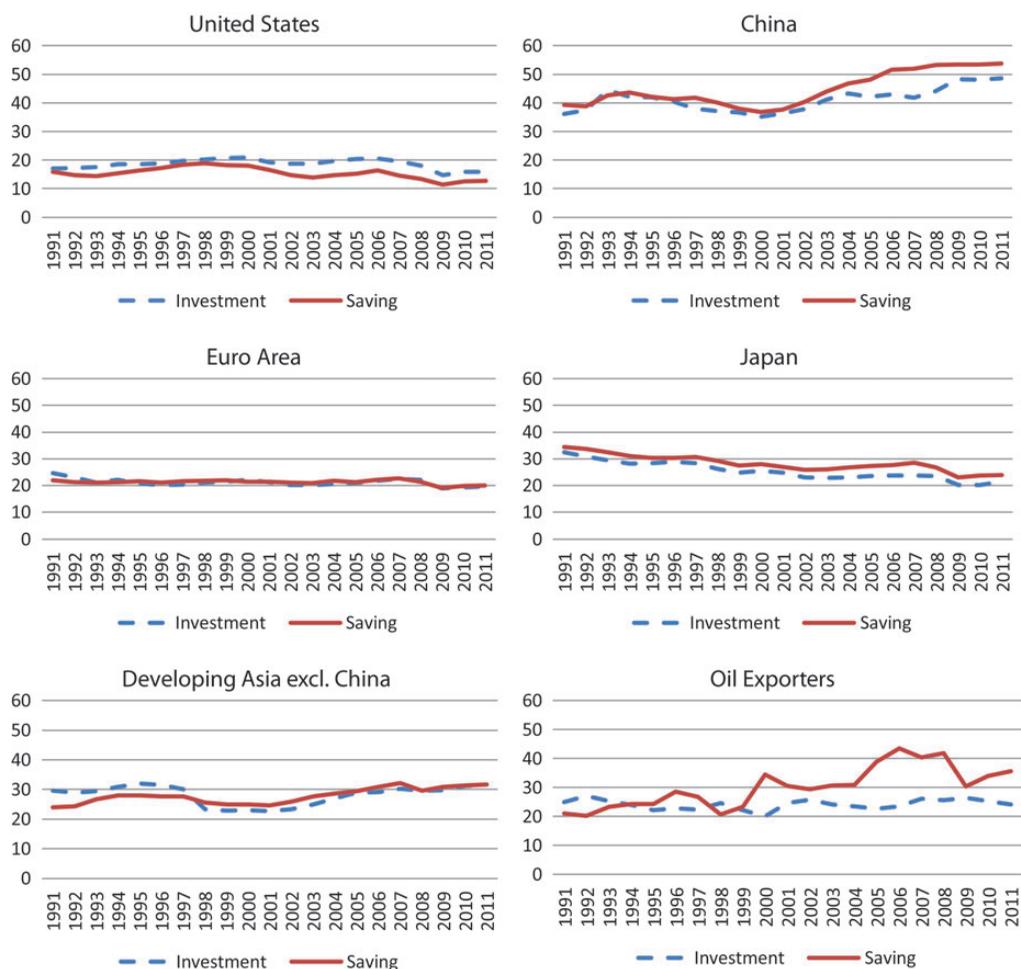
magnitude, beginning in the late 1990s. The latter episode peaked in 2006. By the measure shown, however, overall imbalances remained well above historical standards in 2010.

The figure shows that the U.S. current account deficit increased virtually without interruption since the mid-1990s and exceeded 1 percent of the world GDP after 1999. It peaked in 2005 and 2006 at over 1.5 percent of the world GDP. Thereafter, the U.S. external deficit declined below 1 percent of the world GDP in 2008 to 2009, but preliminary estimates suggest that it increased again in 2010. The figure also highlights the key difference between the episodes of the 1980s and 2000s. In the former episode, the main counterparts to the U.S. deficits were provided by the surpluses of Japan and, to a more modest extent, the EU. In contrast, in the 2000s, the surpluses of China and oil-exporting countries gradually became the largest. Since 2005, combined with the persistent surplus of the rest of emerging Asia, global imbalances have largely taken the form of massive capital flows from developing countries to rich countries, particularly the United States.<sup>2</sup> The bottom panel of figure 1 shows that after 2006, the bilateral deficit with China accounted for the largest portion of the overall current account deficit of the United States. Hence, the popular view of global imbalances as a problem of the United States as compared with China has begun to make sense only in the last few years.

Because the current account surplus is identical to the difference between savings and investment, the trends in these two variables across countries and regions provide information about the sources of changes in their external imbalances. Figure 2 presents the trends in saving and investment rates (as a percentage of the GDP of the country or group of countries in question) from 1991 to the present. As the figure illustrates, in most cases, the largest changes have come from the saving side. In the case of the United States, saving rates fell steadily after 1997 and declined at an accelerating pace in the years preceding the global crisis. As a result, in 2009, the saving rate reached its lowest level in two decades, at approximately 11 percent of the GDP. An examination of the trends in saving and investment reveals that from the end of the 1990s to 2003, a falling saving rate was the principal cause of the increasing external deficits of the United States. After 2003, the swings in the investment rate, which rose steadily prior to the crisis and subsequently collapsed, also contributed significantly to the trends in the current account.

The low saving rate of the United States contrasts with the extremely high levels observed in China, where saving accounts for over 50 percent of the GDP. Since 2000, both saving and investment rates have risen sharply in China, but the former has outpaced the latter. The result has been a major increase in China's current account surplus, which peaked at 10 percent of the GDP in 2007. Closer analysis reveals an increase in China's total saving over the 2000s

**Figure 2.** Investment and Saving Rates (% of GDP)



Source: World Economic Outlook – 2011 data are projected

of approximately 17 percent of the GDP. This rise reflects two forces: large increases in household and corporate savings (both near 20 percent of the GDP in recent years) and, to a lesser extent, increases in government saving (Prasad 2011). The contrast between China and the United States is also apparent from the perspective of consumption. U.S. aggregate consumption hovered at about 85 percent of the GDP in the 1990s and 2000s, with a slightly increasing trend. China's consumption ratio was much lower in the 1990s, under 65 percent, but it declined over the 2000s to a low of just 47 percent in 2008.

In industrial countries other than the United States, saving and investment rates have undergone changes that are more modest. In the EU, these rates have remained

roughly stable, at slightly above 20 percent of the GDP. In Japan, both saving and investment rates followed a downward trend over the 1990s but stabilized in the last decade, when the current account continued to show a steady surplus.

Investment rates fell sharply in emerging Asia following the 1997 crisis, leading to large current account surpluses in the early 2000s that subsequently declined as investment recovered.<sup>3</sup>

Finally, saving rates showed a steep increase in commodity-exporting countries because of the persistent rise in world commodity prices. Although this trend was interrupted in 2009 by the sharp decrease in commodity prices induced by the global crisis, rising saving rates led to widening current account surpluses, especially among oil exporters. However, high saving rates resumed as commodity prices recovered, and the combined current account surplus of oil exporters was projected to exceed 10 percent of their GDP in 2011.

## The Origins of Global Imbalances

The trends in saving and investment provide an accounting decomposition of the observed changes in current account balances across the world, but they do not identify the underlying causes. The sources of global imbalances have attracted considerable interest from academics and policy analysts, who have offered different views stressing a variety of causal factors. We will organize these views into two polar extremes, defined by their emphasis on conventional, cyclical macroeconomic forces (e.g., technological shocks, fiscal and monetary policies) relative to the emphasis on slow-changing forces, such as institutional development and deep-seated market or policy-induced distortions.

One view considers imbalances a reflection of asynchronous fluctuations in spending across the world. For want of a better term, we label this the “disequilibrium view” because it considers imbalances to be a temporary (and hence unsustainable) phenomenon pending correction. However, this designation does not imply that imbalances reflect a failure of goods or asset markets to clear or a failure of consumers and investors to act in their own best interests. The second view takes the opposite perspective: global imbalances reflect deep-seated structural forces and may be self-sustaining for an extended period. We label this perspective the “equilibrium view.”

These contrasting perspectives also differ in their relative emphasis on goods and assets markets. The former view focuses primarily on goods markets and takes a conventional current-account perspective, in which global imbalances results primarily from international asymmetries in the balance between demand and supply for (tradable) goods. The latter view, in contrast, centers on the capital account and stresses international asymmetries in the supply and demand for international assets.

In reality, the distinction between cyclical and structural factors is less than clear. Both sets of factors have likely influenced the recent episode of global imbalances. Thus, to a large extent, the two views should be regarded as mutually complementary rather than competing. However, their different emphases lead to somewhat different policy implications. The first view suggests the need for standard macroeconomic adjustment of aggregate expenditure and exchange rates. The second view calls for medium-term correction of underlying structural distortions. Next, we review the major aspects of both perspectives.<sup>4</sup>

### *Standard Macroeconomic Forces: The “Disequilibrium” View*

Conventional demand/ supply-based explanations of global imbalances have focused primarily on the macroeconomic factors responsible for the decline in the U.S. aggregate saving rates described above. This phenomenon reflected both declining public saving and, especially, declining private saving. The roles of real and financial shocks as well as macroeconomic policies have received considerable attention.

Some observers have argued that the declining savings and rising deficits of the United States in the 1990s were largely driven by acceleration in productivity growth, attributable to the information and communication technology revolution (e.g., [Cova et al. 2008](#)). In this conventional view, the anticipation of higher future growth would have prompted increased spending and deficits in the United States in a manner consistent with optimal forward-looking behavior ([Engel and Rogers 2006](#)). However, the persistent decline in private saving, particularly in the 2000s, was related to a financial cycle of unusual size, characterized by rapid growth of credit to consumers and escalating asset prices, particularly in housing. The credit and asset boom was propelled, in part, by persistently low interest rates. Some scholars (e.g., [Taylor 2009](#)) blamed these rates on an excessively expansionary monetary policy that encouraged excessive risk taking by banks and other financial institutions. Households responded by increasing consumption spending from their newfound wealth. Indeed, the asset boom resulted in a significant increase in the wealth-income ratio of the United States despite the low saving rates ([Bosworth and Collins 2010](#)). In turn, expansionary U.S. fiscal policy led to reduced public saving, which contributed to the decline in aggregate saving and the current account deficit. However, the precise extent of this effect remains disputed. Available empirical estimates suggest that the impact of public deficits on external deficits is far less than one for one (e.g., [Chinn and Ito 2007](#)).

In the conventional view, these shifts in aggregate expenditure and the external imbalances to which they contributed cannot be sustained. This conclusion follows from the routine application of the intertemporal budget constraint. Excluding the possibility of default and abstracting from capital gains and losses

on external assets and liabilities, the intertemporal budget constraint of a country dictates that its net liability position against the rest of the world at any given time cannot exceed the present value of its future trade surpluses.

In principle, however, a country can continue to run current account deficits for an extended period of time as long as it is capable of running sufficiently large trade surpluses in the future. This may be the case for developing countries that borrow to invest and accumulate capital and repay their debts once they reach a higher stage of development (e.g., Kraay et al. 2005). Likewise, an advanced country could run current account deficits if, in the future, it is expected to grow faster than the rest of the world. In effect, the advanced country finances its consumption by borrowing against its future income (as argued by Engel and Rogers 2006 for the United States).

In any case, the U.S. net foreign asset position has undergone a steep decline. The United States moved from a creditor position amounting to 10 percent of its GDP at the beginning of the 1980s to a debtor position close to 25 percent of its GDP in 2009. In absolute terms, this debtor position is the largest in the world. Reversal of this trend demands a change in the sign of the U.S. trade balance, which, in turn, requires a depreciation of the dollar to increase net exports. In an extreme but theoretically possible (if somewhat alarming) scenario, these adjustments may occur in the form of a sudden stop of capital flows into the United States and a collapse of the dollar (e.g., Mussa 2004; Roubini 2008).

In fact, the dollar has depreciated substantially in real (i.e., purchasing power) terms since 2002, but subsequent U.S. external deficits have remained quite large. This phenomenon suggests that further real-dollar depreciation may be forthcoming. The required magnitude of the trade-balance correction and the depreciation necessary to achieve it have been the object of detailed analysis in numerous studies (see, e.g., Obstfeld and Rogoff 2007, 2009 and Méjean, Rabanal and Sandri 2011).

Most of these studies follow convention in assuming that the U.S. external imbalance demands a *real* adjustment—a reversal of the trade deficit. However, recent literature has argued that *financial* adjustment can also play an important role. The change in the net foreign asset position of a country consists of (1) the trade balance and (2) the total return on its net foreign assets. The latter component includes not only the yields on foreign assets and liabilities, which are included in the current account, but also the capital gains and losses arising from changes in their prices, which are omitted from the current account. Because of this important omission, the time path of the net foreign asset position is no longer determined by the current account balance alone.

Financial adjustment had received little attention until a few years ago, but new research has stressed its quantitative importance in the case of the United States for two reasons. The first reason is the difference between the yields on U.S. assets and liabilities. A number of observers find that foreign assets held by

U.S. investors yield higher returns than U.S. assets held by foreign investors (Hausmann and Sturzenegger 2006; Gourinchas and Rey 2007a, b; Forbes 2010). Although seemingly counterintuitive, this situation makes it possible for the United States to earn positive returns on its foreign asset portfolio even if the country's net asset position is negative—a phenomenon that appears to be unique to the United States (Habib 2010).

The second reason is related to the effect of changes in the dollar exchange rate on the U.S. asset position. The key here is that a real depreciation of the dollar generates a wealth transfer in favor of the United States because its external liabilities are denominated in dollars whereas its assets are denominated in other currencies (typically those of the issuing countries). This is exactly the reverse of the usual situation in emerging markets indebted in foreign currency when their exchange rate undergoes a real depreciation.

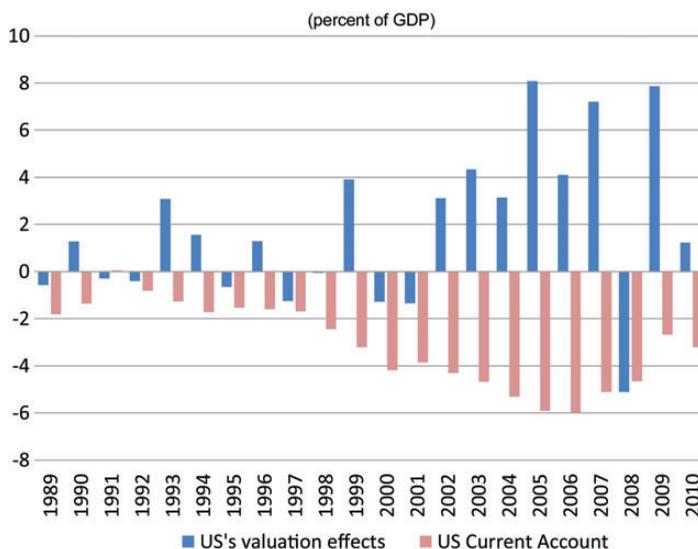
Thus, dollar depreciation has a double effect on the U.S. external asset position. On the one hand, it generates real adjustment through an improving trade account balance. On the other hand, it generates financial adjustment through capital gains (i.e., losses for the rest of the world). The important insight here is that assessments of the trade surplus required to put the external position of the United States on a sustainable trajectory can easily overstate the extent of the necessary policy correction if they do not take this second effect into consideration.

The empirical magnitude of financial adjustment remains controversial, particularly because of the lack of detailed data on the yields and prices of international assets and liabilities (see, e.g., Gourinchas and Rey 2007a; Curcuro, Thomas and Warnock 2008; and Gourinchas, Rey and Govillot 2010). However, the massive increase in cross-border asset holdings over the last two decades logically implies that valuation effects must have become more important. Indeed, there is evidence that, in recent years, valuation changes have dominated the observed changes in countries' net foreign asset positions, dwarfing the effects of current account imbalances (Devereux and Sutherland 2010).

Figure 3 provides information on the growing importance of such valuation effects in the case of the United States. The graph shows the trends in the U.S. current account balance and the valuation effects due to asset price movements and exchange rate changes over the last two decades, with both expressed as percentages of GDP. The change in the net foreign-asset position is given by the sum of the two series.

The figure shows that the United States has enjoyed large valuation gains since 2002, which peaked at 8 percent of its GDP in 2005—an order of magnitude similar to that of the current account deficits of those years. Two main factors were at play. The first factor was the relative decline in U.S. stock market prices since 2000, which generated capital losses for foreign investors (Kraay and Ventura 2005). The second factor was the depreciation of the dollar after 2002, which

**Figure 3.** U.S. Current Account Balance and Valuation Effects (Percent of GDP)



Source: Bureau of Economic Analysis

allowed the U.S. net foreign asset position to increase from 2003 to 2007 despite the record-high current account deficits of those years. In 2008, however, the United States suffered a large negative valuation effect because of international investors' "flight to quality," which caused an appreciation of U.S. assets and the dollar.<sup>5</sup> However, valuation effects turned in favor of the United States again in 2009. Looking forward, the implication is that the depreciation of the dollar required to ensure the sustainability of the external position of the United States is likely to be much more modest than what would be necessary if external adjustment had to occur only from the trade balance. Indeed, over the last twenty years, the correlation between the current account balance and the change in net foreign assets in the United States has been only 0.15, which implies that the current account plays a relatively minor role in shaping the country's net foreign asset position.

### *Structural Factors: The "Equilibrium" View*

In contrast to the view of global imbalances as driven by international shifts in aggregate demand for goods, the "equilibrium view" traces them to structural factors that have caused a shift in the international pattern of excess demand for assets. These structural factors change only gradually, implying that global imbalances may be sustained for a long time.

Admittedly, this is a very broad generalization, in part because it encompasses a variety of views on the ultimate determinants of global imbalances. However, one feature common to all of these views is the emphasis on assets markets and the capital account, in contrast to the analytical emphasis placed on the supply and demand for goods and the current account in the preceding section. Consider first the supply of international assets. An influential view holds that global imbalances largely reflect the differential ability of advanced and emerging countries to supply high-quality assets for international savers (e.g., [Caballero, Fahri and Gourinchas 2008a, b](#)). In particular, the financial underdevelopment of emerging countries prevents them from generating financial instruments that are attractive to their savers for two simple reasons: the yields on local assets are too volatile, and they bear expropriation risks, as exemplified by the financial crises of the 1990s. The result is that international savers tilt their portfolios toward the assets of countries with more advanced financial markets, particularly the United States. A growth acceleration in emerging countries (or an oil price boom) that increases their wealth and saving—the forces behind the so-called “global saving glut” ([Bernanke 2005](#))—leads them to expand their holdings of U.S. assets.<sup>6</sup> This expansion is achieved through U.S. current account deficits that raise the volume of U.S. assets available to international investors. This process can persist as long as its driving force—the lagging development of financial markets in emerging countries—remains unchanged. As a result, capital flows “uphill,” from poor countries to rich countries.

An analogous line of reasoning stresses international asymmetries in the demand for assets, rather than their supply. These asymmetries may be driven by the response of individual savers to idiosyncratic risk. Underdeveloped financial markets in emerging countries offer individuals few options to hedge income shocks or borrow against future income, and these few options force savers to raise their precautionary wealth accumulation to ward off idiosyncratic income risk ([Mendoza, Quadrini and Ríos-Rull 2009](#)). This mechanism affects not only households but also entrepreneurs facing risky investment projects ([Angeletos 2007](#); [Sandri 2010](#)). Similarly, the underdevelopment of the social-protection system in emerging countries encourages precautionary wealth accumulation against unemployment or retirement risk ([Carroll and Jeanne 2009](#)).

Other things equal, the result is that savers in emerging countries tend to save more than those in industrialized countries. Financial integration between emerging economies with underdeveloped financial markets and weak social protection systems and rich countries with more advanced financial markets and social-protection systems leads to a global equilibrium in which the former countries acquire a creditor position whereas the latter countries are net debtors. As long as the gap between both sets of countries, in terms of the degree of development

of financial markets or social-protection systems, remains unaltered, global imbalances and uphill capital flows can persist.

These theoretical arguments seem consistent with the empirical observation that, other things equal, during the 2000s, private investors from emerging countries with less-developed financial markets typically allocated larger shares of their portfolios to U.S. assets (Forbes 2010). Related to this observation, many analysts attribute China's high levels of corporate saving to the limited development of domestic financial markets, which leads firms to fund the bulk of their investment through retained earnings. However, the weak governance framework of large state-owned enterprises, which faced no obligation to distribute dividends until recently, has undoubtedly played a role as well. The empirical evidence also suggests that Chinese households' limited access to borrowing in a context of rising income risk (following the pro-market reforms of the 1990s) and the weakening of the social protection system over the same period are key factors in the rising levels of household saving in China (Chamon, Liu and Prasad 2010; Chamon and Prasad 2010). Likewise, rising public saving in the 2000s was largely driven by the decline, relative to GDP, of net government transfers, notably pension and social insurance payments.

Deliberate policy choices by national economic authorities may be another factor in international asymmetries in the demand for assets. Policy makers may decide to engage in foreign asset hoarding to support a "new mercantilist" development strategy based on export-led growth. This objective is facilitated by an undervalued real exchange rate and foreign asset accumulation to preserve export competitiveness. Historically, Japan and Germany have followed the export-led growth model. This model has been adopted more recently by a number of emerging markets, particularly in East Asia, with China prominently among them.<sup>7</sup> Indeed, some observers have argued that global imbalances arise primarily from the export-led growth strategy of Asian economies (e.g., Adams and Park 2009). This development strategy defines the so-called "Bretton Woods II" system (Dooley, Folkerts-Landau and Garber 2004), in which emerging Asian countries play the role of producers of last resort, and advanced countries—led by the United States—are the consumers of last resort whose deficits are financed by capital inflows from Asia.<sup>8</sup> Proponents of this view claim that Bretton Woods II can remain in operation indefinitely as long as emerging Asia maintains the export-led strategy and is willing to accumulate claims on the United States and other advanced countries.

The logic of this strategy is that, under appropriate circumstances (related mainly to the presence of strong positive externalities in the production of tradable goods), the mercantilist development model may succeed in accelerating structural transformation and economic growth. In this vein, Rodrik (2008) offers suggestive evidence of a positive correlation between exchange rate

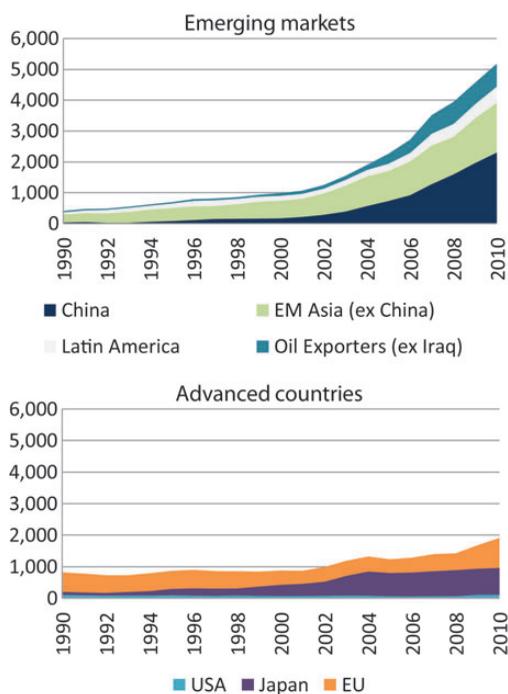
undervaluation and economic growth. Nevertheless, the merits of the export-led growth strategy from the welfare standpoint are not obvious because the large-scale accumulation of external (and, typically, low-yield) assets involves major opportunity costs in terms of foregone consumption (see [Korinek and Servén 2010](#) for details). Policy makers' ability to sustain the export-led strategy may be threatened by its diminishing growth returns as positive externalities become exhausted, by society's rising demands for higher consumption, and by the increased ability of private investors to offset the authorities' attempts at continued foreign asset hoarding, brought about by deepening financial integration ([Eichengreen 2004](#)).

Emerging-market policy makers may also engage in foreign asset hoarding as a precautionary policy. In the absence of mechanisms for international diversification of aggregate risk, emerging countries integrated into the global financial system must resort to self-insurance against external shocks, such as disruptions of international capital flows of the kind observed in the crises of Asia and Russia in the 1990s. Thus, emerging countries accumulate external assets—preferably short-term instruments—upon which they can draw in the event of a “sudden stop.” Unless the global financial system generates new mechanisms to hedge these risks, this precautionary foreign asset accumulation is unlikely to cease.<sup>9</sup>

Empirically, the massive accumulation of international reserves by emerging economies during the last decade suggests that the policy of foreign asset hoarding has played an important role. [Figure 4](#) depicts the foreign reserve holdings of industrial and emerging countries since the early 1990s. Between 1998 and 2010, the reserve holdings of the latter, measured in constant prices, increased six-fold, whereas those of industrial countries rose less than 100 percent. As a result, the volume of international reserves held by emerging markets at present greatly exceeds that of industrial countries. Indeed, at the end of 2010, China's foreign reserve stock exceeded that of all industrial countries combined. The rest of emerging Asia has also dramatically increased its reserve holdings. However, the phenomenon is not confined to Asia; Latin American economies and oil-exporting countries have also accumulated large volumes of international reserves over the last decade.

Nevertheless, the question remains whether reserve accumulation was primarily driven by caution against the volatility of international capital flows or by the pursuit of competitive exchange rates. Disentangling the relative importance of these two factors is not an easy task; their respective roles are likely to vary across countries and over time. [Aizenman and Lee \(2007\)](#) attempt to address this question in a regression framework using data from 49 countries for 1980 to 2000. They conclude that both motives were at work during that period but that the precautionary saving motive was more important. In line with this finding, [Durdu, Mendoza, and Terrones \(2007\)](#) show that foreign reserve holdings of the countries that suffered “sudden stops” in the 1990s and 2000s were substantially

**Figure 4.** Foreign Exchange Reserves of Industrial and Emerging Countries (Billions, constant 2000 US\$)

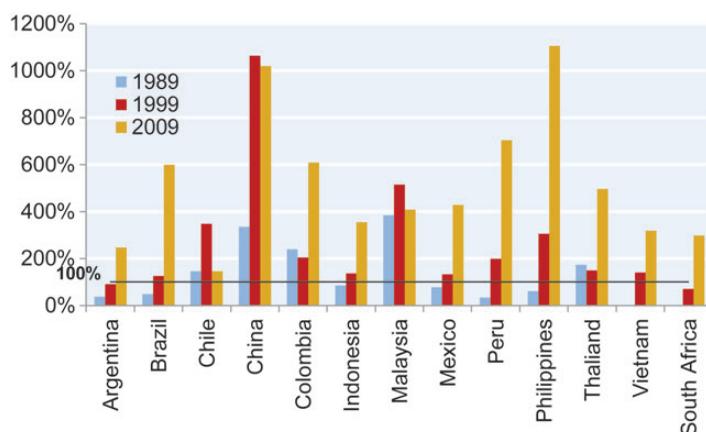


Source: IMF

larger after the sudden stop than before in all but one country. The change was quite significant—on average, they doubled. This evidence seems to lend support to the precautionary motive. However, there are clear indications that in a number of emerging countries, notably in Asia, the reserve buildup has been too large to be justified by the precautionary motive alone (see, e.g., [Jeanne and Ranciere 2011](#)). Figure 5 corroborates this view. In all emerging markets shown in the figure, foreign reserves amply exceeded short-term external debt in 2009, in most cases by a huge margin—over tenfold in China and fivefold or more in a number of other countries. This finding suggests that mercantilist reasons for foreign asset hoarding likely played a prominent role in a number of countries, especially China.

Although this evidence strongly suggests that deliberate policy choices were a major factor in the accumulation of foreign assets by emerging markets, it does not necessarily imply that they were the only factor. If this had been the case, capital inflows from emerging markets to the United States would likely reflect a dominant role of official flows over private flows. The available information,

**Figure 5.** Emerging Market Foreign Reserves (As % of Short-term Debt)



Source: World Development Indicators

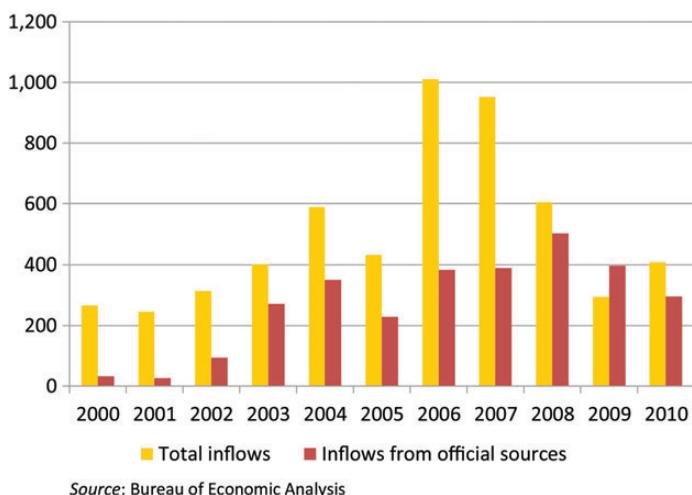
shown in figure 6, should be taken with some caution, but overall, it suggests a mixed picture. Purchases of U.S. assets by official entities (central banks and government bodies) from emerging markets in Latin America, Asia, and the Middle East became increasingly large in the 2000s. Indeed, after the onset of the crisis in 2007, such purchases became the sole source of inflows from these countries. However, over the decade as a whole, the total volume of emerging-market official inflows to the United States was roughly on par with that of private inflows, and the latter actually exceeded the former in the precrisis years from 2000 to 2007.<sup>10</sup>

Finally, the above discussion highlights the role of a variety of forces that contributed to increasing emerging countries' excess demand for quality assets. However, it is important to keep in mind the effect of financial innovation in advanced countries within a context of lagging, and even weakening, regulation of the financial system, particularly the United States. These factors contributed to the credit boom and the housing bubble, which, as many observers have noted, were key aspects of the collapse of private saving and the widening of the U.S. current account deficit. Indeed, some empirical exercises suggest that the rise in aggregate expenditure derived from the private sector's bubble-driven wealth increase and enhanced access to credit may quantitatively explain the bulk of the U.S. external imbalances of the 2000s (Laibson and Mollerstrom 2010). From a political-economy perspective, Rajan (2010) has argued that the hands-off regulatory stance at the root of the credit boom can be partly viewed as a response to the longstanding increase in U.S. income inequality. According to this view, given the limited political support for redistributive policies, enhanced access to cheap

---

**Figure 6.** Gross Capital Inflows into the U.S. from Emerging Markets (US\$ Billions)

---



credit would have offered an alternative way to sustain consumption growth by large segments of the population despite their stagnant or diminishing real incomes.

## Future of Global Imbalances

Evaluating competing views about the causes of global imbalances enables us to better assess the outlook for global imbalances after the crisis, particularly in terms of the need for policy action to reduce their magnitude. We next provide an overview of the imbalances after the global crisis, and we then examine various forms of international spillovers that may call for multilateral action to limit global imbalances. The key principle is that policy intervention is warranted to the extent that the imbalances are driven by welfare-reducing distortions, but in such case, the primary target of the intervention should be the distortions rather than the imbalances themselves.

### *Global Imbalances after the Global Crisis*

It is unquestionable that global imbalances have narrowed since the onset of the global crisis. The U.S. current account deficit, measured as a proportion of the GDP, has declined sharply, from 6.1 percent of the GDP in the second quarter of 2006 to 3.5 percent at the end of 2010. On a global scale, the crisis initially led

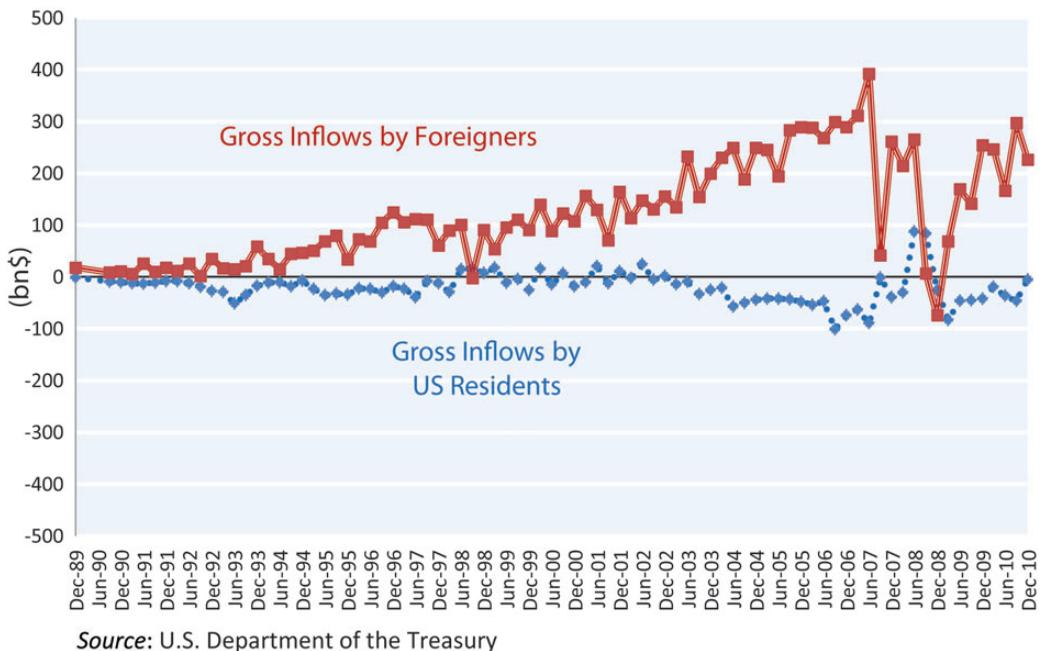
to an abrupt fall of international capital flows and to the collapse of world trade and oil and commodity prices. The latter, in turn, caused a sharp reduction in the surplus of oil-exporting countries. China's surplus also declined under the additional pressure of a major fiscal stimulus package deployed by the authorities in response to the crisis.

Is this the beginning of the end for global imbalances? It is difficult to provide a conclusive answer. As the preceding section suggests, the future of global imbalances depends on a constellation of real and financial forces whose evolution is difficult to predict. On the one hand, to the extent that their key determinants remain largely unchanged, the imbalances may resume after the crisis. Indeed, figure 1a shows that the U.S. current account deficit widened slightly in 2010. On the other hand, prospective and ongoing policy changes and structural reforms, summarized below, could lead to a sustained narrowing of imbalances.

In this regard, it is important to emphasize that the crisis did not involve a "sudden stop" of capital flows to the United States, which could have given rise to a disorderly unraveling of global imbalances and an abrupt depreciation of the dollar. Instead of the depreciation that many had predicted, the dollar experienced an initial appreciation as a result of international investors' flight to safety, which led them to shelter in low-risk U.S. Treasury debt at the expense of risky assets, from corporate debt to emerging-market assets. In effect, the dollar became the reserve currency of last resort, and the U.S. government became the borrower of last resort (Gorinachas, Rey and Govillot 2010). This "safe haven" effect of U.S. assets has recently resurfaced in the context of the Eurozone's sovereign debt turmoil.

In the aftermath of the global crash, there are some indications that the precrisis pattern of capital flows to the United States may be resuming. This pattern is illustrated in figure 7, which updates the information in *De la Torre, Schmukler, and Servén (2009)*. The flows shown in the figure comprise long-term financial instruments only, but their time profile is revealing. The top line captures the inflow of capital from nonresident investors. Until 2008, it is always positive, reflecting an upward trend interrupted in late 2006. The bottom line captures the inflow of capital from resident investors. For the most part, these flows take negative values, indicating capital outflows by U.S. investors. However, at the onset of the subprime turmoil in mid-2007, these patterns changed abruptly: capital inflows from nonresidents collapsed, and outflows from residents changed sign, reflecting capital repatriation by residents to stem losses in domestic markets or to seek safe haven from the global turbulence. Since 2009, however, the data reveal an incipient or emerging return to the precrisis pattern: both capital inflows of foreign investors and outflows of resident investors appear to return to their earlier trends. This finding indicates the persistence of the deep determinants of global imbalances and suggests that they may not disappear anytime soon.

**Figure 7.** Gross Capital Inflows to the U.S. (in long-term securities) (US\$ Billions)



Looking ahead, the global liquidity crisis is likely to further encourage precautionary accumulation of foreign assets by emerging markets. In fact, even at the height of the turmoil, some economies appeared reluctant to deploy their vast reserves for fear of weakening the confidence of international investors (Aizenman and Sun 2009; Aizenman and Hutchison 2010). The experience of these countries in the crisis could prompt them to hold even larger stocks of liquid foreign assets in the future, contributing to the “uphill” pattern of global capital flows.

### *What Role for Policies?*

Should policies be deployed to eliminate, or at least reduce, global imbalances? In theory, the size or persistence of external imbalances does not provide firm ground for the adoption of policies to reduce them. Theory does not offer clear guidance as to whether a country’s current account gap (or global imbalances, more broadly) is sustainable. Theoretical concepts of sustainability typically involve an assessment of whether intertemporal budget constraints are met. Therefore, these concepts raise a timely and important question: “Are deficits today commensurate with eventual future surpluses and capital gains to prevent the economy’s net foreign asset position from becoming increasingly negative?”

In effect, any given current account position can be sustainable, in principle, given sufficient future adjustment. However, the medium- or longer-run sustainability of a country's current account depends on its ability to repay, which, in turn, depends on a host of other factors in addition to current account balances or debt levels.

On conceptual grounds, there is nothing wrong with current account imbalances as long as they reflect socially optimal saving and investment decisions. As noted, in principle, it would be entirely desirable for a country to incur external deficits in the face of productivity improvements or rising future income. Likewise, a country would be well advised to run large external surpluses if it enjoys a large, but temporary, terms-of-trade windfall. However, if saving and investment decisions primarily reflect structural and policy-imposed distortions rather than socially optimal intertemporal choices, then policy action may be warranted (e.g., [Blanchard and Milesi-Ferretti 2009](#)).

Examples of such important distortions are the lack of insurance mechanisms at the national or global levels (highlighted in the previous section); inappropriate regulation of the financial system, which most observers place among the key causes of the U.S. bubble-driven overconsumption of the 2000s and of the ensuing financial crash; or its polar opposite, financial system underdevelopment that contributes to over-saving in emerging markets. In addition, externalities across national economies may make individual countries' imbalances a global (i.e., systemic) concern and justify policy intervention.

In practice, it is often difficult to disentangle the causes of external imbalances. Policy makers may need to take precautionary action even in the absence of a conclusive diagnosis, especially if an abrupt unraveling of the imbalances is likely to entail large costs and create major economic disruption on a national or global scale. Even then, however, intervention should be guided as much as possible by an understanding of the distortions responsible for the socially undesirable pattern of imbalances. The primary objective should be to counter the distortions rather than the imbalances.

The role of domestic distortions is most visible in the cases of China and the United States. For China, there is broad consensus that strengthening the weak social-safety net and developing domestic financial markets should decrease saving rates from their current record highs. Improved coverage and quality of health, pensions, and other social-insurance schemes will reduce households' need for self-insurance through precautionary saving. Indeed, some empirical estimates suggest that the effect could be quantitatively significant. For example, in the case of China, whose initial level of social expenditure is quite low, a 1 percent increase in public social expenditure could lead to a 2 percent decline in household saving, with both expressed as a ratio to GDP ([Baldacci et al. 2010](#)).

The development of domestic financial markets and the enhancement of households and firms' access to finance should likewise reduce private saving, not only in China, but also in other high-saving emerging markets. The relaxation of borrowing constraints lowers households' precautionary savings and the saving for housing purchases and education expenditures. For firms, financial market development reduces the need to retain earnings in anticipation of the potential arrival of good investment projects and might boost aggregate investment by relaxing binding credit constraints on capital expenditures. Overall, the international evidence suggests that the development of financial markets leads to falling saving rates (e.g., Loayza, Schmidt-Hebbel and Servén 2000) and possibly to a (modest) rise in investment rates (OECD 2011), implying an unambiguous decline in the current account surplus (or an increase in the deficit).<sup>11</sup>

In addition to their contribution to narrowing external imbalances, these reforms are likely to have positive effects on welfare and possibly on growth. For example, households' enhanced access to risk-management mechanisms through financial markets and social insurance will likely reduce the variability of their consumption and increase their welfare. In addition, long-run growth should be positively affected by the rise in human capital accumulation associated with improved health and education services and by the improved efficiency in the allocation of investment brought about by financial market development.

Domestic reforms of this type are already underway in China. Recent policy announcements in connection with the Twelfth Five-year Plan anticipate a gradual shift away from the producer-biased growth model of the 2000s toward a more consumption-friendly framework, with higher real wages, increased government social expenditures, and strengthened social-safety nets.<sup>12</sup> These factors are likely to halt the decline in disposable household income relative to GDP that characterized the last decade and may begin to reduce the incentives for household saving. The declining trend in the ratio of aggregate consumption to GDP that characterized the 2000s appears to have stopped in 2009. Finally, the rise in firm saving is likely to end under the pressure of newly introduced dividend-payment requirements for state-owned enterprises as well as the gradual improvement in access to external financing for investment by private firms.

In the United States, the distortions have been the opposite of those in China. As noted, the private-saving collapse was largely driven by the credit boom and the housing bubble in a context of weak financial regulation. The decrease in household wealth caused by the bursting of the asset bubble and the post-crisis credit contraction have led to an incipient reversal of the steep decline in personal saving, with household saving rates rising by some 2 to 3 percentage points post-crisis. The rising trend is likely to continue in coming years as households partially rebuild their wealth stocks. This process should be helped by reforms underway to strengthen the lax financial-system regulation that fueled the credit and

consumption boom. However, the scope and effectiveness of such reforms have yet to be established. For the U.S. public sector, the record-high deficits of recent years are set to decline as the crisis-driven fiscal stimulus is gradually phased out. However, the long-term viability of public finances, and thus of external accounts, will remain threatened by another key distortion, the mounting cost of the massively inefficient healthcare system, whose prospects for deep reform remain distant.<sup>13</sup> This discussion relates to domestic distortions, but global imbalances also reflect systemic distortions, most notably the lack of global insurance mechanisms, which prompts countries to engage in foreign-asset hoarding as self-insurance against external shocks. Self-insurance in the form of low-yield reserve accumulation is costly for the countries involved because it entails major opportunity costs. These costs become considerable when reserve stocks reach their current proportions in emerging markets (e.g., [Rodrik 2008](#)).

Precautionary reserve hoarding imposes negative externalities across countries. First, individual countries ignore the fact that their reserve accumulation contributes to depressing the rate of return on reserve assets, adding to the opportunity costs of all other countries. Second, the decline in rates of return on safe assets poses risks for global financial stability because it tends to encourage investor risk taking (the “search for yield”) and creates favorable conditions for the emergence of asset bubbles like those at the core of the global crisis ([Caballero, Fahri and Gourinchas 2008b](#)). The deterioration in the quality of lenders’ portfolios encouraged by cheap funding also makes the economy more vulnerable to adverse shocks ([Eden 2012](#)). Paradoxically, this situation amplifies external volatility and prompt additional precautionary hoarding. The implication of these spillover effects is that the self-insurance motive tends to result in the socially excessive accumulation of foreign assets.

The development of global insurance mechanisms offers a more efficient alternative to self-insurance, thus reducing the incentives for the precautionary buildup of foreign reserve stocks (and current account surpluses) in emerging economies. The global crisis has given new impetus to these mechanisms at the global and regional levels. These mechanisms include expanding the IMF’s Flexible Credit Line and High-Access Precautionary Arrangement as well as regional currency-swap schemes, such as the Chiang Mai Initiative in East Asia<sup>14</sup> and bilateral swap arrangements between national monetary authorities and major central banks, such as the Federal Reserve, the European Central Bank, and the Bank of Japan. However, the scale of these mechanisms remains modest, and strong multilateral action to expand them is necessary before they can play a credible risk-management role comparable to that of self-insurance through the hoarding of reserves.

There are other channels through which the external imbalances of some countries could impose negative externalities on others. One channel that is

frequently mentioned involves the international transmission of financial turbulence. The underlying idea is that global imbalances could trigger a financial crash in a “systemic” deficit country, understood as a large country with extensive financial links with the rest of the world, such as the United States. This situation could arise as the result of a “sudden stop” or as a consequence of the stress placed on the financial system by the intermediation of large capital inflows (e.g., [Portes 2009](#)). Either scenario might have severely disruptive effects on other countries through the ripple effects of a collapse in international asset prices in a world ridden with financial frictions. However, these spillover effects are likely to be ignored by domestic authorities when assessing the risks posed by their external gaps and the need for preventive policy action. Thus, individual countries’ external imbalances may become “excessive” from a global perspective.

This reasoning sounds persuasive, because the propagation of financial turmoil poses very real dangers to the global economy, as the crisis has clearly shown. However, it is not obvious why external imbalances should make those dangers worse. In particular, the international transmission of asset price shocks is driven by *gross* cross-border asset positions (e.g., [Devereux and Yetman 2010](#)) whereas external imbalances are related to the time path of *net* foreign asset positions, and the empirical evidence shows that there is no significant link between the trends in gross and net assets (Broner, Lorenzoni, and Schmukler Forthcoming). In addition, history shows a mixed record regarding the unwinding of large global imbalances ([Bracke et al. 2008](#)). For example, the Gold Standard episode had a mostly orderly unwinding. In contrast, the imbalances of the 1970s eventually triggered the debt crisis of the 1980s, and the imbalances of the 1980s led to a mixed outcome; they were followed by a prolonged crisis in the main surplus country (Japan) but had little effect on the main deficit country (the United States).

Negative externalities may also arise from the “export-led” growth strategy followed by some countries, which involves expenditure compression and foreign asset accumulation in support of undervalued real exchange rates—or, equivalently, controls on external borrowing imposed with the same objective ([Korinek 2012](#)). In effect, this strategy attempts to divert export demand from competitors and may prompt retaliatory action in a “currency war” of competitive real devaluations that may negatively affect all participants. In fact, a policy of undervaluation is equivalent to the adoption of import tariffs and export subsidies, but these measures represent unfair trade practices under WTO rules, whereas undervaluation does not ([Rodrik 2008](#); [Korinek and Servén 2010](#)).

Related to the issue of undervaluation is the adverse spillover from large external surpluses, which arises through their negative contribution to global demand. This adverse spillover is especially a concern in the post-crisis situation, in which global interest rates remain at record lows ([Blanchard and Milesi-Ferretti 2011](#)).

In normal circumstances, a reduction in the investment or an increase in the saving of deficit countries would lead to a reduction in global aggregate demand *ex ante*, which would be offset *ex post* by a decline in global interest rates. However, when the world economy is close to a liquidity trap, global interest rates cannot fall further, and the result is a decrease in global demand and output. This outcome could easily be avoided if surplus countries were to raise their aggregate demand. However, this need not be the case: although foreign investors can force deficit countries to adjust by refusing to continue lending to them, markets rarely put pressure on surplus countries (e.g., China) to correct their imbalances. With major industrial countries facing a protracted process of deleveraging and slow growth after the world crisis, demand expansion by surplus emerging markets is likely to be a critical factor in restoring global growth.

These various forms of international spillovers seem to call for multilateral action regarding global imbalances. In truth, however, it is very difficult to gauge their quantitative significance from the perspective of growth or welfare, and the practical challenges involved in multilateral action should not be underestimated. Notwithstanding the qualitative arguments above, it would be exceedingly difficult to establish whether a given deficit or surplus is “excessive” from the global perspective or if a country’s depreciated real exchange rate reflects an attempt at unfair competition rather than, for example, a genuine national preference for frugality.

In this context, some current multilateral initiatives (e.g., the G20’s “mutual assessment process” or the multilateral surveillance recently entrusted to the IMF) represent efforts to lay the groundwork for consensus guidelines on countries’ external balances, exchange rates, foreign assets, and so on. However, it is not clear how much progress they can achieve. In the end, it is unlikely that major countries would agree to set domestic policies (fiscal, monetary, or financial) to maintain their current account balances or real exchange rates to the levels deemed acceptable by other countries. Furthermore, no global governance mechanisms exist to enforce multilateral statements of intention to that effect. The ongoing European crisis shows that collective action can be extremely difficult and complex, even when these mechanisms are present.

In other words, perhaps not too much should be expected from attempts at coordinated rebalancing, aside from the domestic policy actions chosen by individual countries in their own self-interest. Of course, many of the country-level reforms to address domestic distortions, such as those outlined earlier to foster financial system development and social protection in emerging markets and to enhance financial regulation in advanced economies, are steps in the right direction to mitigate the external effects summarized above. However, the extent to which these reforms can achieve a significant, speedy, and sustained correction of global imbalances remains an open question.

## Conclusion

This paper has reviewed contrasting analytical perspectives in the academic and policy debate on the roots of global imbalances, their likely path in the post-crisis world, and the potential need for policy action to hasten their elimination in the aftermath of the crisis.

Although some observers have argued that the latest episode of global imbalances largely reflects macroeconomic policies and other cyclical factors affecting world goods markets, others hold that the imbalances are primarily driven by an excess global demand for safe assets. The excess demand reflects structural and policy-induced distortions in domestic and global financial markets as well as the action of neo-mercantilist growth strategies. In reality, the evidence suggests that the widening global imbalances of the last decade likely resulted from a combination of both sets of factors.

Global imbalances have narrowed after the crisis under the influence of expansionary demand-management policies in emerging markets as well as the incipient (if modest) rise in U.S. private saving rates. The unwinding of the U.S. fiscal stimulus will likely add to this trend. However, failure to address the domestic and global distortions that, according to the equilibrium view, lie at the root of global imbalances would likely lead to their return. There are reasons to think that excess global demand for quality assets has risen post-crisis as a result of the destruction of previously safe assets—first, with the collapse of mortgage-based securitization and, more recently, with the sovereign debt turmoil in Europe—and to countries' unrelenting demand for self-insurance after the global crash (Caballero 2010). Other things equal, this new surge of global demand for quality assets will likely contribute to the return of global imbalances. Indeed, the incipient return of U.S. capital inflows to their precrisis patterns suggests that this return is already underway.

This paper has reviewed the case for policy action toward the elimination of global imbalances. In principle, current theory offers little guidance to determine what constitutes an excessive current account deficit or surplus. Nevertheless, to the extent that the imbalances reflect structural or policy-induced distortions, actions to remove them are justified from the welfare perspective. However, their primary target would most appropriately be the distortions at work rather than the imbalances themselves, although the magnitude of the latter should decline as a result. A number of domestic policy initiatives recently undertaken in advanced and emerging markets address at major distortions that have direct effects on external imbalances.

Finally, this paper has examined the case for international policy coordination regarding global imbalances. Conceptually, the case for coordination hinges on the existence of negative externalities through which a country's current account

imbalance imposes costs on other countries. This paper has also reviewed some financial and real mechanisms through which such negative spillovers would likely arise. Because the spillovers are likely to be ignored by individual countries when choosing their own desired current account balances, multilateral policy actions may be warranted. However, the economic significance of these external effects, and thus the potential gains from policy coordination, are, in most cases, difficult to gauge. Moreover, if history is any guide, prospects do not seem promising that major countries would agree to set their macroeconomic policies and structural reforms to maintain their current account balances or real exchange rates at the levels deemed acceptable by other countries. Thus, not too much should be expected from policy coordination. The future of global imbalances is likely to be determined by whatever measures countries decide to adopt in their own self-interests. Whether such measures will result in a significant and sustained narrowing of global imbalances remains to be seen.

## Notes

We are indebted to Ramón Guzmán, Aart Kraay, Anton Korinek, William Martin, Martin Ravallion, Sergio Schmukler, the editor, and three anonymous referees for helpful comments on earlier drafts. We also thank conference participants at the 2010 Meeting of Latin American Economists and seminar participants at the Getulio Vargas Foundation and the Universidad de San Sebastián. Junko Sekine provided able research assistance. Any errors are our own responsibility. The views expressed here are ours alone and do not necessarily reflect those of the World Bank, its Executive Directors, or the countries that they represent.

1. See, e.g., [Meissner and Taylor \(2006\)](#) and [Bracke et al. \(2008\)](#).
2. However, the near balance of the EU's current account, shown in figure 1, conceals a large contrast between Germany, with persistently large current account surpluses, and the rest of the EU, with external deficits in recent years.
3. Thus, East Asia's rising surpluses had initially much more to do with the postcrisis investment slump than with the "saving glut" pointed out in [Bernanke \(2005\)](#). See, for example, [Blanchard and Milesi-Ferretti \(2009\)](#) and [Rajan \(2010\)](#).
4. [Eichengreen \(2006\)](#) and [Eichengreen and Park \(2006\)](#) summarize different views on the causes of global imbalances. The two perspectives highlighted in the text roughly correspond to those outlined by [Bracke et al. \(2008\)](#).
5. This has been termed the "exorbitant duty" of the United States by [Gourinchas, Rey, and Govillot \(2010\)](#), which is the counterpart in crisis times of the "exorbitant privilege" that the United States enjoys in normal times, described earlier in the text.
6. The widening growth differential between emerging and advanced countries, and thus their enhanced role for global growth—what has been recently termed "multipolarity" of the world economy—add to this trend; see [Caballero, Fahri and Gourinchas \(2008a\)](#).
7. There have been numerous empirical assessments of the degree of misalignment of the renminbi. For example, [Cline and Williamson \(2008\)](#) surveyed 18 studies, of which only one concluded that the renminbi was overvalued.
8. [Rajan \(2010\)](#) characterizes the current assignment of roles in the world economy in a similar way, adding the United Kingdom and some smaller EU economies among the "consumers of last

resort” and defining a third group of emerging markets building large net creditor positions for self-insurance reasons (see later in the text) or as a result of booming commodity prices (in the case of OPEC and other commodity exporters).

9. This is closely related to the explanations based on asset demand by Mendoza, Quadrini, and Rios-Rull (2009) and Carroll and Jeanne (2009), summarized previously. In both cases, the underlying force is the lack of adequate insurance mechanisms. The main difference is that the focus now is on country-level insurance against external shocks rather than on individual insurance against idiosyncratic shocks.

10. Total gross capital inflows into the United States were likewise dominated by nonofficial flows until 2007; see Borio and Disyatat (2011).

11. Chinn and Ito (2007) argue that the net impact of financial development on the current account depends on the degree of legal development and financial openness. For emerging countries with low levels of legal development and financial openness, financial development may be associated with higher surpluses.

12. In health, for example, reforms were announced in 2009 that aimed to achieve universal and affordable basic health care by 2020. In education, the government introduced a new long-term strategy in 2010 to improve schooling for students under the age of 6 years and over the age of 15 years.

13. In 2011, Medicare and Medicaid, along with Social Security, accounted for 43 percent of the U.S. federal budget. This figure is projected to increase in the future (Congressional Budget Office 2011).

14. The Chiang Mai Initiative is a multilateral currency swap arrangement among the ten members of the Association of Southeast Asia Nations, China (including Hong Kong), Japan and South Korea. This initiative draws from a foreign reserve pool worth US\$120 billion and was launched on March 24, 2010.

## References

- Adams, C., and D. Park. 2009. “Causes and Consequences of Global Imbalances: Perspective from Developing Asia.” *Asian Development Review* 26 (1): 19–47.
- Aizenman, J., and M. Hutchison. 2010. “Exchange Market Pressure and Absorption by International Reserves: Emerging Markets and Fear of Reserve Loss during the 2008-09 Crisis.” NBER Working Paper 16260. National Bureau of Economic Research, Cambridge, MA.
- Aizenman, J., and J. Lee. 2007. “International Reserves: Precautionary vs Mercantilist Views.” *Open Economies Review* 18: 191–214.
- Aizenman, J., and Y. Sun. 2009. “The Financial Crisis and Sizable International Reserves Depletion: From ‘Fear of Floating’ to the ‘Fear of Losing International Reserves?’” NBER Working Paper 15308. National Bureau of Economic Research, Cambridge, MA.
- Angeletos, G. 2007. “Uninsured Idiosyncratic Investment Risk and Aggregate Saving.” *Review of Economic Dynamics* 10 (1): 1–30.
- Baldacci, E., G. Callegari, D. Coady, D. Ding, M. Kumar, P. Tommasino, and J. Woo. 2010. “Public Expenditures in Social Programs and Household Consumption in China.” IMF Working Paper 10/69. International Monetary Fund, Washington, DC.
- Bernanke, B. 2005. “The Global Saving Glut and the U.S. Current Account Deficit.” Speech presented at the Sandridge Lecture, Virginia Association of Economics, Richmond, Virginia, March 10, 2005, and the Homer Jones Lecture, St. Speech, Board of Governors of the Federal Reserve System (US).

- Blanchard, O., and G. M. Milesi-Ferretti. 2009. "Global Imbalances: In Midstream?" IMF Staff Position Note SPN/09/29. International Monetary Fund, Washington, DC.
- . 2011. "(Why) Should Current Account Balances Be Reduced?" IMF Staff Discussion Note. International Monetary Fund, Washington, DC.
- Borio, C., and P. Disyatat. 2011. "Global Imbalances and the Financial Crisis: Link or No Link." BIS Working Paper 346. Bank for International Settlements, Basel, Switzerland.
- Bosworth, B., and S. Collins. 2010. "Rebalancing the U.S. Economy in a Post-Crisis World." ADBI Working Papers No 236. Asian Development Bank Institute, Tokyo, Japan.
- Bracke, T., M. Bussiere, M. Fidora, and R. Straub. 2008. "A Framework for Assessing Global Imbalances." European Central Bank Occasional Paper 78. Frankfurt am Main, Germany.
- Broner, E., G. Lorenzoni, and S. Schmukler. Forthcoming. "Why Do Emerging Economies Borrow Short Term?" *Journal of the European Economic Association*.
- Caballero, R. 2010. "The "Other" Imbalance and the Global Crisis." NBER Working Paper No. 15636. National Bureau of Economic Research, Cambridge, MA.
- Caballero, R., E. Farhi, and P. Gourinchas. 2008a. "An Equilibrium Model of Global Imbalances and Low Interest Rates." *American Economic Review* 98 (1): 358–93.
- . 2008b. "Financial Crash, Commodity Prices, and Global Imbalances." *Brookings Papers on Economic Activity* 2008: 1–55.
- Carroll, C., and O. Jeanne. 2009. "A Tractable Model of Precautionary Reserves, Net Foreign Assets, or Sovereign Wealth Funds." NBER Working Paper 15228. National Bureau of Economic Research, Cambridge, MA.
- Chamon, M., K. Liu, and E. Prasad. 2010. "Income Uncertainty and Household Savings in China." NBER Working Papers 16565. National Bureau of Economic Research, Cambridge, MA.
- Chamon, M., and E. Prasad. 2010. "Why are Saving Rates of Urban Households in China Rising?" *American Economic Journal: Macroeconomics* 2 (1): 93–130.
- Chinn, M., and H. Ito. 2007. "Current Account Balances, Financial Development and Institutions: Assaying the World's 'Saving Glut.'" *Journal of International Money and Finance* 26 (4): 546–69.
- Cline, W., and J. Williamson. 2008. "Estimates of the Equilibrium Exchange Rate of the Renminbi: Is there a Consensus and if Not, Why Not?" In M. Goldstein and N. Lardy, eds., *Debating China's Exchange Rate Policy*, 131–54. Washington, DC: Peterson Institute of International Economics.
- Congressional Budget Office. 2011. "The Budget and Economic Outlook: Fiscal Years 2011 to 2021." <http://www.cbo.gov/publication/21999>
- Cova, P., M. Pisani, N. Batini, and A. Rebucci. 2008. "Productivity and Global Imbalances: The Role of Non-Tradable Total Factor Productivity in Advanced Economies." *IMF Staff Papers* 55: 312–25.
- Curcuru, S., C. Thomas, and E. Warnock. 2008. "Current Account Sustainability and Relative Reliability." NBER Working Paper 14295. National Bureau of Economic Research, Cambridge, MA.
- De la Torre, A., S. Schmukler, and L. Servén. 2009. "Back to Global Imbalances." *World Bank Research Digest* 3 (4): 3–4.
- Devereux, M., and A. Sutherland. 2010. "Valuation Effects and the Dynamics of Net External Assets." *Journal of International Economics* 80 (1): 129–43.
- Devereux, M., and J. Yetman. 2010. "Leverage Constraints and the International Transmission of Shocks." *Journal of Money, Credit and Banking* 42:71–105.
- Dooley, M., D. Folkerts-Landau, and P. Garber. 2004. "The Revived Bretton Woods System: The Effects of Periphery Intervention and Reserve Management on Interest Rates and Exchange Rates

- in Center Countries.” NBER Working Paper 10332. National Bureau of Economic Research, Cambridge, MA.
- Durdu, C., E. Mendoza, and M. Terrones. 2007. “Precautionary Demand for Foreign Assets in Sudden Stop Economies: An Assessment of the New Mercantilism.” *Journal of Development Economics* 89:194–209.
- Eden, M. 2012. “Financial Distortions and the Distribution of Global Volatility.” World Bank Policy Research Working Paper 5929. Washington, DC.
- Eichengreen, B. 2004. “Global Imbalances and the Lessons of Bretton Woods.” NBER Working Paper 10497. National Bureau of Economic Research, Cambridge, MA.
- . 2006. “Global Imbalances: The New Economy, the Dark Matter, the Savvy Investor, and the Standard Analysis.” *Journal of Policy Modeling* 28:645–52.
- Eichengreen, B., and Y. Park. 2006. “Global Imbalances and Emerging Markets.” In J.J. Teunissen and A. Akkerman, eds., *Global Imbalances and the US Debt Problem: Should Developing Countries Support the US Dollar?*, 14–44. The Hague, The Netherlands: Fondad.
- Engel, C., and J. Rogers. 2006. “The U.S. Current Account Deficit and the Expected Value of World Output.” *Journal of Monetary Economics* 53:1063–93.
- Forbes, K. 2010. “Why do Foreigners Invest in the United States?” *Journal of International Economics* 80:3–21.
- Gourinchas, P., and H. Rey. 2007a. “International Financial Adjustment.” *Journal of Political Economy* 115:115.
- . 2007b. “From World Banker to World Venture Capitalist: The US External Adjustment and the Exorbitant Privilege.” In R. Clarida, ed., *G7 Current Account Imbalances: Sustainability and Adjustment*, 11–55. Chicago: University of Chicago Press.
- Gourinchas, P., H. Rey, and N. Govillot. 2010. “Exorbitant Privilege and Exorbitant Duty.” IMES Discussion Paper Series number 10-E-20. Institute for Monetary and Economic Studies, Tokyo, Japan.
- Habib, M. 2010. “Excess Returns on Net Foreign Assets: The Exorbitant Privilege from a Global Perspective.” European Central Bank Working Paper 1158. Frankfurt am Main, Germany.
- Hausmann, R., and F. Sturzenegger. 2006. “Global Imbalances or Bad Accounting?: The Missing Dark Matter in the Wealth of Nations.” Working Paper No. 124. Center for International Development, Harvard University, Cambridge, MA.
- Jeanne, O., and R. Ranciere. 2011. “The Optimal Level of International Reserves for Emerging Market Countries: A New Formula and Some Applications.” *Economic Journal* 121:905–30.
- Korinek, A. 2012. “Capital Controls and Currency Wars.” Working Paper. University of Maryland, College Park, MD.
- Korinek, A., and L. Servén. 2010. “Undervaluation through Foreign Reserve Accumulation: Static Losses, Dynamic Gains.” World Bank Policy Research Working Paper 5250. Washington, DC.
- Kraay, A., N. Loayza, L. Servén, and J. Ventura. 2005. “Country Portfolios.” *Journal of the European Economic Association* 3:914–45.
- Kraay, A., and J. Ventura. 2005. “The Dot-Com Bubble, the Bush Deficits, and the US Current Account.” NBER Working Paper No. 11543. National Bureau of Economic Research, Cambridge, MA.
- Laibson, D., and J. Mollerstrom. 2010. “Capital Flows, Consumption Booms and Asset Bubbles: A Behavioural Alternative to the Savings-Glut Hypothesis.” NBER Working Paper 15759. National Bureau of Economic Research, Cambridge, MA.

- Loayza, N., K. Schmidt-Hebbel, and L. Servén. 2000. "What Drives Private Saving Across the World?" *Review of Economics and Statistics* 82:165–81.
- Meissner, C., and A. Taylor. 2006. "Losing our Marbles in the New Century?: The Great Rebalancing in Historical Perspective." NBER Working Paper 12580. National Bureau of Economic Research, Cambridge, MA.
- Méjean, I., P. Rabanal, and D. Sandri 2011. "Current Account Rebalancing and Real Exchange Rate Adjustment between the U.S. and Emerging Asia." IMF Working Paper 11/46. International Monetary Fund, Washington, DC.
- Mendoza, E., V. Quadrini, and J. Rios-Rull. 2009. "Financial Integration, Financial Deepness and Global Imbalances." *Journal of Political Economy* 117:371–416.
- Mussa, M. 2004. "Exchange Rate Adjustments Needed to Reduce Global Payments Imbalances." In C.F. Bergsten and J. Williamson, eds., *Dollar Adjustment: How Far? Against What?*, 113–38. Washington, DC: Institute for International Economics.
- Obstfeld., M., and K. Rogoff. 2007. "The Unsustainable U.S. Current Account Position, Revisited." In R.H. Clarida, ed., *G7 Current Account Imbalances: Sustainability and Adjustment*, 339–376. Cambridge, MA: National Bureau of Economic Research.
- \_\_\_\_\_. 2009. "Global Imbalances and the Financial Crisis: Products of a Common Cause." Discussion Papers 7606. Center for Economic and Policy Research, London, UK.
- OECD 2011. "The Impact of Structural Reforms on Current Account Imbalances." OECD Economics Department Policy Notes No. 3. Organisation for Economic Co-operation and Development, Paris, France.
- Portes, R. 2009. "Global Imbalances." In M. Dewatripont, X. Freixas, and R. Portes, eds., *Macroeconomic Stability and Financial Regulation*, 19–26. London, UK: Center for Economic and Policy Research.
- Prasad, E. 2011. "Rebalancing Growth in Asia." *International Finance* 14 (1): 27–66.
- Rajan, R. 2010. *Fault Lines*. Princeton, NJ: Princeton University Press.
- Rodrik, D. 2008. "The Real Exchange Rate and Economic Growth." *Brookings Papers on Economic Activity* 2008:365–412.
- Roubini, N. 2008. "Will the Bretton Woods 2 (BW2) Regime Collapse Like the Original BrettonWoods Regime Did? The Coming End Game of BW2." <http://www.economonitor.com/nouriel/2008/07/06/will-the-bretton-woods-2-bw2-regime-collapse-like-the-original-bretton-woods-regime-did-the-coming-end-game-of-bw2/>
- Sandri, D. 2010. "Growth and Capital Flows with Risky Entrepreneurship." IMF Working Paper WP/10/37. International Monetary Fund, Washington, DC.
- Taylor, J. 2009. *Getting Off Track*. Stanford, CA: Hoover Institution Press.

---

# Multidimensional Poverty Analysis: Looking for a Middle Ground

---

Francisco H. G. Ferreira • Maria Ana Lugo

---

*Widespread agreement that poverty is a multifaceted phenomenon encompassing deprivations in multiple dimensions clashes with the vociferous disagreement about how best to measure these deprivations. Drawing on the recent literature, this short paper reviews three methodological alternatives to the false dichotomy between scalar indices of multidimensional poverty, on the one hand, and a “dashboard” approach that considers only marginal distributions, on the other. These alternatives include simple Venn diagrams of the overlap of deprivations across dimensions, multivariate stochastic dominance analysis, and the analysis of copula functions, which capture the extent of interdependency across dimensions. Examples are provided from the literature on both developing and developed countries. JEL codes: I32, O15*

Over the last 10 years, interest in multidimensional poverty measurement has grown steadily. Since the pioneering works of Bourguignon and Chakravarty (2003) and Tsui (2002), a number of approaches have been proposed to measure or analyze deprivation in multiple dimensions. This rapidly growing literature now includes Alkire and Foster (2011a), Chakravarty, Deutsch, and Silber (2008), Deutsch and Silber (2005), Duclos, Sahn, and Younger (2006), and Maasoumi and Lugo (2008), among others.

Multidimensional poverty analysis has been transformed from a purely academic discussion into a broad domestic and international policy debate, both within and between many countries. In December 2009, for example, Mexico’s National Council for the Evaluation of Social Policy (CONEVAL) adopted a multidimensional index as the country’s official poverty measure.<sup>1</sup> In 2011, the government of Colombia followed suit by adopting a poverty-reduction strategy that focused on five separate dimensions and relied on a variant of Alkire and Foster’s (2011a)

approach to quantify progress in the reduction of poverty. Internationally, the Multidimensional Poverty Index of [Alkire and Santos \(2010\)](#), which was reported for over 100 countries in the UNDP's *Human Development Report 2010*, has also gained prominence ([UNDP 2010](#)).

The driving force behind this rising popularity is the acknowledgment that poverty involves much more than just low income. Low consumption and inadequate living standards are at the heart of the concept of “poverty,” to be sure. The associations conjured up by that term also include aspects of poor health, such as a shortened lifespan; limited access to education, knowledge, and information; and powerlessness in various domains. Poor people often mention nonincome dimensions as being crucial to their perceptions of their own hardships. The following quote from a person from Georgia illustrates this point: “Poverty is lack of freedom, enslaved by crushing daily burden, by depression and fear of what the future will bring” (quoted in [Narayan et al. 2000](#), p. 37). It is now quite common for poverty analysts, whether from academia, the World Bank, or other agencies, to face enquiries about how best to summarize information on “multidimensional poverty” in a particular country.

However, the various new multidimensional poverty indices have not been universally welcomed. Serious criticism of these indices centers on the manner in which information on deprivation is aggregated across dimensions. One powerful critique is that the weights used to aggregate across dimensions lack the intrinsic meaning associated with prices, which are used to add the components of consumption expenditure (or, implicitly, its dual, the incomes used to finance consumption). Under the law of one price, and given relatively weak assumptions on preferences, relative prices are equal to the rate at which consumers themselves—regardless of their income levels and allowing for different utility functions—are willing to trade one such component (e.g., bread) for another (e.g., a bicycle).

Of course, there are a number of practical reasons why prices may not be ideal welfare weights. These reasons range from the existence of externalities to the fact that price data are often geographically coarse, so actual price variation in space is missing from the information available to the researcher. Nevertheless, as [Ravallion \(2011, p. 247\)](#) argues,

It is widely agreed that prices can be missing for some goods and deceptive for others. There are continuing challenges facing applied economists in addressing these problems. However, it is one thing to recognize that markets or prices are missing or imperfect, and quite another to ignore them in welfare and poverty measurement. There is a peculiar inconsistency in the literature on multidimensional indices of poverty whereby prices are regarded as an unreliable guide to the tradeoffs, and are largely

ignored, while the actual weights being assumed in lieu of prices are not made explicit in the same space as prices.

Because multidimensional indices adopt arbitrary weights, which are often equal weights across dimensions, they inherently involve specific tradeoffs between the constituent components of welfare in the mathematical sense: A certain extra amount of one component will exactly offset the change in another component to leave the index unchanged. These tradeoffs are seldom stated explicitly, and it is not obvious that they are frequently revised by public debate (Ravallion 2011).

## What Is the Disagreement?

The debate between proponents and skeptics of multidimensional poverty measurement was featured in the Forum Section of the June 2011 issue of the *Journal of Economic Inequality* (e.g., Alkire and Foster 2011b; Ravallion 2011; Lustig 2011). These articles suggest strong agreement on at least one basic point: “poverty is multidimensional” (Ravallion 2011, p. 236). There is little dispute that deprivation exists in multiple domains, which are often correlated. Therefore, considering information on these various dimensions (rather than on incomes or consumption expenditures only) is likely to be useful for designing policies that effectively combat poverty.

The disagreement between proponents and skeptics involves how best to *measure* this multidimensional poverty—that is, how best to convey information about the extent of these various deprivations in a way that is useful for both analysts and policy makers. Some studies, such as those by Alkire and Foster (2011a) and Maasoumi and Lugo (2008), have proposed scalar indices that seek to combine information from various dimensions into a single number. A key advantage of such scalar indices is that they generate a complete ordering of countries, regions, or individuals, even when the rankings conflict across individual dimensions. The Multidimensional Poverty Index (or the Human Development Index, in the space of attainments) attracts a good deal of international attention, in large part because it ranks countries according to how well they perform on various dimensions through a simple summarizing tool.

In contrast, Ravallion (2011) suggests a “dashboard” approach, whereby “we may need to focus our efforts and resources on developing the best possible *distinct* measures of the various dimensions of poverty [...] aiming for a credible set of ‘multiple indices’ rather than a single ‘multidimensional index’” (Ravallion 2011, p. 13; our emphasis). An important limitation of multidimensional indices, as noted earlier, is that they require the use of relative weights for each dimension, which are chosen somewhat arbitrarily by the analyst. Other analysts, policy makers, and the public may disagree with these specific weights. Instead of

imposing a specific weighting system, the dashboard approach allows the user to attach more or less importance to any particular dimension.

## Is There a Policy-relevant Middle Ground for Multidimensional Poverty Analysis?

In this short paper, we argue that this debate—a single index versus a dashboard—is a false dichotomy. In other words, we suggest that the analysis of multidimensional poverty can and should move beyond both scalar indices and dashboards of deprivation based on marginal distributions. Multidimensional poverty analysis is interesting because the joint distribution of achievements provides more information than that provided by its margins. The dependency structure in a joint distribution—how closely correlated various achievements (or deprivations) are—can affect how we assess poverty in a society or how we compare it across time periods, even given identical margins. A dashboard that reports on poverty indices for each dimension separately may overlook this aspect of the joint distribution.

This point was made eloquently by [Duclos, Sahn, and Younger \(2006\)](#). In one instance, these authors compare the joint distributions of two dimensions of health (nutrition and survival probabilities) in Cameroon and Madagascar and find that *both* marginal distributions in Cameroon first-order dominate the corresponding marginal distributions in Madagascar. However, there is no dominance of the *joint* distribution. In this example, a dashboard approach would lead a researcher to naively conclude that poverty was unambiguously greater in Madagascar, whereas a truly multidimensional assessment (i.e., one that took into account the correlations) would enable the researcher to conclude that no clear ranking was possible.

The discrepancy arises because the correlation between the two dimensions may differ substantially from one place to another, so the cumulative concentration of deprivations could make overall poverty worse in a place that has better marginal distributions. As [Duclos et al.](#) put it, “It is possible for a set of univariate analyses done independently for each dimension of well-being to conclude that poverty in A is lower than poverty in B while a multivariate analysis concludes the opposite, and vice-versa. The key to these possibilities is the interaction of the various dimensions of well-being in the poverty measure and their correlation in the sampled populations” ([Duclos, Sahn and Younger 2006](#), p. 945).

In the remainder of this paper, we briefly describe three alternative empirical approaches for the analysis of the dependency structure in joint distributions. All three illustrations are drawn from the recent literature, and we make no claim to

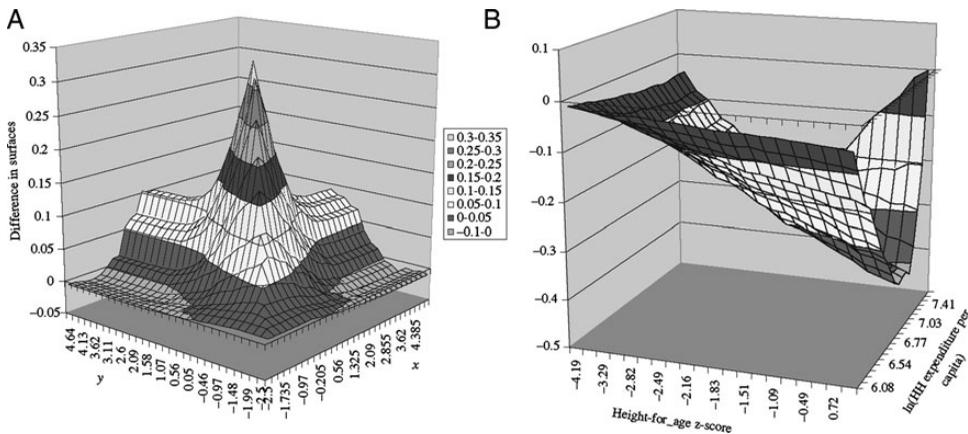
originality. Each is suitable to a specific purpose, but all of the approaches focus on interactions among the dimensions, so the menu of options may be useful not only to analysts but also to policy makers.

The first approach is the set of multivariate stochastic dominance techniques proposed by Duclos et al. (2006), which enable poverty analysts to investigate joint distributions of multiple deprivations without making the specific assumptions about tradeoffs that understandably worry Ravallion (2011). If the correlation between deprivations matters—as surely it must—then this approach must be regarded as superior to the dashboard approach, which considers only the marginal distributions.

Multivariate stochastic dominance compares two multidimensional distributions to determine whether one distribution dominates the other (i.e., consistently lies always above or always below) for all reasonable poverty frontiers. If this is the case, then one could conclude that poverty in, say, A is always higher (or lower) than poverty in B for all additive poverty indices. Below, we reproduce two examples given in Duclos et al. (2006). These graphs represent the bidimensional dominance surfaces—that is, the difference between two joint distributions.

The graph on the left of Figure 1 represents the difference in the surfaces of two hypothetical distributions. In this case, “[a]lthough differences in the univariate dominance curves in both dimensions clearly cross the origin (at the extreme left and right of the figure), there is a significant interior section where the first surface is entirely above the second” (Duclos et al. 2006, p. 954). In other words, there is intersection dominance without marginal dominance.

**Figure 1.** Differences in Dominance Surfaces



Source: Reprinted from Duclos, J.-Y., D. Sahn, and S. Younger. 2006. “Robust Multidimensional Poverty Comparisons.” *Economic Journal* 116 (514): 943-68.

The graph on the right of Figure 1 depicts the dominance surface comparing rural and urban children in Vietnam on two dimensions: household expenditure per capita and height-for-age Z scores. This case indicates that over almost the entire range of expenditures and stunting, rural children are poorer than their urban counterparts. This dominance is found for all reasonable poverty lines, so the finding that “rural children are poorer than urban ones is valid for almost any intersection, union or intermediate poverty frontier” (Duclos et al. 2006, p. 959).

Of course, it is true that multivariate stochastic dominance analysis tends to be of limited use when the number of dimensions or margins is large. In these cases, a seldom-discussed natural alternative involves complementing the dashboard approach with a direct representation of the dependency structure. In poverty studies, the degree of interdependence can be presented in terms of the extent of the overlap between individuals who are identified as deprived under the various criteria. For instance, if poverty were defined by three dimensions (such as education, health, and income), the dependency could be illustrated—at least in part—by the proportion of individuals who were deprived in all three dimensions, those who were deprived in (different) pairs of dimensions, or those who were deprived in only one dimension. Atkinson and Lugo (2010) provide such an example for the case of Tanzania.

Table 1 below (reproducing Table 5 in Atkinson and Lugo 2010) presents information on deprivations in three dimensions of well-being measured at the household level and the extent of overlap between them. The dimensions chosen are school attendance of children between 5 and 16 years old, access to safe sources of drinking water (piped or protected), and an indicator of ownership of durable assets. These indicators are closely related to the goals set by the government in the Tanzanian National Strategy and are obtained at the household level from the (same) Household Budget Surveys. The example shows the following:

Between 2001 and 2007, school attendance and availability of durable assets have improved significantly, whereas access to protected sources of drinking water has deteriorated. Despite the latter, the combined effect is to reduce the proportion of Tanzanians who suffer from any of the three forms of deprivation: this has fallen from 90 per cent to 80 per cent. Equally there has been a fall in the proportion deprived on all 3 dimensions: from 19 per cent to 10 per cent. At the same time, there has been an increase in one category of the deprived: those lacking only access to water (Atkinson and Lugo 2010, p. 15).

The implication of these results for assessing the progress of the country between the two years is clear: “If it were decided that access to water were the sole concern, then deprivation would have increased from 46 per cent in 2001 to 51

**Table 1.** Deprivations in schooling, access to protected water and durable assets in Tanzania

<i>Proportion of the individuals living in households . . .</i>	<i>2001</i>	<i>2007</i>
School deprived: at least one child 5–16 years old not in school	55.2	34.5
Assets deprived: no car and fewer than on "small asset"	66.6	47.4
Water deprived: no access to piped or protected source of drinking water	45.8	50.7
Distributions of individuals		
Not deprived in school, water or assets	10.3	20.1
Only school deprived	8.9	7.5
Only water deprived	4.7	14.4
Only assets deprived	17.3	15.1
School and water deprived	9.5	10.7
Water and assets deprived	12.5	15.9
School and assets deprived	17.6	6.6
School, water and assets deprived	19.1	9.8

*Source:* HBS 2001 and 2007. Reproduced from Atkinson and Lugo (2010), "Growth, poverty and distribution in Tanzania." International Growth Centre, Working Paper 10/0831. November. Page 15.

*Note:* small assets include television, radio, telephone (including mobile phones), refrigerator bicycle, and motorcycle.

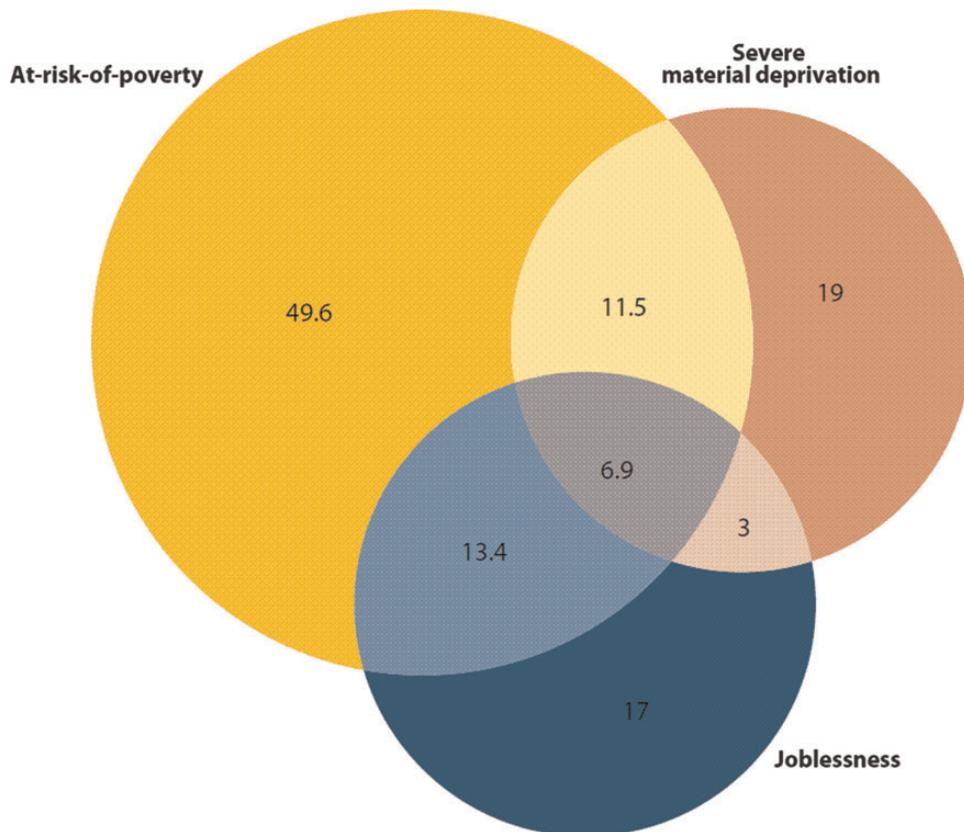
per cent in 2007. [...] [If instead, one wanted to] assess overall performance simply in terms of the proportion deprived on all dimensions [...] this proportion has fallen, indicating definite progress" (Atkinson and Lugo 2010, p. 15).

An effective way of diagrammatically presenting this kind of information on the degree of overlap across dimension-specific deprivations is through Venn diagrams, as Atkinson et al. (2010) do for the EU-27, the 27 countries of the European Union. The larger the overlap between deprivations is, the greater the extent of interdependence will be. Figure 2 (reprinted from Atkinson et al. 2010) shows the number of people who are at risk of poverty (EU definition), the number of people who are materially deprived, and the number of people aged 0–59 years who are living in jobless households.<sup>2</sup> Data come from European Union Statistics on Income and Living Conditions surveys. The authors note, "A little over 80 million people live in households at risk of poverty, a further 40 million live in households that are not at risk of poverty but are defined as jobless and/or materially deprived." Indeed, "well over two-thirds are identified under only one of the criteria." This especially insightful example suggests that, in this particular context, policies that are directed exclusively toward one of the indicators may fail to reduce the degree of deprivation of a large proportion of households. The authors astutely conclude that it is important not only to monitor the three indicators but also to understand that the degree of overlap among them will help to shape policies to address these shortfalls.

---

**Figure 2.** Multiple Indicators from the Europe 2020 Target

---



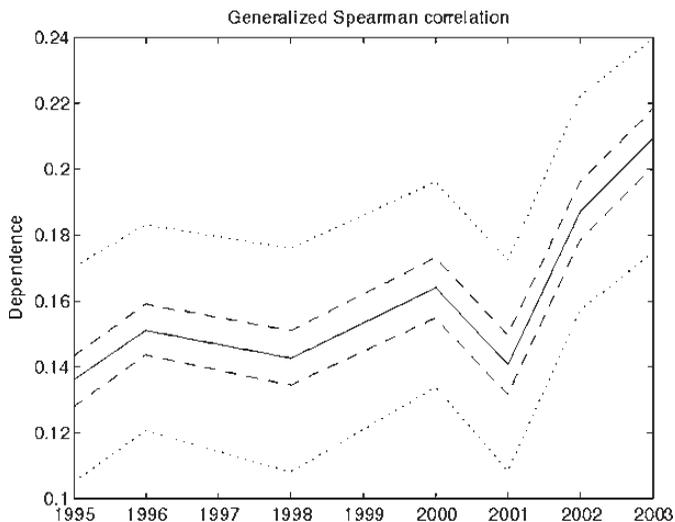
Note: Figures for EU-27 in million of persons.

Source: EU-SILC, Eurostat-CEPS/INSTEAD calculations. Reprinted from Atkinson, A. B., E. Marlier, F. Monatigne, and A. Reinstadler. 2010. "Income Poverty and Income Inequality." In A.B. Atkinson and E. Marlier, eds., *Income and Living Conditions in Europe*, 127. Luxemburg: Eurostat.

---

When the objective is to evaluate well-being rather than deprivation, the representation of interdependencies between dimensions is less straightforward, and some "lateral thinking" (Atkinson 2011) might be necessary. Sklar's Theorem, from the statistics literature, tells us that any joint distribution function can be decomposed into the marginal distribution functions of each dimension (corresponding to the dashboard approach) and a copula function, which captures the degree of interdependency between the dimensions. Copulas, which are the third approach we want to highlight, have been used to study the relationship between health and income (Decancq 2009, Quinn 2007). Formally, the copula function is the joint distribution function of  $n$  vectors whose elements are the *relative position* of every individual in each of the  $n$  dimensions of well-being. As in the case

**Figure 3.** The Evolution of the Dependence between the Dimensions of Well-being over Time in Russia



Source: RLMS 1995-2003. Reprinted from Decancq, K. 2009. "Copula-based Measurement of Dependence Between Dimensions of Well-being." HEDG Working Paper 09/32. University of York, Health Economics Resource Centre, York, UK.

of the margins, once the copula function is constructed, a stochastic dominance analysis could be performed to assess whether the degree of interdependency between the components has changed in an unambiguous way. In addition, and in the same way that information on unidimensional inequality can be summarized by an index (such as the Gini coefficient or the mean log deviation), measures of rank correlation allow us to order distributions unambiguously in terms of their degree of interdependency (Decancq 2009).

As an example, we reproduce from Decancq (2009) a graph showing the evolution of the Spearman correlation coefficient, a commonly used measure of rank correlation, in Russia between 1995 and 2003 (Figure 3). The author considers three dimensions, standard of living, health, and schooling, represented respectively by household income, self-assessed health, and years of schooling as primary ranking variables (three other variables are used as secondary ranking variables in case of ties). For context, the Russian Human Development Index increased from the beginning to the end of the period, albeit with a deterioration around the time of the 1998 financial crises.

In this figure, the dashed line represents the 95 percent confidence interval obtained by Monte Carlo simulations, whereas the dotted line represents an alternative computation of the same confidence interval by bootstrapping. The figure

shows that the degree of dependence across these three dimensions of well-being has increased throughout most of the period; as average well-being was rising, so was the degree of interdependency between the dimensions of well-being.

Although they differ in technical complexity, these three alternative techniques (Venn diagrammatic representations of the dependency structure, multivariate stochastic dominance, and the analysis of copula functions) seem to avoid the disadvantages associated with both the scalar indices and the dashboard approach. Like the dashboard approach (but unlike scalar indices), these techniques do not require the use of generally arbitrary weights to aggregate across dimensions, with their unpalatable implications in terms of tradeoffs. Like scalar indices (but unlike the dashboard approach), these techniques incorporate information about how deprivations are *jointly* distributed and allow analysts to take into account different levels or changes in the extent of overlap or correlation between them. This is a menu of analytical approaches that represents an advantageous middle ground in multidimensional poverty analysis and that may add value to some of the poverty analyses currently undertaken at the World Bank (and elsewhere).

There is one important caveat, however. This kind of analysis requires that information on the various dimensions be observed for each unit of observation (typically, the individual or the household). In other words, the dimensions must be observed in the same survey (or census) or, at least, in different surveys that cover the same set of households and contain common identifiers. Otherwise, it is clearly impossible to observe the joint distribution.

An especially important argument proposed by [Ravallion \(2011\)](#) for the dashboard approach is that the best data on separate dimensions (say, health status and consumption expenditures) are often found in different data sets, from which no joint distribution can be constructed. In these cases, a tradeoff may arise between data quality and information on the joint distribution, and such tradeoffs must be evaluated on a case-by-case basis. If the quality of information is thought to differ only marginally—if, for example, anthropometric information about children in a health survey is thought to be superior to that contained in an Living Standards Measurement Study-type survey, but the latter has some information—some analysts may be sufficiently interested in the extent of the overlap of deprivation in health and consumption to incur the cost of using the slightly worse data. Conversely, analysts who place sufficient weight on the accuracy of the information on marginal distribution would choose to analyze each margin separately, drawing from the best survey in each case.<sup>3</sup>

In many cases, of course, reliable information on relevant dimensions, such as health status, anthropometrics, education, and consumption, exists in the same survey. This is the case, for instance, for the Russia Longitudinal Monitoring Survey used by [Decancq \(2009\)](#), the Indonesia Family Life Survey, and various Living Standards Measurement Study-type surveys. In these cases, it is difficult to

find a justifiable excuse for any analyst who fails to complement his marginal analysis with information on the dependency structure. Furthermore, one should not underestimate the power of the demand for better data. If there are certain aspects of well-being and deprivation that are not regularly captured in household surveys but whose joint distribution with other dimensions (such as income) is of real policy interest, then more frequent analysis of the kind we suggest might encourage statistical institutes or other data providers to collect information on such aspects.

## What about Scalar Multidimensional Poverty Indices?

Finally, we turn briefly to the question of whether the analytical approaches we propose should completely preclude the computation of scalar multidimensional poverty indices. It is clear that they do not preclude the dashboard approach in that considering the marginal distributions is inherent in considering the joint distributions. The three approaches described above are best understood as complementing the dashboard approach with information on the dependency structure between the dimensions, when this information is available.

What role do we see for the definition and computation of *scalar* multidimensional poverty *indices*? We see this question as another instance of choice given a tradeoff. A scalar multidimensional index provides a complete ordering, with the ability to rank two years, countries, or regions, even when their joint distributions (or copulas) cross. Just as in unidimensional poverty or inequality analysis, the ability to generate a complete ordering comes at a cost in terms of specific functional form assumptions. In the case of multidimensional well-being and poverty, this price is high, for two key reasons.

The first reason is that the identification step (in the sense of [Sen 1976](#)) is considerably more complex for multidimensional poverty than for unidimensional poverty. This complexity results not only because one has to define a threshold for each individual dimension but also because a difficult choice must be made at the identification step about how many deprivations constitute poverty. This situation recalls the fundamental choice between the union and intersection approaches (e.g., [Bourguignon and Chakravarty 2003](#)), the establishment of an intermediate cutoff in the number of dimensions ([Alkire and Foster 2011a](#)), or whether the depth of deprivation in one dimension should be allowed to offset well-being in another dimension at the identification stage (e.g., [Duclos et al. 2006](#)). In this paper, we have largely ignored issues of identification, although these issues are clearly important if one decides to pursue the scalar-index route.

The second reason is the need for weights to aggregate across dimensions. Even conditional on a particular identification algorithm, the issue of weights remains.

Whether a price is worth paying is therefore likely to depend on, first, the importance of the ability to rank for the purpose of the analysis at hand and, second, the arbitrariness (and number) of the weights.

Therefore, it is difficult to take a definitive position in the abstract. As in the case of unidimensional poverty measurement, it seems that two extreme positions are untenable. The first position would be to say that a particular multidimensional poverty index is *the* one true measure of poverty. The second position would be to argue that *any* particular index that makes an unpalatable assumption is inadmissible. In the same way that we recognize the limitations of the headcount index but continue to report it (ideally, alongside other measures and more disaggregated and robust analysis), there are clear uses for various indices of multidimensional poverty.

In our view, such scalar indices would be most useful if they relied, to the greatest possible extent, on (shadow or market) prices to aggregate across different goods and services. Only the aspects of well-being for which there can truly be no sensible estimate of relative prices should be treated as separate dimensions. Although food, cooking utensils, toilets, clothing, and vehicles may be resources that affect different “functionings” (in the sense of Amartya Sen), they are best treated as components of a single dimension of well-being—command over private goods—whose internal weights are given by relative prices.

Equilibrium prices, as noted earlier (and in elementary first-year economics courses), reflect people’s preferences and constraints as well as market structures. If individual choices can be reasonably approximated by the result of maximizing a utility function subject to a budget constraint, and if the law of one price holds, then it follows that (regardless of people’s individual utility functions or income levels) the rate at which every person trades off one unit of a good against one unit of another good will be the same. In other words, relative prices reflect the marginal rate of substitution between the goods.

Naturally, price data are often problematic. Very often, the available data on prices do not accurately represent the prices actually encountered by individuals for the indicators that we would include in the analysis. This is largely because there is price heterogeneity across regions (or even cities) within a country and in the quality of products and services, whereas price data tend to be collected at a higher level of aggregation and for a representative (average) item in each category. However, it is questionable whether these problems are so severe that choosing arbitrary weights across dimensions, such as the availability of cooking utensils and, say, the ownership of certain means of transport, would yield a preferable metric. In this regard, Ravallion’s (2011) arguments for greater reliance on market prices seem overwhelmingly compelling.

The true value of the analysis of multidimensional poverty (or well-being) lies in the existence of certain aspects of well-being that we deem important but for which there can be no sensible estimate of relative prices. It is reasonable to

include in this category items such as political and personal freedoms, health, and, arguably, education. Although various inputs into the production of health and education are marketable, this is not true of all of them. Health is influenced by environmental quality and by a number of other public goods. Education, once embodied in human capital, generates so many externalities that it is difficult to think that school fees or costs are suitable approximations of its true shadow price. These things—not the material from which a ceiling is made or the kind of stove one uses—are the true *dimensions* of welfare.

In these cases, multidimensional analysis becomes particularly relevant. If, for the purpose at hand, the analyst decides that the price of selecting weights is worth paying (to obtain a complete ordering), then the question is whether one can choose relative weights that at least attempt to represent the existing tradeoff between the different components of deprivation (or well-being). There are various approaches to setting these weights; some are based exclusively on the observed distribution of attributes, others are based on people's opinions, and others use both sorts of information (Decancq and Lugo, 2013). Regardless of the weighting scheme and the precise functional form chosen, multidimensional indices of deprivation (or well-being) should be allowed to be sensitive to the essence of the multidimensional approach, that is, to the degree of dependency between its components.

## Conclusions

There is widespread agreement that poverty is a multifaceted phenomenon. Income shortfalls, which translate into an inability to consume certain basic commodities, are central to this phenomenon. However, income poverty is typically associated with deprivation in other realms, such as health, education, social status, and political power, which are more difficult to price. These associations or correlations between the constituent dimensions of poverty vary over time and from place to place, and they are often believed to be significant. Recent advances in multidimensional poverty analysis seek to capture these interactions, and revealed preference seems to suggest that they are of interest to policy makers in many developing countries.

Despite this widespread agreement on the essential fact that poverty is multidimensional, there has been a lively debate about whether this implies that scalar indices should be constructed that summarize information on these various dimensions into a single number or whether multiple indices should be provided, one for each dimension, in a dashboard approach. Drawing on the existing literature, we have argued here that such a dichotomous view misses the point. The most interesting aspects of the multidimensionality of poverty arise from the

interdependence among dimensions. The joint distribution of dimensions over the population contains more information than the corresponding marginal distributions, and the correlation patterns in that joint distribution may change how we compare poverty across two countries or time periods. This dependency structure is overlooked entirely by the dashboard approach and is often obscured by scalar indices.

Drawing on examples from both developed and developing countries, we provide three alternative approaches that allow researchers and policy analysts to focus on the dependency structure of a joint distribution. The first approach is stochastic dominance analysis, which permits partial orderings across joint distributions that are robust not only to poverty lines and welfare weights (as in the unidimensional case) but also to dimension weights. The second approach is a representation of the overlap of deprivations over the population by means of simple tabulations or Venn diagrams. Given the agreement on the identification criterion along each dimension, this extremely simple tool can complement the dashboard statistics on the marginal distributions in informative ways. The third approach involves the use of copula functions to study the multivariate association among different components of well-being across two or more joint distributions.

Finally, we argue that multidimensional poverty indices, like most other tools, can be accommodated in the economist's toolkit, and the risk of serious injury decreases with reliance on relative prices and a focus on a few core, truly irreducible dimensions. If such indices are used, fodder will remain for future controversies on the choice of weights, poverty lines, and functional forms, but these controversies would largely be a distraction from what really matters to policy makers: the pattern of associations and overlaps across the core dimensions of well-being.

## Notes

Francisco H. G. Ferreira is with the Development Research Group at the World Bank and the Institute for the Study of Labour (IZA); email address: fferreira@worldbank.org. Maria Ana Lugo is in the Poverty, Equity and Gender Group at the Latin America Region of the World Bank. This paper is an expanded version of a short comment that was published as [Ferreira \(2011\)](#). We are grateful to three anonymous referees and to Sabina Alkire, Peter Lanjouw, Nora Lustig, and Martin Ravallion for comments on earlier versions of the paper. We are also grateful for illuminating conversations with James Foster on this subject. All remaining errors are ours. This paper is a product of the Equity and Development Research Project (P099861). The views expressed here are those of the authors. They should not be attributed to the World Bank, its Executive Directors, or the countries they represent.

1. The Executive Secretary of the National Council for the Evaluation of Social Policy, Dr. Gonzalo Hernández Licona, is quoted as saying, "Mexico is proud to be the first country in the world to measure poverty, not narrowly on economic grounds alone, but to take full account of crucial social components of poverty such as quality of housing and access to healthcare and food,

which are all too often neglected by established poverty measures.” See <http://www.ophi.org.uk/launch-of-mexico%E2%80%99s-new-poverty-measure/>.

2. Persons “at risk of poverty” are defined as those who have an equivalized disposable income below 60 percent of the national median equivalized disposable income, after social transfers. Material deprivation covers indicators relating either to economic strain or to the ownership of durables. Severely materially deprived persons cannot afford at least four of the following: to pay rent or utility bills; to keep their home adequately warm; to pay unexpected expenses; to eat meat, fish, or a protein equivalent every second day; a one-week holiday away from home; a car; a washing machine; a color TV; or a telephone. Finally, a “jobless household” is one in which none of the members aged 18–59 years are working or the members aged 18–59 years have very limited work attachment.

3. The question of whether data on a particular variable are more reliable in one survey than in another may, in some cases, be investigated empirically. At a minimum, survey-to-survey imputation techniques can be used to assess the sensitivity of results to using information from different sources. This remains an area for future work.

## References

- Alkire, S., and J. Foster. 2011a. “Counting and Multidimensional Poverty Measurement.” *Journal of Public Economics* 95 (7): 476–87.
- . 2011b. “Understandings and Misunderstandings of Multidimensional Poverty Measurement.” *Journal of Economic Inequality* 9 (2): 289–314.
- Alkire, S., and M.E. Santos. 2010. “Acute Multidimensional Poverty: A New Index for Developing Countries.” OPHI Working Paper Series #38. Oxford University, Department of International Development, Oxford Poverty and Human Development Initiative, Oxford, UK.
- Atkinson, A.B. 2011. “On Lateral Thinking.” *Journal of Economic Inequality* 9 (3): 319–28.
- Atkinson, A.B., E. Marlier, F. Monastigne, and A. Reinstadler. 2010. “Income Poverty and Income Inequality.” In A. Atkinson and E. Marlier, eds., *Income and Living Conditions in Europe* 287–398. Luxemburg: Eurostat.
- Atkinson, A.B., and M.A. Lugo. 2010. “Growth, Poverty and Distribution in Tanzania.” International Growth Centre Working Paper 10/0831. London School of Economics and Political Science and Oxford University, Department for International Development, London and Oxford, UK.
- Bourguignon, E., and S. Chakravarty. 2003. “The Measurement of Multidimensional Poverty.” *Journal of Economic Inequality* 1 (1): 25–49.
- Chakravarty, S., J. Deutsch, and J. Silber. 2008. “On the Watts Multidimensional Poverty Index and its Decomposition.” *World Development* 36 (6): 1067–77.
- Decancq, K. 2009. “Copula-Based Measurement of Dependence Between Dimensions of Well-Being.” HEDG Working Paper 09/32. University of York, Health Economics Resource Centre, York, UK.
- Decancq, K., and M.A. Lugo. 2013. “Weights in Multidimensional Indices of Well-Being: An Overview.” *Econometric Reviews* 32 (1): 7–34.
- Deutsch, J., and J. Silber. 2005. “Measuring Multidimensional Poverty: An Empirical Comparison of Various Approaches.” *Review of Income and Wealth* 51 (1): 145–74.
- Duclos, J.-Y., D. Sahn, and S. Younger. 2006. “Robust Multidimensional Poverty Comparisons.” *Economic Journal* 116 (514): 943–68.
- Ferreira, F.H.G. 2011. “Poverty is Multidimensional. But What Are We Going To Do About It?” *Journal of Economic Inequality* 9 (3): 493–95.

- Lustig, N. 2011. "Multidimensional Indices of Achievements and Poverty: What Do We Gain and What Do We Lose? An Introduction to the JOEI Forum on Multidimensional Poverty." *Journal of Economic Inequality* 9 (2): 227–34.
- Maasoumi, E., and M.A. Lugo. 2008. "The Information Basis of Multivariate Poverty Assessments." In N. Kakwani and J. Silber, eds., *Quantitative Approaches to Multidimensional Poverty Measurement*. New York: Palgrave Macmillan.
- Narayan, D., R. Patel, K. Schafft, A. Rademacher, and S. Koch-Schulte. 2000. *Voices of the Poor: Can Anyone Hear Us?* Washington, DC: World Bank.
- Quinn, C. 2007. "Using Copulas to Measure Association between Ordinal Measures of Health and Income." HEDG Working Paper 07/24. University of York, Health Economics Resource Centre, York, UK.
- Ravallion, M. 2011. "On Multidimensional Indices of Poverty." *Journal of Economic Inequality* 9 (2): 235–48.
- Sen, A.K. 1976. "Poverty: An Ordinal Approach to Measurement." *Econometrica* 44 (2):219–31.
- Tsui, K.-Y. 2002. "Multidimensional Poverty Indices." *Social Choice and Welfare* 19 (1): 69–93.
- UNDP. 2010. *Human Development Report 2010: The Real Wealth of Nations: Pathways to Human Development*. New York: Palgrave Macmillan.

---

# The Impact of Health Insurance Schemes for the Informal Sector in Low- and Middle-Income Countries: A Systematic Review

---

Arnab Acharya • Sukumar Vellakkal • Fiona Taylor • Edoardo Masset •  
Ambika Satija • Margaret Burke • Shah Ebrahim

---

*This paper summarizes the literature on the impact of state subsidized or social health insurance schemes that have been offered, mostly on a voluntary basis, to the informal sector in low- and middle-income countries. A substantial number of papers provide estimations of average treatment on the treated effect for insured persons. We summarize papers that correct for the problem of self-selection into insurance and papers that estimate the average intention to treat effect. Summarizing the literature was difficult because of the lack of (1) uniformity in the use of meaningful definitions of outcomes that indicate welfare improvements and (2) clarity in the consideration of selection issues. We find the uptake of insurance schemes, in many cases, to be less than expected. In general, we find no strong evidence of an impact on utilization, protection from financial risk, and health status. However, a few insurance schemes afford significant protection from high levels of out-of-pocket expenditures. In these cases, however, the impact on the poor is weaker. More information is needed to understand the reasons for low enrollment and to explain the limited impact of health insurance among the insured.*

JEL codes: I10, I15

The World Bank Research Observer

© The Author 2012. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

doi:10.1093/wbro/lks009

Advance Access publication November 9, 2012

28:236–266

## Introduction

In a seminal study, [Townsend \(1994\)](#) showed that in rural India, a health crisis in a household induced significant declines in the level of household consumption that were more severe than those associated with other financial crises. Townsend examined households' ability to "smooth consumption," to maintain a stable consumption level over a period of time. The inability to smooth consumption over time because of a health crisis has been found in other developing countries, partly as a result of the inability to afford appropriate and effective care to ensure the recovery of health and partly as a result of the reduced labor supply ([Gertler and Gruber 2002](#); [Deaton 1997](#)). Recent accounting methods indicate that for as many as 1.3 billion people in low- and middle-income countries, financial constraints are a major barrier to access to health care ([Xu et al. 2003](#); [Preker et al. 2004](#)).

Among the solutions proposed within various low- and middle-income countries to reduce costs to households at the point of care are the establishment or extension of national or social health insurance (SHI) in which service providers are paid from a designated government fund, which is partially funded through taxes. These SHIs are primarily intended for those employed outside of the formal sector. The [WHO \(2010\)](#) and the World Bank ([Hsiao and Shaw 2007](#)) have endorsed the restriction of out-of-pocket (OOP) expenditures for health care at the time of use through the prepayment of insurance as an important step toward averting the financial hardship associated with paying for health care. The impact of insurance, even when properly implemented, is not clear from the outset in low- and middle-income countries. Awareness of and trust in public programs, distance to health care facilities and institutional rigidities within the health care system can play major roles in limiting insurance enrollment and its related effects ([Wagstaff 2007](#); [Basinga et al. 2010](#)).

We conduct a systematic review of the literature on the extent to which insurance schemes enhance access to care and offer protection from financial risk to households in the informal sector. We also report on how these schemes may improve health. Previous reviews of the impact of health insurance in low- and middle-income countries for the poor include studies by [Ekman \(2004\)](#) and [Lagarde and Palmer \(2009\)](#). Both of these reviews focused on community-level insurance. The latter examined the impact of removing user fees. [Ekman \(2004\)](#) focused on community-based health insurance in low-income countries and concluded that community-based health insurance provides some financial protection by reducing OOP spending. Several countries have gone beyond community risk sharing, and many new insurance schemes have been introduced since 2004. In parallel, there is a growing interest in evaluating the impact of SHI programs in low- and middle-income countries ([Wagstaff 2009](#)). Thus, a new review of the

impact of health insurance for people in the informal sector could inform policy on the extent to which insurance provides greater access to necessary care, reduces the financial burden of care, and improves health.

The paper is organized as follows. In Section 2, we explain the review methodology of the current study. Section 3 describes methodologies that should be used to assess the impact of health insurance when it is offered on a voluntary basis. In Section 4, we describe the schemes found in the literature. Section 5 examines enrollment into health insurance schemes and its related impact. Section 6 concludes the paper.

## Methodology

The methodology aims to critically examine the evidence on whether health insurance schemes, once implemented, (1) are adopted, (2) provide greater access, (3) provide financial protection, and (4) improve health status among the intended beneficiaries. The last three points describe welfare impact. We summarize the methodology by defining the types of insurance that are of interest.

### *Types of Insurance*

If universal health care coverage is to be financed through prepayment for insurance, health insurance should have the following characteristics: (1) compulsory contributions to the risk pool (otherwise, the rich and healthy will opt out); (2) large numbers of people in the risk pool because pools with a small number of people cannot spread risk sufficiently and are too small to handle large health costs; and (3) pooled funds that are generally subsidized from government revenue (WHO 2010) due to the large numbers of poor people. The standard SHI schemes in developed countries mandate enrollment for people who are fully employed by imposing a mixture of tax on the employer and direct payment of insurance premiums. This effort is accompanied by requiring enrollment from people who are self-employed or unemployed, with varying degrees of revenue funding and cross-subsidies (OECD 2004).

SHI has also been mandated for formal sector workers in a number of developing countries. To achieve universal health care coverage, an institutional structure that emphasizes payment to providers for services delivered has been offered to those beyond the formal workforce in Vietnam, Nigeria, Tanzania, Ghana, India, and China over the last 15 years. Although schemes for the poor may share the same administrative structure as SHI for the formal sector, the former usually offer a reduced package or restrictions on providers. Alternatively, we find free-standing schemes (separate from SHI) that offer financial protection to the poor

through subsidized, usually voluntary household enrollment into a defined benefits arrangement (Wagstaff 2009). Insurance is offered at premiums that are considerably below the actuarially fair price. Given that most employment in low- and middle-income countries is informal, governments manage compulsory insurance in the formal sector with limited avenues to cross-subsidize the informal sector. Thus, health insurance is tax financed, although funding for it may be ring-fenced.

Starting from a point of low utilization by the formal sector, both free care and prepayment of financing for care are implemented to encourage use for illness or to increase contact with health workers to facilitate better delivery of preventive care. From the perspective of the user, prepayment insurance schemes, whether highly subsidized or zero-entry fee, should be understood differently than free care. First, even if the entry fee is small, some households may not be able to afford the fee. For example, Sparrow, Suryahadi, and Widayanti (2008) report that in Indonesia, it may be costly to provide photos of household members for insurance cards. Second, at the point of care, there may be copayments to limit frivolous care-seeking behavior because of the extremely low marginal cost of seeking care if zero copayments prevail. Reimbursement mechanisms may play a role; many families have little cash, are credit constrained (Pitt and Khandkar 1998) and cannot take on the financial burden at the point of care when reimbursement is often delayed (Shi et al. 2010).

From a governmental perspective, insurance allows a separation between the provision and the funding of care. Despite a public sector that offers care largely free of cost to recipients, the insurance system can take advantage of the pluralism in the supply of medical care that prevails in most low- and middle-income countries. The use of the private sector through government financing of health insurance would reduce the administrative burden within the allocation and subvention processes, the incidence of side payments and, perhaps, corruption.

When insurance is offered free, it often involves prescribed care and care givers. Some of these may be offered at the national level and some at the community level. Some community-level insurance with subsidized entry fees may have limited risk pooling because a specific community may be small. Most insurance that we examined only offered a set of well-defined interventions; thus, limited risk pooling at the community level may yield the same coverage as nationally sponsored insurance in terms of illness. Nationally sponsored insurance may allow for wider access to providers.

We examine studies of schemes that meet all of the following criteria:

- (1) Schemes that seek to offer financial protection for people facing health shocks to cover health care costs; these schemes involve some tax financing

- (2) Schemes that include a component in which poorer households can or must enroll through some formal mechanism at a rate much lower than the actuarial cost of the package or even free of charge; in return, these households receive a defined package of health care benefits
- (3) Schemes that are offered in one of two ways:
  - (a) Nationally managed and considered an extension of existing SHI
  - (b) Managed at the community level (limiting the risk-pooling population), either through a local government or nongovernmental organization (usually with government sponsorship), often called community-based health insurance

### *Defining the Impact of Insurance*

We presume that the impact evaluation of a project should provide two essential pieces of information to policy makers: (1) Is the program implementable? (2) Once implemented, does the program achieve a set of desirable outcomes? In the case of a health insurance program, if the adoption of a program is high, then the program *at least* approached proper implementation. Furthermore, policy makers are interested in the impact on those adopting the insurance, the average treatment on the treated, or the average impact on those who actually adopted the program. Insurance may affect those who do not adopt it by, for example, affecting the price of health care, which would be part of the total impact of the insurance program. [Morduch \(2006\)](#) reports that if richer individuals adopt a social program disproportionately at zero private cost, the program can be considered a large income transfer that may affect the prices of all goods within a relevant economy. The impact on prices is smaller when the poor adopt social goods because the adoption may not replace previous large expenditures. This type of general equilibrium impact has been largely ignored, probably because of the stable unit treatment value assumption where only those intended to be affected by the program become the subject of evaluation (see [Imbens and Wooldridge 2009](#)), which focuses on the intended target of a program.

The intention to treat effect can also be measured as the impact of insurance on individuals offered the insurance, regardless of whether it was adopted. An intention to treat measure would not approximate the impact of health insurance for those who adopt insurance if insurance uptake is not at a very high level, a common situation in most of the studies.

We report on three outcomes: (1) utilization of health care, (2) financial protection, and (3) health status. Willingness to pay for health insurance schemes, obtained *ex post* once the benefits have been realized, can be used to measure welfare impact. *Ex ante* willingness to pay for insurance is likely to be positive ([Gustafsson-Wright, Asfaw and van der Gaag 2009](#)). However, recipients of health services may

not always be able to accurately assess benefits. Furthermore, the severity of income constraints for the poor may not elicit well-considered responses.

### *Inclusion Criteria*

The studies on which we report must measure or report impact through a comparator, using either a contemporaneous control or a constructed control from data containing similar information collected over a similar time period. Inclusion criteria are as follows:

- (1) Randomized controlled trials
- (2) Quasi-randomized controlled trials in which methods of allocating are not random but create a matched control group through either:
  - (a) a propensity score matching method or
  - (b) a regression discontinuity design
- (3) Controlled before-and-after studies or difference-in-differences; the pre- and postintervention periods for study and control groups should be the same, and the choice of the control site should be similar in terms of socioeconomic characteristics and/or should have no major differences in the baseline
- (4) Regression studies that consider the probability of selection into treatment through the instrumental variable method
- (5) Qualitative studies focused on exploring the impact of health insurance and meeting a checklist

However, no qualitative studies that explored the impact of insurance were found.

### *Search Method*

A number of electronic databases<sup>1</sup> were searched using keywords related to health insurance, health care, and low- and middle-income countries. This search yielded 4756 references, including numerous duplicates and studies detailing general health issues in low- and middle-income countries. We filtered by titles and abstracts to reduce the number of relevant studies to 64. Of these, 35 were related to the impact of SHI on low- and middle-income countries. Inclusion criteria were met in 24 studies. Further examination found that five studies used poor identification strategies when measured against the standard methods recommended for impact evaluation studies (Imbens and Wooldridge 2009).

### *Summarizing the Results*

It is difficult—and, more important, misleading—to aggregate the outcome measures that we found into some form of meta-analysis. This difficulty arises for three

reasons: (1) many of the outcome measures are different (for example, the time intervals varied); (2) the insurance schemes were different, as outlined in table 2; and (3) the estimations of the impact depended on the functional form or the estimation method and the unit of measure, such as the period in which data were measured.

The unit of measure can shape the results, especially with regard to health expenditures (Das, Hammer, Sánchez-Paramo 2011). It can also dictate different estimation methods. For example, OOP expenditures as a share of income can be modeled through probit, whereas OOP expenditures may be modeled using linear estimation methods. In addition, when magnitudes are reported, they should be understood within the context of the study; magnitudes have limited generalisability outside of a study, even for the same insurance scheme within the same region. Thus, only trends are reported.

## Identification Issues

Although low enrollment fees should have attracted universal adoption, in most cases, enrollment rates were low. Low enrollment may induce selection effects, and selection into insurance may ultimately affect the outcome. One way that selection may affect outcome is through adverse selection: ill individuals select themselves into insurance at premium levels, which individuals in good health find the premium too costly given their expectation of their health care needs (Rothschild and Stiglitz 1976).<sup>2</sup> The pool of the insured may be sicker than the pool of the uninsured. The expectation of becoming healthy influences the adoption of insurance, which can be an efficient way to obtain care. In contrast to the possibility of adverse selection, given the low costs of entry for most insurance, it may simply be that better-informed individuals enroll. Better-informed individuals may also be more educated, may have larger incomes, and may be healthier than those who do not adopt insurance.

Thus, if one examines the average impact of insurance on those who adopt insurance (i.e., average treatment on the treated), then a simple comparison of insured and noninsured individuals does not yield appropriate results. The comparison is flawed because the noninsured group may not have had an opportunity to enroll, and this group includes those who would adopt insurance if offered as well as those who would not adopt insurance. The insured and noninsured groups may differ with regard to the factors that may affect outcomes. In light of the possibility of selection into insurance, the threat to validity is high. As a result, the impact may differ significantly when no adjustment is made for this type of selection, especially if average treatment on the treated is reported (Imbens and Wooldridge 2009).<sup>3</sup> Inclusion criteria have focused on studies with counterfactuals, we now discuss identification issues applied to the inclusion criteria. We briefly describe

statistical procedures to obtain average treatment on the treated in the present context. As indicated, 19 studies properly addressed identification issues.

### *Randomized Studies*

Even if insurance is offered through random means when uptake is low, it cannot be assumed that people who adopt insurance are similar to those who do not. An adjustment is needed even in this randomized setting. The most standard approach is to determine the local average treatment effect. If the stable unit treatment value assumption holds, as the instrument, mainly the assignment to treatment, is exogenous, then the local average treatment effect estimates the impact of those who comply with the offer but would not be treated otherwise (Angrist and Pischke, 2009). Thornton used the instrumental variable approach, although different approaches have also been used.<sup>4</sup>

### *Matching*

If insurance uptake occurs in a nonexperimental setting, a popular method known as propensity score matching can be used. Impact is measured by comparing the outcomes of insured individuals with the outcomes of nonparticipants. This measurement derives weights for the outcomes for nonparticipants according to the degree of similarities between the two groups as judged through observed factors, which are reduced to a single metric (Rosenbaum and Rubin 1983): the propensity to be enrolled in a program. Wagstaff et al. (2009) emphasize that unobservable heterogeneity may be stronger between those who adopt insurance and those who refuse insurance when both groups are offered insurance than between those who have never been offered insurance and those who adopt insurance when offered. In this case, the matching method should use comparators chosen from those who have never been offered insurance.

### *Instrumental Variable Approach*

A number of authors have used instrumental variable methods to determine that individuals who adopt insurance are not easily comparable to those without insurance. Insurance status is dependent on a variable that affects only entry into insurance, not any of the outcomes that may be affected by insurance. For studies that use the instrumental variable method, participating in insurance can be considered a problem of endogeneity or of selecting into insurance. Wagstaff and Lindelow (2008) and Sosa-Rubi, Galarraga and Harris (2009a) model selection into insurance as a problem of endogeneity; individuals anticipate the impact of insurance, and this expectation of the impact shapes the uptake of insurance.

These studies use instrumental variable methods to correct endogeneity. Instrumental variable methods can correct policy endogeneity (Dow and Schmeer 2003) by including policy with the expectation of certain types of results.

### *Regression Discontinuity Design*

When programs are targeted to a group at a measurable threshold income, the regression discontinuity design approach compares health care-related outcomes for those who are eligible at the margin with those who are just above eligibility. Individuals who do not qualify for enrollment in insurance because they are marginally on the other side of eligibility constitute the control group. The impact of regression discontinuity design yields an intention to treat estimation; some individuals who are eligible may not actually have insurance although they intended to receive it.

Finally, a study by Wagstaff (2010) subtracts two previous difference-in-differences outcome measures from two later difference-in-differences measures (using available data for three periods) and regresses this variable with similar differences in the independent variables and insurance status.

Several studies model the insurance effect through multiple observations of individuals in the sample and individual heterogeneity over time. This model is usually performed by inputting factors for a specific individual effect with the underlying assumption that any correlation between the error term and the insurance status arises from the correlation between time-variant unobservable factors (perhaps such as health) and insurance status. However, the time-invariance assumption is unlikely to hold because health conditions indeed fluctuate to influence insurance uptake.<sup>5</sup>

## Description of the Studies

Table 1 provides descriptions of the health insurance schemes from the 19 studies in addition to the corresponding data and methodologies. No study attempted to link the various outcomes of interest to any specific insurance features. However, it is instructive to note which types of schemes were evaluated. Reports of the impact of health insurance are from Burkina Faso (one study) Costa Rica (one study), Georgia (one study), Ghana (one study), India (one study) Nicaragua (one study), Colombia (two studies), Mexico (three studies), Vietnam (four studies), and China (four studies). Three studies, from Burkina Faso, China, and India, reported on community-based health insurance with government support.

Not all studies reported enrollment. Studies on impact evaluations obtained results through (1) a randomized trial (three studies); (2) propensity score matching (nine studies); (3) instrumental variable estimation, to consider either

**Table 1.** Description of the Studies

	Study country, and year of launch	(1) Name of the scheme (2) National or regional	Health insurance <sup>1</sup>		Data	Methodology	Funding source
			(1) Benefit package (2) Target beneficiaries (3) Premium (4) Cost-sharing arrangements (5) Enrollment rate				
1	Aggarwal (2010); Karnataka, India; 2003	(1) The Yeshasvini Health Insurance Programme (2) Regional (state level) scheme	(1) Covers surgical procedures of high-cost, low-probability, highly catastrophic medical events and free out-patient department care (2) Rural farmers of cooperative societies and informal sector workers (3) Premium INR 120 (USD 2.4) per person (4) No copayment (5) 3.0 million in 2008–09		Survey of 4109 households; block matching to the insured; cross-sectional; sample includes households that have never been offered insurance	Propensity score matching; clear selection equation and balancing results; not clear if the presence of zero expenditure is taken into account	Bill and Melinda Gates Foundation through the Global Development Network
2	Bauhoff et al. (2011); Georgia; 2006	(1) The Medical Insurance Programme for the Poor (2) National	(1) Most emergency outpatient care and set of planned and emergency inpatient care (2) Poor, 20 percent of Georgian population (3) Fully funded through the general government budget (4) No copayments (5) Low enrollment		Survey of 3500 households	Intention to treat estimation regression discontinuity design where enrollment is not very high; estimations for OOP expenditures is presented through generalized linear model with log link to account for those who undertook zero expenditure	Georgia Health and Social Project Implementation Center
3	Dow and Schmeer (2003); Costa Rica; 1970s	(1) National health insurance (2) National	(1) Primary and secondary health care (2) Lower socioeconomic groups (3) 73 percent of children by 1984 (4) <b>Not reported</b> (5) <b>Not reported</b>		Vital statistics registries, Census data; a panel for 88 to 97 regional data	Fixed-effect model of health outcomes on a region for infant mortality rates using Cox binary transformation for infant mortality rates as the dependent variable for the region	National Institute of Child Health and Human Development Grants

Continued

Table 1. Continued

	Study country, and year of launch	(1) Name of the scheme (2) National or regional	<u>Health insurance<sup>1</sup></u>		Data	Methodology	Funding source
			(1) Benefit package (2) Target beneficiaries (3) Premium (4) Cost-sharing arrangements (5) Enrollment rate				
4	Gnawali et al. (2009); Senegal; 2004 to 2006	(1) Community-based health insurance (2) Regional	(1) Consultation, essential and generic drugs, laboratory tests, limited inpatient hospital stays (2) People in the informal sector, including the poor (3) XAF 1500 per adult and XAF 500 per child per annum in a household (USD 1 = XAF 655) (4) No copayment (5) 5.2 percent in 2004, 6.3 percent in 2005, and 5.2 percent in 2006	Cluster randomized with 33 clusters, involving 4936 households; step-wedge design; complete information is found for 1309 households; not clear where the uninsured came from	Propensity score matching model to account for very low uptake; no balancing table, but selection model is present	The German Research Foundation	
5	Mensah et al. (2010); Ghana; 2003	(1) National Health Insurance Scheme (2) National	(1) Specified package of general out-patient services, in-patient services, oral health, eye care (2) General population, including people in the informal sector (3) Sliding scale: free for core poor, between USD 8 and USD 52 (4) <b>Not reported</b> (5) 55 percent of the total national population by August 2007	Survey by researcher; 393 insured women and 1689 uninsured women, randomized samples from regions; control group matched in the area; reports on only 565 women who were pregnant; small sample	Used propensity score matching; presents clear selection equation and covariate balance tables; results from different matching methods that are generally consistent	Global Development Network and the Bill and Melinda Gates Foundation	

6	Thornton and Field (2010); Nicaragua; 2007	(1) Nicaraguan Social Security Institute's health insurance programme (Seguro Facultativo de Salud); (2) National (study experiment on increasing enrollment)	(1) Preventive, diagnostic, maternity, and curative health services (2) Informal sector (3) USD 15 per month (4) No copayment at the time of service (5) 20 percent of the sample (in the experiment)	Insurance offered at randomly selected market booth; uses a pre-experiment baseline, 2610 households; loss to follow up 7 percent	Local average treatment effect; selection into the insurance is modeled as an ordinary least squares with the insurance offer as the instrumental variable; the outcome measure is differences-in-difference; results for those who enrolled given the enrollment procedure; no accounting for zero OOP expenditure	The United States Agency for International Development's Private Sector Partnerships-One project and the Global Development Network
7	Trujillo et al. (2005); Colombia; 1993	(1) Subsidized health insurance program; (2) National	(1) Basic health care services (2) Low-income families (3) Government funded (4) A coinsurance rate that varies between 5 percent and 30 percent according to the individual's income (5) <b>Not reported</b>	1997 Colombia Living Standards Survey; 5559 insured through social health insurance system and 16,732 uninsured; may not have been insurance	Propensity score matching, selection equation presented (no balancing results) and instrumental variable estimation are compared	The University of Central Florida
8	Miller et al. (2009); Colombia; 1993	(1) Régimen Subsidiado (Subsidised Regime), National Health Insurance (2) National	(1) Primary care, inpatient care (2) Poor (3) Fully funded through the general government budget (4) Low level of coinsurance (5) <b>Not reported</b>	Colombian household surveys (the Encuestas de Calidad de Vida and the Demographic and Health Surveys); nearly 4300 families eligible and marginally ineligible	Intention to treat estimation of constructed eligibility from a survey; uses regression discontinuity design; the analysis also uses an instrumental variable of constructed value for eligibility on actual enrollment; no accounting for zero expenditure	The Economic and Social Research Council; UK Inter-American Development Bank; National Institute of Child Health and Human Development and the Stanford Center on Demography and Economics of Health and Aging
9	King et al. (2009); Mexico; 2005	(1) Mexican <i>Seguro Popular de Salud</i> (Universal Health Insurance program-SP) (2) National (study experiment on increasing enrollment)	(1) A package to treat the diseases responsible for approximately 95 percent of the burden (2) People in the informal sector (3) Fully government funded (sliding scale by income, free for the poor) (4) <b>Not reported</b> (5) Approximately 3.5 million families	Negotiated 74 paired clusters to participate with one from a pair randomly assigned to intense insurance uptake campaign before national enrollment; survey in 50 pairs; 32,515 households	Presents intention to treat estimations and the effect on experimental compliers of average causal effect; the outcome measured is differences-in-difference; the complier results should be understood as specific to the study; no accounting for OOP expenditure of zero	National Institute of Public Health of Mexico; the Mexican Ministry of Health; the National Institutes of Aging and the National Science Foundation; United States

Continued

Table 1. Continued

Study country, and year of launch	(1) Name of the scheme (2) National or regional	Health insurance <sup>1</sup>			Data	Methodology	Funding source
		(1) Benefit package	(2) Target beneficiaries	(3) Premium			
10 Sosa-Rubi et al. (2009a) Mexico; 2005	See King et al. (2009)				National Health and Nutrition Survey 2006; Sample of 3890 women who delivered between 2001 and 2006; no one with employer insurance or private insurance was included; complete data for all women	Innovative multinomial probit that takes into account endogenous enrollment through instrumenting by time of the introduction of SP, assuming there is no problem of policy endogeneity.	The Health Ministry of Mexico (SSA)
11 Sosa-Rubi et al. (2009b) Mexico; 2005	See King et al. (2009)				National Health and Nutrition Survey 2006; 1491 adults with diabetes were chosen; no pregnant women or women with access to social security services; complete data for all adults (see Sosa-Rubi, 2009a)	Propensity score matching for those with and without SP insurance; presentation of bias reduction and selection equation	The Health Ministry of Mexico (SSA)
12 Wagstaff (2007); Vietnam; 2003	(1) Health care fund for the poor (2) National	(1) Covered inpatient and outpatient care only at public providers until 2005; some preventive care (2) All poor households and selected other groups (3) Fully subsidized (4) No copayment (5) As of 2006, the program covered approximately 60 percent of those eligible, accounting for approximately 23 percent of the population			Vietnam Household Living Standards Survey 2004	Propensity score matching method on single-period data; selection equation and balancing is well presented; not clear whether those in the control group were offered insurance; results dependent on propensity score matching weights chosen; no accounting for zero expenditure	Not reported

13	Axelsson et al. (2009); Vietnam; 2003	(1) Health care fund for the poor (2) National	(1–4) See Wagstaff (2007) (5) Study reported 18 percent enrollment for 2002	Vietnam Household Living Standards Survey data 2002 (preprogram) and 2004 (postprogram); cross-sectional 10,232 and 4112 panel	Uses both single-period and differences-in-difference outcome measures; uses survey data as in Wagstaff (2007); selection equation and balancing results are presented from propensity score matching; no accounting for zero expenditure	Not reported
14	Wagstaff (2010); Vietnam; 2003	The same as Wagstaff (2007)	(1–5) See Wagstaff (2007)	The panel 2002, 2004, and 2006 Vietnam Household Living Standards Survey Multipurpose household survey; 1689 households in all three waves	Triple difference is the outcome variable; the triple difference is regressed on covariates; no accounting for zero expenditure	Not reported
15	Jowett et al. (2004); Vietnam; 1992, discontinued, not currently in place	(1) Voluntary health insurance program (2) National	(1) <b>Not reported</b> (2) Self-employed individuals, farmers, schoolchildren (3) Fully subsidized (4) A copayment of 20 percent with exceptions (5) 9.7 percent of target group	Data were collected through a household survey designed specifically to evaluate the impact of the scheme; analysis from 2631 households	Two-stage multinomial logit model to examine the type of facility used; instrumental variable for selection into insurance is used; appropriateness of instrumental variable tested; no theory given for unusual instrumental variable	United Kingdom's Department for International Development
16	Wagstaff and Lindelow (2008); China; 1996 onward	(1) Multiple health insurance schemes: Labour Insurance Scheme and Government Insurance Scheme (2) National	(1) <b>Not reported</b> (2) General population (3) Not free (4) <b>Not reported</b> (5) 90 percent of the population covered in 1970, but decreased to 20 percent for rural population and 40 percent for urban population from 1980 onward; increased to 90 percent of urban workers by 2003	Three surveys: (1) China Health and Nutrition Survey in 1991, 1993, 1997, and 2000; (2) Gansu Survey of Children and Families in 2000 and 2003; (3) World Bank Health VIII project baseline survey in 1998; total sample was 18,200 adults	Instrumental variable is used to take account of selection with probit for catastrophic measure and then panel data are used; fixed effect is only used for logit with no instrumental variable because these are OOP expenditures, a <b>generalized linear</b> model with instrumental variable is used to consider zero expenditure by some	The Spencer Foundation Small and Major Grants Programme; Fogarty International Center at the National Institutes for Health; World Bank Research Committee

Continued

Table 1. Continued

Study country, and year of launch	(1) Name of the scheme (2) National or regional	<u>Health insurance<sup>1</sup></u>			Data	Methodology	Funding source
		(1) Benefit package (2) Target beneficiaries (3) Premium (4) Cost-sharing arrangements (5) Enrollment rate					
17 Wagstaff et al. (2009); China; 2003	(1) NCMS (2) National	(1) Heterogeneity in the benefit package across counties and coverage modes; all counties cover inpatient care, some cover outpatient (2) Rural population (3) The minimum premium requirement was a CNY 10 (per person) beneficiary contribution from households, supplemented by government subsidy (4) Deductibles, ceilings, and coinsurance rates (5) <b>Not reported</b>			Two data sets: (1) The 2003 round of the National Health Service Survey of the Ministry of Health; follow up in 2005; (2) Routine Health Facility Survey from the Ministry of Health administrative data; total households > 8000	Use propensity score matching to match the insured with those who have never been insured; show balancing results but no selection equation; subgroup analyses are presented by regressing individual treatment effect (weighted through propensity score) on income groups; most likely estimation of cost is for those receiving medical care.	The World Bank and the United Kingdom's Department for International Development
18 Lei and Lin (2009); China; 2003	(1) NCMS (2) National	(1–4) See Wagstaff et al. (2009) (5) For NMS, 85.7 percent of the rural population were covered in 2008			Longitudinal sample drawn from the China Health and Nutrition Survey for 2000, 2004, and 2006; different analyses use different panel and thus have different data sizes; differences-in-difference is only for a panel of 3225 individuals	Differences-in-difference using propensity score matching along with instrumental variable estimations and fixed-effect panel are used on panel data; balancing results are presented with no selection equations; not clear if OOP expenditures includes zero expenditures	Not reported
19 Wang et al. (2009); China; 2003 to 2006	(1) Rural Mutual Health Care in China; a social experiment (2) Regional and community-based in rural area (China's western provinces)	(1) Both outpatient services and hospital services (2) Villagers, including farmers (3) Annual premium of at least CNY 10 (4) No copayment (5) 1173 households			The Rural Mutual Health Care experiment adopted a pre-post treatment-control study design of those not offered insurance; panel of 1665 insured and 1745 uninsured individuals	Propensity score matching models with varied matching and subgroup analyses are presented, as is balancing after matching	Not reported

<sup>1</sup>No entry indicates not reported in the study.

endogeneity at the individual level or regional program placement (four studies); (4) the use of a regression discontinuity design on eligibility to obtain intention to treat (two studies); and (5) double difference-in-differences from three periods with regression (one study).

The data used in these studies ranged from program-designated data sets to routinely collected available data at the national level gathered to measure a range of indicators of wellbeing.

## Findings

We first report enrollment, and then, we report intention to treat or average treatment on the treated estimations of whether insurance is likely to have resulted in welfare improvement. Table 2 summarizes the outcomes.

### *Enrollment and Its Determinants*

The enrollment rate partially reflects whether a health insurance program can be implemented. Our review did not conduct a systematic search to identify studies that report enrollment. For three papers (Gnawali et al. 2009; King et al. 2009; Thornton and Field 2010), the evaluation was conducted for programs that were designed to enhance enrollment. The activities did not enhance enrollment.

Enrollment rates varied. For the Vietnam Health Care Fund for the Poor (VHCFP), introduced in 2003, which includes free enrollment and no copayment with specified access to care, country-wide enrollment reached 60 percent by 2006 (Wagstaff 2010). The New Cooperative Medical Scheme (NCMS) in China showed regional variations of 48 to 99 percent (Wagstaff 2009). By 2007 (i.e., within the first four years), national enrollment in Ghana was at 55 percent (Mensah, Oppong and Schmidt 2010). Bauhoff, Hotchkiss and Smith (2010) report low enrollment in Georgia in a collected sample.

The enrollment patterns and determinants of enrollment in health insurance schemes are similar to those observed for enrollment into insurance schemes to provide protection from adverse shocks in general (Gine and Yang 2007). We summarize the factors affecting enrollment from studies that reported determinants:<sup>6</sup>

- (1) No clear demographic patterns emerge; in some cases, positive enrollment factors include female-headed households and elderly headed households, family size, and composition
- (2) Positive effect of education (except in Colombia; Miller, Pinto and Vera-Hernandez, 2009)

**Table 2.** Summary of Findings<sup>†,‡</sup>

<i>Utilization</i>	<i>Financial Protection</i>	<i>Health Status</i>
<b>1. Aggarwal (2010) India (Yeshasvini Community-based Health Insurance)</b>		
Those with health insurance decided to use health facility in greater numbers and with greater frequency; increase from outpatient service usage, including outpatient surgery; no higher usage in frequency of hospitalization; less usage of government services	Overall, medical expenses were actually higher for the insured, with the poor experiencing no change; for hospitalization, expenditures are significantly lower for the insured; also reported is the incidence of burrowing for hospital care, which is smaller for the insured	<b>n.a.</b>
<b>2. Bauhoff et al. (2011) Georgia (Targeted Scheme for the Poor)</b>		
No impact on utilization from intention to treat estimations	No robust evidence of lower expenditures among insured outpatients' expenditures, except for the elderly; lower expenditure among insured for inpatient care	<b>n.a.</b>
<b>3. Dow and Schmeer (2003) Cost Rica (National Insurance Expansion)</b>		
<b>n.a.</b>	<b>n.a.</b>	No impact on decline in community infant mortality rates from increased proportion of population insured over time
<b>4. Gnawali et al. (2009) Burkina Faso (Community-based Health Insurance)</b>		
Overall, there is a significant positive impact on health care utilization; more outpatient visits, but no significant impact on inpatient care utilization; the higher outpatient utilization is only significant among the richest group	<b>n.a.</b>	<b>n.a.</b>
<b>5. Mensah et al. (2010) Ghana (National Health Insurance Scheme)</b>		
The insured women who are enrolled are more likely to give birth in hospitals and to receive higher levels of prenatal care, preventive health check ups, and attention from trained health professionals	<b>n.a.</b>	Three types of health status are reported, two of which (infant death and birth complications) are significant under specific matching weights; the difference in infant death is likely to suffer from small sample size

**6. Thornton and Field (2010) Nicaragua**

Local average treatment effect measure; no significant effect on overall health care utilization; fairly substantial substitution away from use of public and private facilities to health care facilities covered by insurance; social security hospitals

Local average treatment effect measure; overall decline in OOP expenditures decreased by a smaller amount than the actual premium for the insured; no significant result for OOP spending for the insured reported; the sample is too small to note effect in catastrophic spending

**7. Miller et al. (2009) Colombia (Targeted Scheme for the Poor)**

Intention to treat estimations; substantial higher use of traditionally underutilized preventive services for those with health insurance

Intention to treat estimations; no significant effect on average outpatient expenditures; lowers inpatient expenditures and lowers incidence of high-end expenditures among the insured

**n.a.****8. Trujillo et al. (2005) Colombia (Targeted Scheme for the Poor)**

Greatly increased medical care utilization among the country's poor, including children, women, and the elderly

**n.a.****n.a.****9. King et al. (2009) Mexico (SP)**

No effect in utilization

Intention to treat and complier effects for low-asset holders show lower OOP expenditures for the SP-insured with low assets for overall, inpatient, and outpatient care; female-headed households had lower inpatient OOP expenditures; all insured had lower inpatient and outpatient care, but no significant effect was found for overall care (including drug costs)

**n.a.***Continued*

**Table 2.** *Continued*

<i>Utilization</i>	<i>Financial Protection</i>	<i>Health Status</i>
<b>10. Sosa-Rubi et al. (2009a) Mexico (SP)</b> <b>n.a.</b>	Different types of facility utilizations are reported; these imply different costs; those with SP prefer SP facilities and the cheapest care over private care and non-SP government hospital service care, which costs more than SP care but is cheaper than private care; private care is preferred to non-SP government care	<b>n.a.</b>
<b>11. Sosa-Rubi et al. (2009b) Mexico (SP)</b> Those with SP had better access to diabetes care; they had higher rates of insulin shots, regular tests, and physician visits	<b>n.a.</b>	Higher proportion of the insured with glucose control, and lower proportion with very poor glucose control
<b>12. Wagstaff (2007) Vietnam (VHCFP)</b> Increase in both outpatient and inpatient utilization but substantially increased inpatient care utilization; impact on utilization among the poor is even less noticeable	Results robust to different matching techniques and samples; there is no effect on overall OOP expenditures; however, there is a lowered risk of high or catastrophic OOP expenditures; even with this protection among the insured, one-third still faced catastrophic expenditures; the poor may have received more risk protection from high expenditures	<b>n.a.</b>
<b>13. Axelson et al. (2009) Vietnam (VHCFP)</b> Small but positive impact on overall health care utilization; the insured do not have greater difference in utilization of inpatient care; statistically significant effect is present only for outpatient visits in community hospital	From the cross-sectional study, the OOP expenditures were higher for the insured; differences-in-difference measure of two periods showed a larger reduction in expenditures for the insured for inpatient care and a reduction in catastrophic expenditures of 20 percent; the opposite result was found for outpatients	<b>n.a.</b>

**14. Wagstaff (2010) Vietnam (VHCFP)**

No impact on the use of either outpatient or inpatient health care services

Triple-differencing estimates yield insurance effects with a sizable reduction of expenditures; poorer groups also experienced a significant reduction

**n.a.****15. Jowett et al. (2004) Vietnam (health insurance not presently in place)**

Overall, insured patients are more likely to use outpatient facilities, public providers, and inpatient services

**n.a.****n.a.****16. Wagstaff and Lindelow (2008) China (Basic Medical Insurance Programme)**

The results suggest that the insured may use health services more frequently; analysis suggests that insurance facilitates the use of higher-level services

The results vary for different data sets and specifications; the general picture that emerges is that insurance results in a lower OOP expenditures and is likely to increase the probability of incurring catastrophic expenditures at different threshold levels

**n.a.****17. Wagstaff et al. (2009) China (NCMS)**

In the analysis by regions, the scheme increased outpatient and inpatient utilization; households with insurance have more doctor visits and inpatient spells; results differ by health centers; the richest quintile responded more favorably

The results vary; for delivery services, all regions showed lower costs for the insured; for the overall OOP expenditures, a mixed result emerges; total OOP expenditures increased in most regions, except two, mirroring the case for inpatient visits

**n.a.***Continued*

**Table 2.** *Continued*

<i>Utilization</i>	<i>Financial Protection</i>	<i>Health Status</i>
<p><b>18. Lei and Lin (2009) China (NCMS)</b>            No significant evidence on increase in the utilization of formal medical service; however, utilization significantly decreases the use of traditional Chinese folk doctors and increases the utilization of preventive care, particularly general physical examinations</p>	<p>Authors detect no impact on expenditures from any of the estimations presented in the work</p>	<p>Among the estimation presented, the propensity score matching shows marginally higher health for the insured</p>
<p><b>19. Wang et al. (2009) China (Rural Mutual Health Care, Community-based Health Insurance)</b>  <b>n.a.</b></p>	<p><b>n.a.</b></p>	<p>EQ-5D dimensions and specific dimensions were reported for the entire population with the Differences-in-Difference measure; decrease in illness is significantly higher among the insured for all dimensions and for pain/discomfort and anxiety and depression from two types of propensity score matching</p>
<p><sup>†</sup>All measures are the average treatment on the treated unless specified.  <sup>‡</sup>Empty cells indicate no information for the category.</p>		

- (3) No influence of initial conditions, such as chronic illness (except in Colombia; [Miller et al. 2009](#))
- (4) No influence of distance to health centers or rural residency (except in Mexico, where people in rural areas sign up more frequently; [Sosa-Rubi, Galarraga and Harris, 2009](#))

The initial health condition did not matter. A detailed study by [Gine, Townsend and Vickery \(2008\)](#) on the uptake of agricultural insurance against bad weather showed that adverse selection played a small role in uptake. None of the studies we examined explicitly included a variable for trust in government or financial institutions, levels of risk aversion, availability of care, or understanding of insurance. Not all studies reported enrollment, even though the issue of who enroll influences the outcome of a social program.

### *Utilization*

The studies report whether the use of overall or specific types of health care was higher for insured people than for uninsured people within a specific time interval. Studies conducted across multiple time periods compare changes across two groups. To measure any incidence of utilization, studies use logit or probit; to measure the impact of insurance on the number of incidences per household or person, count data models can be used. Most studies reported on both inpatient and outpatient care. Choice of facility, which has cost implications, was also reported through a multinomial model.

Membership in health insurance schemes may lead to overuse of health care as a result of two types of moral hazards: overuse because the cost of any given point of contact with the health care system for the insured is low or nearly zero and overuse because insurance involves a third-party payer, which can encourage greater health care utilization. Thus, the utilization rate may not reflect actual welfare gains. There was no estimation of unnecessary care in any of the studies. Where there was a financial barrier to care, increased care gained through insurance is likely to indicate unambiguous welfare improvement.

In the case of Ghana, [Mensah, Opong and Schmidt \(2010\)](#) report a higher utilization rate for pregnancy care among the insured, although the sample is small. For Nicaragua ([Thornton and Field 2010](#)) and Georgia ([Bauhoff et al. 2010](#)), insurance targeted mostly to the poor did not induce higher utilization, although the study in Georgia reported higher utilization by those with higher assets. In Burkina Faso ([Gnawali et al. 2009](#)) and India ([Aggarwal 2010](#)), the two community-managed schemes, there were overall increases in health care use, but there was no impact on inpatient utilization.

Trujillo, Portillo and Vernon (2005) and Miller, Pinto and Vera-Hernandez (2009) indicate positive effects for the same insurance program in Colombia at different times. The insured received care more often, and the latter study reported a higher use of preventive care after changes to the payment structure for the provider. Both studies report no difference in inpatient care for insured and noninsured groups.

For Mexico's *Seguro Popular* (SP), studies report differing results: King et al. (2009) report no higher utilization for the insured for all health care, whereas Sosa-Rubi, Galarraga and Lopez-Riduaura (2009) report that diabetics insured under SP have better access to diabetic care compared to the corresponding figures for diabetics who are uninsured.

Three studies of VHCFP from Vietnam report conflicting results. Wagstaff (2007) reported higher utilization rates for inpatient and outpatient care, with substantially higher inpatient care. Axelson et al. (2009) reported a small increase in overall utilization, mostly because of increased outpatient care. Although both papers use propensity score matching, they use different data. A subsequent study by Wagstaff (2010), which used a different data and methodology, found no effect of insurance on utilization. The results from both papers by Wagstaff are not robust to functional specifications. For insurance prior to VHCFP, Jowett, Deolalikar and Mattinsson (2004) use instrumental variables on the decision to seek care and the type of health center used. They report that the insured are more likely to use health services and public services than the uninsured.

Contradictory results emerge from two studies on China's NCMS. Wagstaff et al. (2009) show that in China, the insured, including the insured poor, use health services more often in comparison to the noninsured. Lei and Lin (2009) show no overall effect for utilization but find a drop in the use of traditional care and an increase in preventive care.

We cannot claim that insurance yields a higher probability of care seeking. It is particularly telling that different results can be obtained for the same insurance. Of the 15 studies reporting utilization, nine studies report a higher utilization rate among the insured. Recall that increased usage may not always indicate welfare improvement.

### *Financial Protection*

Insurance should protect the insured from incurring high levels of health care costs. An effective health care system that includes insurance and other forms of social protection should provide much broader financial risk protection. None of the insurance schemes offered protection for financial loss due to reduced labor supply, which is among the main reasons for the lack of consumption smoothing, as noted by Gertler and Gruber (2002). Miller et al. (2009) mention, in passing,

the important issue of whether health insurance can go beyond reducing health care costs to eliminate significant adverse effects of health shocks so that households can maintain their standard consumption and saving bundles (Townsend 1994; Chetty and Looney 2006). Most studies addressed only the issue of OOP expenditures and do not include insurance premiums or entry fees into insurance. Some studies used the measure of catastrophic payment, defined as a threshold proportion of all expenditures (or some type of income measure, which is usually imprecise), that is spent on health. The denominator varied across studies, as did the threshold levels. Because a reduction in the average level of OOP expenditures for a household would reflect a reduction in high-level OOP expenditures, many studies reported on this value. One way of describing the impact of insurance on financial risk protection is to examine the right tail of the distribution of OOP expenditures. Distributional analysis may require the use of quantile regression methods that help to analyze the occurrence of high levels of expenditures at different income levels.

One indicator of improved wellbeing is found by measuring the reprieve from high levels of OOP expenditures by the poor in comparison to populations without insurance. This indicator was not clearly identified in any of the studies. Low levels of increased spending may actually indicate greater contact with health services, which may occur through insurance. Comparisons between insured and noninsured groups at the average level of OOP expenditures may not yield a clear measure of welfare.

Nonetheless, the studies compared the average expenditure between the insured and the noninsured at the household level as well as the incidence or probability of incurring high or catastrophic expenditures, measured at different thresholds. Some studies on the determinants of expenditures for hospital care noted the large fraction of zeroes because many people do not use hospital care. Although a two-part model can be used by first considering the likelihood of the use of health services, this model was not incorporated in most of the studies reported here. It was difficult to discern whether some studies reported costs only for those who adopted health care, which may be an observed indicator for being ill. One disadvantage of such an approach is that among the poor, some people may not use health care at all, even when they are ill.

For Georgia, Bauhoff et al. (2010) report lower levels of OOP expenditures for the insured, with a larger impact for inpatient care. Thornton and Field (2010) use baseline data to show that insurance does not provide cost savings in Nicaragua when the cost of insurance is taken into account. Aggarwal's (2010) study of community insurance in India shows a favorable impact for overall care among the insured but found no effect for inpatient care. Miller et al. (2009) find overall lower OOP expenditures and a lower incidence of high-level expenditures for the insured.

Sosa-Rubi et al. (2009a) report that pregnant women with Mexican SP insurance use SP-sponsored state services, the care with the lowest OOP expenditure. Nevertheless, the evidence is unclear because there is a preference among the insured for expensive private care over cheaper types of state-sponsored care. The urban poor seem to have benefitted the most from SP. King et al. (2009) confirm that for all types of care, OOP expenditures are lower for the insured under SP insurance.

For Vietna's VHCFP, Wagstaff (2007) shows no overall impact on OOP expenditures for the insured. However, there is some protection for high levels of expenditures, with the poor experiencing a small effect. The results are susceptible to the matching methods used. Axelson (2009) uses data from two periods and finds a protective effect of insurance; however, there is no impact when a single cross-sectional data period is examined. Wagstaff (2010) uses data from three periods and finds strong and robust measures of a greater decline in OOP expenditures for the insured.

Lei and Lin (2009) do not find a significantly lower level of OOP expenditures for people insured under China's NCMS. Wagstaff et al. (2009) note weak evidence for lower OOP expenditures for the insured under the NCMS; however, this evidence is sensitive to matching methods. For deliveries, the insured received protection, although this protection was weaker for the poorer population. Wagstaff and Lindelow (2008) use a number of econometric specifications through instrumental variable analyses and report that people insured in a (now discontinued) Chinese health insurance scheme actually experienced higher levels of catastrophic payments, measured at various threshold levels.

Only four of 16 studies reporting on costs provided conclusive indications of lower average OOP expenditures for the insured. Seven studies provided mixed results, and two showed no effect. Five studies reported a lower incidence of catastrophic OOP expenditures.

### *Health Status*

Surprisingly, only six studies reported on health measures. It is presumed that health insurance would induce greater access for the insured and thereby lead to better health. With the exception of the study on the health insurance scheme from Colombia (Miller et al. 2009), no study reported that supply-side improvement accompanied the introduction of insurance. If health insurance implementation is not accompanied by a significant improvement in the quality of supply and does not lead to greater utilization, then we should not expect health improvement. Financial protection is the main aim of insurance. However, if a range of health outcomes improves or death rates decline for the insured, then it is possible that we can attribute better health outcomes to health insurance.

Mensah et al. (2010) show lower levels of infant death, although these levels are not statistically significant. Wang et al. (2009) use EQ-5D, a standardized index value instrument for use as a measure of a wide range of health conditions, to report on a community-based health insurance program in China. They find that the scheme had positive effects on health status for all insured people and for the poor. Measuring regional changes, Dow and Schmeer (2003) find no correlation in changes in infant mortality as regional insurance uptake improves. For Nicaragua, Thornton and Field (2010) show no improvement in health. Sosa-Rubi et al. (2009b) examine Mexico's SP insurance and show improved glucose control among diabetics with insurance than those without insurance. Lei and Lin (2009) find no improvement in health status for the China's NCMS.

## Discussion and Conclusion

We now summarize our conclusions and note some methodological issues. We offer very little in terms of broad results regarding the impact of insurance, once implemented, on the intended beneficiaries.

Studies reporting on enrollment showed that low enrollment is commonly observed for many of the insurance schemes; enrollment seems to be related to perceptions, education, and cultural factors rather than to factors related to health and health care, such as initial health status and distance to health centers. The study from Nicaragua indicated that there was considerable confusion about coverage. We do not observe a pattern regarding enrollment and outcome; for example, China and Vietnam had high enrollment. Nevertheless, there is no indication that insurance worked well for the participants, although more recent analysis shows positive results from Vietnam. Given the low coverage, policies could include incentives for insurance or could even mandate required enrollment. Enrollment may also be low because the administrative implementation process may be poor.

It is perhaps most important to prevent high levels of OOP expenditures through insurance. There is some evidence that this may be the case. Some of the studies that report only average expenditures could not capture this effect. The present method of setting catastrophic expenditures at various levels of income is arbitrary and complicates comparisons among studies. One option is to examine the expenditure distributions of the insured and the noninsured, particularly at high levels of expenditures. However, this technique is of limited use if there are high levels of selection into insurance. The possibility of quantile regression methods can be explored (Angrist and Pischke 2009).

Counterintuitively, for most of the health insurance schemes, the poorest among the insured fared less well. One reason that average expenditures may not

be lower for the insured poor than for the uninsured poor is that the latter may not seek any care or may give up on care altogether when the appropriate care is well beyond reach without insurance. However, if this were the predominant reason why the poor do not receive the full effect of insurance, then insurance should induce higher levels of health-seeking behavior. We do not find this to be the case. One reason for the low level of health-seeking behavior may be a lack of understanding of insurance or the existence of hidden charges other than those covered by insurance.

Two studies explicitly note that, although a causal link was not established, features of the payment scheme may have affected the outcome. In the older insurance scheme in China, fee-for-service may have given rise to cost (Wagstaff and Lindelow 2008). In the case of Colombia, the incentive structures in the providers' contract may have given rise to higher use of preventive care (Miller et al. 2009). Studies should identify health system and household economic factors that may determine impact. For example, mechanisms for copayments, expectations from reimbursement policies, and the presence of various financial mechanisms have been shown to produce variations in uptake, utilization, and health improvement in the U.S. market (Newhouse and the Insurance Experiment Group 1993; Deb, Trivedi, and Zimmer 2006). In the low- and middle-income country study settings, the full range of variations may not exist as it does in the U.S. market. Qualitative studies may shed some light in this regard. Health insurance schemes differed sufficiently from one another in this review to avoid implying any relationship between the specificities of health insurance schemes and outcomes.

In this review, we found that many studies used data collected for purposes other than the evaluation of insurance schemes. Thus, important questions may be missing from general living standard surveys to allow the assessment of the welfare implications of insurance, such as detailed questions on illnesses. Longitudinal surveys would be more robust in capturing selection effects and the extent to which health insurance schemes provide risk protection against health shocks.

Both the development of rigorous impact evaluation methodology for social programs (Imbens and Wooldridge 2009) and the introduction of health insurance schemes for the poor in low- and middle-income countries are new phenomena. Our report of 19 studies may be encouraging. However, for impact studies to be useful for future considerations of health insurance schemes, greater attention must be given to the rigor and uniformity of welfare measurements, especially in terms of risk protection and evaluation methodologies. In the future, examining a larger number of studies would allow for meta-analyses (regressions), which would facilitate more conclusive remarks regarding program features and

outcomes. Further, given that health insurance schemes can differ and exist in different contexts, studies should emphasize pathways through which programs affect outcomes. Although we did not identify such studies, both qualitative and quantitative methods can be used to trace these pathways. Policy makers would benefit from a greater number of rigorous studies that examine the pathways through which programs are likely to affect welfare.

## Notes

Arnab Acharya is with the Department of Public Health Policy, London School of Hygiene & Tropical Medicine. Sukumar Vellakkal and Ambika Satija are with the South Asia Network for Chronic Disease, Public Health Foundation of India. Fiona Taylor and Shah Ebrahim are with the Department of Epidemiology & Population Health, London School of Hygiene & Tropical Medicine. Edoardo Masset is with the Institute of Development Studies. Margaret Burke is with the Department of Social Medicine, University of Bristol. The corresponding author is Arnab Acharya; email address: [Arnab.Acharya@lshtm.ac.uk](mailto:Arnab.Acharya@lshtm.ac.uk). We would like to thank Rachel Blackman, Billy Stewart, Julia Watson, and Robert Yates from Department for International Development, United Kingdom for their comments. We would also like to thank Sarah Alkenbrack, Kara Hanson, Jorge Hombrados, Anne Mills, Chris Millet, Timothy Powell-Jackson, Ajay Tandon, Richard Smith, Hugh Waddington, Howard White, and Winnie Yip as well as seminar participants at the London International Development Centre, London School of Hygiene and Tropical Medicine at a session at the 8th World Congress on Health Economics, Toronto and from the South Asia Network for Chronic Disease. We greatly appreciate Wendy Wisbaum's editorial assistance. All errors remain the sole responsibilities of the authors. This work was supported by the UK Department for International Development-DFID [PO 40031461].

1. Databases from the 1950s to September 2010 were searched: the Cochrane EPOC group Specialised Register and Library (3, 2010), MEDLINE, EMBASE, ECONLIT, ISI Web of Knowledge, CAB Abstracts, CENTRAL, DARE, ELDIS, and IDEAS as well as websites from the World Bank, the World Health Organization, and the U.S. National Bureau of Economic Research. Expert opinions and searches in key journals yielded findings of additional studies published before July 2011. 2. Adverse selection is one of the reasons mandatory insurance is prescribed in many instances. Voluntary enrollment increases the possibility of adverse selection, which is one of the reasons that enrollment fees must be low and the cost of the program must be subsidized.

3. [Yip and Berman's \(2001\)](#) study was among the early empirical papers on health insurance for developing countries that recognized the selection problem. They addressed the issue through simulation.

4. Studies with random allocation at the cluster level using informal matching methods cannot be justified; see [Devadasan et al. \(2010\)](#).

5. Two of the reviewed studies attempted this approach. [Sepehri, Sarma and Simposon \(2006\)](#) attempt to control selection through the use of fixed-effect or random-effect models for individuals in a panel. Similarly, [Sparrow, Suryahadi, and Widyanti \(2010\)](#) model insurance impact with baseline self-reported health status, which is nearly akin to a fixed-effect model. They report that self-reported illness is likely to be unreflective of actual illness status.

6. Reporting the determinants of insurance involved straightforward identification issues. We include enrollment results from a few studies for which we did not include impact results: [Sun et al. \(2009\)](#), [Dror et al. \(2009\)](#), [Schneider and Diop \(2001\)](#), and [Msuya \(2004\)](#). No studies had high enrollment at the time reported by the study.

## References

- Aggarwal, A. 2010. "Impact evaluation of India's 'Yeshasvini' community based health insurance programme." *Health Economics* 19 (S1): 5–35.
- Angrist, J., A. Imbens, and D.B. Rubin. 1996. "Identification of causal effects using instrumental variable." *Journal of American Statistical Association* 94 (436): 444–72.
- Angrist, J., and J-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Axelsson, H., S. Bales, P. Minh, B. Ekman, and U. Gerdtham. 2009. "Health financing for the poor produces promising short-term effects on utilization and out-of-pocket expenditure: evidence from Vietnam." *International Journal for Equity in Health* 8 (1): 20.
- Basinga, P., P. Gertler, A.L.B. Soucat, J.R. Sturdy, and C.M.J. Vermeersch. 2010. *Paying primary health care centers for performance in Rwanda*. The World Bank: WPS5190.
- Bauhoff, S., D.R. Hotchkiss, and O. Smith. 2011. "The impact of medical insurance for the poor in Georgia: pA regression discontinuity approach." *Health Economics* doi:10.1002/hec.1673.
- Chetty, R., and A. Looney. 2006. "Consumption smoothing and the welfare consequences of social insurance in developing economies." *Journal of Public Economics* 90 (12): 2351–6.
- Das, J., J. Hammer, and C. Sánchez-Paramo. 2011. *The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior*. The World Bank: WPS5778.
- Deaton, A. 1997. *The Analysis of Household Surveys*. Baltimore. Johns Hopkins University Press.
- Deb, P., C. Li, P. Trivedi, and D.M. Zimmer. 2006. "The effect of managed care on use of health care services: Results from two contemporaneous household surveys." *Health Economics*, 15:743–60.
- Devadasan, N., B. Criel, V. Van Damme, S. Manoharan, P.S. Sarma, and P. Van der Stuyft. 2010. "Community health insurance in Gudallur, India, increases access to hospital care." *Health Policy and Planning* 25 (2):145–54.
- Dow, W., and K. Schmeer. 2003. "Health insurance and child mortality in Costa Rica." *Social Science and Medicine* 94 (2): 975–8.
- Dror, D.M., R. Radermacher, S.B. Khadilkar, P. Schout, EX. Hay, A. Singh, and R. Koren. 2009. "Micro insurance: pinnovations in low-cost health insurance." *Health Affairs (Millwood)* 28 (6): 1788–98.
- Ekman, B. 2004. "Community based health insurance in low-income countries: pa systematic review of the evidence." *Health Policy and Planning* 19 (5): 249–70.
- Gertler, P., and J. Gruber. 2002. "Insuring Consumption against Illness." *American Economic Review* 92 (1): 51–70.
- Gine, X., and D. Yang. 2007 *Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi*. The World Bank: Policy Research Working Paper 4425.
- Gine, X., R. Townsend, and J. Vickery. 2008. "Patterns of Rainfall Insurance Participation in Rural India." *The World Bank Economic Review*: 1–28.
- Gnawali, D.P., S. Pokhrel, A. Sié, M. Sanon, M. De Allegri, A. Souares, H. Dong, and R. Sauerborn. 2009. "The effect of community-based health insurance on the utilization of modern health care services: Evidence from Burkina Faso." *Health Policy* 90 (2-3): 214–22.
- Gustafsson-Wright, E., A. Asfaw, and J. van der Gaag. 2009. "Willingness to pay for health insurance: pAn analysis of the potential market for new low-cost health insurance products in Namibia." *Social Science and Medicine* 69:1351–9.
- Hsiao, W., and R.P. Shaw. 2007. *Social Health Insurance for Developing Nations*. Washington DC: The World Bank

- Imbens, G., and J. Wooldridge. 2009. "Recent development in the econometrics of program evaluation." *Journal of Economic Literature* 47 (1): 5–86.
- Jowett, M., A. Deolalikar, and P. Mattinsson. 2004. "Health insurance and treatment seeking behaviour: pevidence from a low-income country." *Health Economics* 13 (9): 845–57.
- King, G., E. Gakidou, K. Imai, K. Lakin, T. Moore, C. Nall, N. Ravishankar, M. Vargas, M. Téllez-Rojo, J. Hernández Ávila, M. Hernández Ávila, and H. Hernández Llamas. 2009. "Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme." *The Lancet* 373 (9673): 1447–54.
- Lagarde, M., and N. Palmer. 2009. "The impact of health financing strategies on access in low and middle income countries." *The Cochrane Collaboration*. Protocol Issue 4, Wiley Publishers [www.thecochranelibrary.com](http://www.thecochranelibrary.com)
- Lei, X., and W. Lin. 2009. "The New Cooperative Medical Scheme in rural China: pdoes more coverage mean more service and better health?" *Health Economics* 18 (S2): S25–S46.
- Mensah, J., J. Oppong, and C. Schmidt. 2010. "An evaluation of the Ghana national health insurance scheme in the context of the health MDGs." *Health Economics* 19 (S1): 95–106.
- Miller, G., D.M. Pinto, and M. Vera-Hernandez. 2009. *High-powered incentives in developing country health insurance: pevidence from Colombia's regimen subsidiado*: NBER Working Paper, 15456.
- Morduch, J. 2006. "Micro-insurance: The Next Revolution?" In Abhijit Banerjee, Roland Benabou, and Dilip Mookherjee, eds., *Understanding Poverty*, New York: Oxford University Press.
- Msuya, J., J.P. Jutting, and A. Esfaw, 2004. "Impacts of community health insurance schemes on health care provision in rural Tanzania." Bonn: *ZEF-Discussion Papers on Development Policy* (82).
- Newhouse, J., and the Insurance Experiment Group. 1993. *Free for All? Lessons from the RAND Health Insurance Experiment*, Cambridge, MA: Harvard University Press.
- OECD. 2004. OECD Study on Private Health Insurance. *OECD Health Project*. Organisation for Economic Cooperation and Development.
- Pitt, M.M., and S.R. Khandker. 1998. "The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter?" *Journal of Political Economy* 106 (5): 958–96.
- Preker, A., G. Carrin, D. Dror, M. Jakab, W. Hsiao, and D. Arhin. 2004. "Rich-Poor Differences in Health Care Financing." In *Health Financing for Poor Pople: Resource Mobilisation and Risk-Sharing*, pp. 3–51, eds. A.S. Preker, and G. Carrin. Washington DC: The World Bank.
- Rosenbaum, P., and D. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70 (1): 41–55.
- Rothschild, M., and J. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90 (4): 630–49.
- Schneider, P., and E. Diop. 2001. *Synopsis of results on the impact of community-based health Insurance on financial accessibility to health care in Rwanda*. HNP Discussion Paper. The World Bank.
- Sepehri, A., S. Sarma, and W. Simposon. 2006. "Does non-profit health insurance reduce financial burden? Evidence from the Vietnam living standards survey panel." *Health Economics* 15 (6): 603–616.
- Shi, W., V. Chongsuvivatwong, A. Geater, J. Zhang, H. Zhang, and D. Brombal. 2010. "The influence of the rural health security schemes on health utilization and household impoverishment in rural China: Data from a household survey of China." *Int J Equity Health* 9:7: doi:10.1186/1475-9276-9-7.
- Sosa-Rubi, S.G., O. Galarraga, and J. Harris. 2009a. *Heterogeneous impact of the 'Seguro Popular' program on the utilization of obstetrical services in Mexico, 2001-2006: A multinomial probit model with a discrete endogenous variable*. Cambridge, MA: National Bureau of Economic Research.

- Sosa-Rubi, S.G., O. Galarraga, and R. Lopez-Riduaura. 2009b. "Diabetes treatment and control: he effect of public health insurance for the poor in Mexico." *Bulletin of World Health Organisation* 87 (7): 512–9.
- Sparrow, R., A. Suryahadi, and W. Widayanti. 2008. *Public Health Insurance for the Poor: Targeting and Impact of Indonesia's Asheskin Programme*. The Hague, Netherlands: Institute of Social Studies.
- Sun, Q., X. Liu, Q. Meng, S. Tang, B. Yu, and R. Tolhurst. 2009. "Evaluating the financial protection of patients with chronic disease by health insurance in rural China." *International Journal of Equity in Health* 9:42.
- Thornton, R., L. Hatt, E. Field, M. Islam, E. Solís, and M.A. González Moncada. 2010. "Social Security Health Insurance for the Informal Sector in Nicaragua: pA Randomised Evaluation." *Health Economics* 19 (S1): 181–206.
- Townsend, R. 1994. "Risk and Insurance in Village India." *Econometrica* 62 (3): 539–91.
- Trujillo, A., J. Portillo, and J. Vernon. 2005. "The impact of subsidized health insurance for the poor: pEvaluating the Colombian experience using propensity score matching." *Journal of Health Care Finance and Economics* 5 (3): 211–39.
- Wagstaff, A. 2009. "Social health insurance re-examined." *Health Economics* 19 (5): 503–17.
- . 2007. *Health insurance for the poor: initial impacts of Vietnam's health care fund for the poor*. The World Bank, Policy Research Working Paper Series: 4134 2007.
- . 2010. "Estimating health insurance under unobserved heterogeneity: pThe case of Vietnams Health Care Fund for the poor." *Health Economics* 19 (2): 189–208.
- Wagstaff, A., and M. Lindelow. 2008. "Can insurance increase financial risk? The curious case of health insurance in China." *Journal of Health Economics* 27 (4): 990–1005.
- Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng. 2009. "Extending health insurance to the rural population: pan impact evaluation of China's new cooperative medical scheme." *Journal of Health Economics* 28 (1): 1–19.
- Wang, H., W. Yip, L. Zhang, and W.C. Hsiao. 2009. "The impact of rural mutual health care on health status: pevaluation of a social experiment in rural China." *Health Economics* 18:S65–S82. doi:10.1002/hec.1465.
- WHO. 2010. *The world health report - Health systems financing: the path to universal coverage*. Geneva, Switzerland: World Bank.
- Xu, K., D. Evans, K. Kawabata, R. Zeramdini, J. Klavus, and C. Murray. 2003. "Household catastrophic health expenditure: pa multi-country analysis." *The Lancet* 362 (9378): 111–7.
- Yip, W., and P. Berman. 2001. "Targeted health insurance in a low income country and its impact on access and equity in access: Egypt's school health insurance." *Health Economics* 10 (3): 207–20.

---

# Rural Households in a Changing Climate

---

Javier E. Baez • Dorothy Kronick • Andrew D. Mason\*

---

*This paper argues that climate change poses two distinct, if related, sets of challenges for poor rural households: challenges related to the increasing frequency and severity of weather shocks and challenges related to long-term shifts in temperature, rainfall patterns, water availability, and other environmental factors. Within this framework, we examine evidence from existing empirical literature to compose an initial picture of household-level strategies for adapting to climate change in rural settings. We find that although households possess numerous strategies for managing climate shocks and shifts, their adaptive capacity is insufficient for the task of maintaining—let alone improving—household welfare. We describe the role of public policy in fortifying the ability of rural households to adapt to a changing climate. JEL codes: Q12, Q54, O13*

## Introduction

How will rural households in developing countries adapt to climate change? Rooted as it is in the hard sciences and in high-income communities, the public dialogue on global environmental change has only recently begun to consider the adaptive capacity of poor rural households. This inattention is not for lack of importance; the welfare of these households concerns millions. Rather, it may be for lack of definition: the issue is so broad and so complex that it seems to defy focused consideration.

One way to conceptualize adaptation in rural communities is to recognize two effectively distinct phenomena associated with climate change: first, increased weather variability and increased frequency and severity of extreme weather events (broadly, “shocks”) and, second, gradual changes in temperature and rainfall patterns (broadly, “shifts”). From a climate perspective, these phenomena are not actually distinct; from the perspective of adaptation among agricultural households, however, dividing the issue into shocks and shifts provides a useful

framework. Events that temporarily disrupt a household's modal environmental conditions can be considered shocks, and those that permanently change the modal conditions can be considered shifts.

Of course, the increasing frequency and severity of extreme weather events is itself a shift because it permanently changes modal environmental conditions. Likewise, gradual changes in temperature and rainfall patterns may manifest as shocks (in the form of droughts and floods, for example). By categorizing the ways in which rural households experience climate change, we clarify the relevance of several established bodies of literature. This literature includes work on ambiguity risk, agricultural technology adoption, reaction to weather-related information, determinants of migration, income diversification, health, and past experiences with permanent environmental changes. This paper synthesizes evidence from this literature to develop a more inclusive and composite picture of household-level strategies for adapting to climate change in rural settings.<sup>1</sup> We also briefly discuss the role of public policy in fortifying households' capacity to adapt.

We describe rural households' numerous strategies for dealing with climate change, and we outline the ways in which these strategies are insufficient for the task of maintaining (let alone improving) household welfare. The threat to welfare derives not only from the direct losses associated with climate shocks and shifts but also from increased uncertainty, which complicates households' attempts to smooth consumption (with credit or insurance, for example), to smooth income (by adjusting agricultural portfolios), and to evaluate options such as migration.

The notion that some households may lack sufficient adaptive capacity or may respond in ways that have adverse long-term consequences suggests a possible role for public policy. To date, the dominant policy framework, to the extent that one exists, has focused on what has been called a no-regrets approach to climate-related adaptation policy (see [Heltberg, Siegel, and Jorgensen 2009](#); [de la Torre, Fajnzylber, and Nash 2009](#); and [World Bank 2010](#)). In the face of the many uncertainties associated with the future effects of climate change, the no-regrets approach emphasizes policies and investments today that can generate positive net benefits "under all future scenarios of climate change and impacts" (Heltberg, Siegel, and Jorgensen 2009). This approach has many merits, though there may also be room for more specific action. Here, we draw attention to country experiences with (i) better systems of information for adaptation and (ii) climate-adapted risk management instruments (such as indexed production insurance), and (iii) social safety nets that are targetable and scalable on the basis of weather events.

The remainder of this paper is organized as follows. Section II defines key concepts related to rural households' exposure, sensitivity, and capacity to adapt to climate change. Drawing on extensive empirical literature on household responses to weather shocks, section III analyzes how—and how well—households may be

expected to cope with the increased frequency and severity of weather shocks associated with climate change. Drawing on several strands of empirical literature on the determinants of household economic mobility, section IV explores how households are likely to adapt to longer-term climatic shifts. Section V discusses directions for public policy, and section VI concludes.

## Exposure, Sensitivity, and Household Capacity to Adapt to Climate Change

In the literature on climate change, the concept of vulnerability comprises three elements: exposure (experience of climate conditions), sensitivity (response of the physical environment to exposure, such as the response of crop yield to temperature and precipitation), and adaptive capacity (discussed below) (IPCC 2007 and many others, notably Turner 2003 and Schröter, Polsky, and Patt 2005). This section briefly describes rural households' exposure and sensitivity to climate change and then defines adaptive capacity, which is the subject of the remainder of the paper.

Substantial effort has been devoted to generating increasingly precise projections of climate trends and the exposure of rural households to these trends. Summarizing this work is beyond the scope of this paper; below, we outline a few points that are useful for our discussion of adaptive capacity. As noted above, we organize our discussion according to two broad categories of climate changes: shocks, the increase in weather variability and the frequency of extreme weather events, and shifts, the long-term rise in global surface temperatures and related gradual changes in sea level and rainfall patterns.

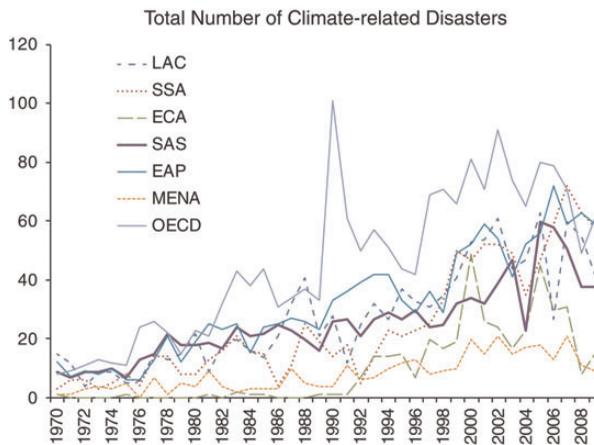
Over the past century, global mean surface temperatures rose by  $0.74^{\circ}\text{C}$ , on average (the absolute global mean surface temperature is approximately  $14^{\circ}\text{C}$ ) (IPCC 2007). The speed of warming accelerated as the century progressed. The expected trajectory of warming over the next hundred years is a matter of considerable debate. Existing models predict that global mean surface temperatures will rise by an additional 1 to  $5^{\circ}\text{C}$  by the year 2050.

Global warming is one determinant of precipitation patterns and extreme weather events, such as storms. Higher surface temperatures induce increased evaporation from soils and bodies of water. The resulting elevated concentration of water vapor leads to more precipitation in some places and more droughts in others. Some precipitation that used to fall as snow now falls as rain. In the view of many researchers, extreme weather events are more frequent and more severe as a result of global warming; there are now more extreme warm nights, heat waves, tropical storms, and hurricanes than there were 50 years ago. Figure 1

illustrates a clear upward trend in the number of weather-related natural disasters. The direction of the trend is clear, though the increase may be driven in part by both better reporting practices and increasing human settlement in risk areas. In regions such as Latin America and the Caribbean as well as East Asia and the Pacific, the average annual frequency of disasters increased fivefold between the 1970s and the 2000s and increased almost twofold in the most recent two decades (during which changes in measurement techniques were less likely to confound interpretation and inference).

Agricultural production is often highly sensitive to the weather shocks and shifts associated with climate change. Higher temperatures and changing rainfall patterns may influence crop yields; in recent years, for example, climatic variability related to El Niño Southern Oscillation (ENSO) affected wheat yields in Mexico, cotton and mango growing cycles in Peru, and the incidence of plant diseases in various other parts of Latin America. Cattle and dairy productivity may also suffer: heat waves have been shown to negatively influence milk production in Argentina. Water supply restrictions also reduce productive potential in agriculture. Ocean acidification may damage fish stocks, and sea-level rise may eventually submerge coastal farmlands. Extreme weather events often destroy crops and arable land.

**Figure 1.** Climate-related Disasters by Region (1970–2009)



*Note:* The count includes events that meet at least one of the following criteria: (1) 10 or more people reported as dead, (2) 100 people reported as affected, (3) a declaration of a state of emergency, or (4) a call for international assistance. LAC (Latin America and the Caribbean), SSA (Sub-Saharan Africa), ECA (East and Central Europe), SAS (South Asia), EAP (East Asia and Pacific), MENA (Middle East and North Africa).

*Source:* World Bank staff calculations based on EM-DAT: The OFDA/CRED International Disaster Database, Catholic University of Louvain

This list is not exhaustive, of course. The point is merely that climate change alters the natural conditions in which agricultural households operate. Although employment in agriculture has fallen over the past half century as developing economies have grown, hundreds of millions of people remain in the sector. For nonagricultural households in rural areas, the environment is sensitive to climate change: roads flood, infrastructure is damaged or destroyed, and diseases flourish or fade.

Perhaps because the term “adaptation” seems relatively straightforward, climate change authorities such as the Intergovernmental Panel on Climate Change and the United Nations Framework Convention on Climate Change Secretariat have not taken it upon themselves to sanction a common definition of the term in the context of climate change (OECD 2006). However, some clarity of definition may be useful for understanding what adaptation means in the context of our discussion. For the purposes of this paper, then, adaptation refers to household-level actions taken (i) in response to observed or expected climate conditions and (ii) with the objective of maintaining or improving household welfare.

Adaptation, in this context, involves a multistage process among household decision makers: first, signal detection, or the identification of a weather or climate event; second, an evaluation of the expected consequences or impacts of that event; third, a response, an observable change in household behavior that stems from the evaluation; finally, feedback, or monitoring the outcomes of the adaptive behavior (Kandlikar and Ribsey 2000). Planting alternative crops in response to changing temperature, moving to a different location to avoid floods or droughts, and insuring against weather-related asset losses could all be considered adaptive actions within this process. Accordingly, we use the term adaptive capacity to denote the extent to which households are able to undertake the processes and actions required to maintain or improve welfare in the face of negative shocks and shifts associated with climate change.

Like the manifestations of climate change, the determinants of adaptive capacity among rural households are numerous. For the purposes of this article, these determinants can be thought to fall into two broad categories: (i) economic and physical factors and (ii) behavioral or psychological factors. The former category includes factors traditionally discussed in development literature, such as people’s levels of human capital, their access to credit, ownership and control of assets (e.g., land, livestock), and physical mobility. The latter category involves individuals’ perceptions of risk and of their own adaptive capacity, which have been the focus of more recent work (e.g., Bryan, Chowdhury, and Mobarak 2011). One view of adaptation to climate change holds that these perceptions are just as important—perhaps more important—than physical and economic factors in determining adaptive capacity (Grothmann and Patt 2005).<sup>2</sup>

## Household Responses to Climate-change Shocks: Protecting Consumption, Income, and Human Capital

Rural households have been exposed to weather shocks for millennia, and scholars have studied households' responses to these shocks for decades. In many ways, household adaptation to weather shocks associated with climate change is likely to mirror past adaptation to storms, floods, and droughts. In this sense, the empirical evidence on household responses to past weather shocks is instructive. At the same time, the increasing frequency and severity of weather shocks associated with climate change, if detected and evaluated, may produce qualitative changes in household responses. Understanding these potential differences is key to contextualizing existing evidence.

First, the increasing frequency and severity of weather shocks will almost certainly alter the economic and physical factors that determine households' adaptive capacity. Beyond direct damage to physical capital, more frequent and extreme weather events may induce higher levels of precautionary savings, which, in turn, may result in lower investment and a reduction in the long-term income stream. Similarly, farmers may reduce overall productivity in an attempt to mitigate risk by diversifying agricultural production portfolios in a suboptimal manner. To our knowledge, there is little empirical evidence on the cumulative effects of repeated weather shocks.

We might consider the damage to physical assets and the necessity of additional saving or risk-management activities to be quantitative or magnitude-based changes in adaptive capacity. The key concept is that climate change will simply scale up the adverse effects of weather shocks as we currently observe them.

Perhaps more significant than these changes in scale, however, are the ways in which increasing frequency and severity of weather shocks may imply qualitative changes in the way that households respond to climate, principally by affecting the behavioral and psychological determinants of adaptive capacity. In particular, trends in climate shocks may reduce households' ability to detect and appropriately evaluate weather-related risk *ex ante*, thereby widening the gap between optimal and actual adaptive behaviors. Recent theoretical and empirical work suggests that ambiguity aversion may negatively affect investments in insurance, for example (Bryan 2010). In Zimbabwe, farmers have been found to systematically underestimate the risk of drought and to underestimate their own capacity to adapt to it (Pat and Gwata 2002). This evidence is suggestive, although little is known empirically about the way in which changes in the frequency or severity of climate shocks affect the appropriateness of adaptive responses.

Despite these potential differences between past and future adaptive behavior and capacity, the existing literature on household responses to weather-related

shocks can provide valuable insights into future adaptive behaviors. The central message that emerges from previous work is that although households employ a range of measures to manage risk and shocks, the (often informal) adaptation mechanisms that they employ are generally insufficient for maintaining their welfare in the face of weather shocks. Moreover, some coping mechanisms are economically inefficient and may have adverse long-term consequences. As discussed above, we expect that the increasing frequency and severity of climate shocks will only magnify these insufficiencies.

### *Consumption Smoothing*

Weather shocks (as compared with household-specific shocks such as illness) pose particular challenges for households intent on smoothing consumption. Because they affect whole communities rather than individuals, they complicate local insurance or reciprocal gift-giving efforts, requiring risk pooling or intervention across a broader set of geographic regions.

To what extent will weather shocks associated with climate change reduce consumption among rural households? A large body of evidence indicates that although rural households' efforts to smooth consumption in the face of weather shocks can significantly mitigate negative effects, the success of these efforts varies dramatically across both communities and households. In a survey of papers based on longitudinal data from southern rural India, for example, Morduch (2003) finds that the effect of transfers between households (i.e., gifts given with the expectation of reciprocity, an active consumption-smoothing strategy) ranges from a 40 percent reduction in total income risk to a 90 percent reduction in total income risk.

Similarly, a study on Thailand finds that variation in rainfall patterns pushed rural households to save more in their efforts to smooth consumption (Paxson 1992). However, a recent study of 162 villages in Ethiopia finds that although some households were able to protect consumption by selling livestock or borrowing from friends and relatives, between 25 and 40 percent of respondents "did nothing" in response to droughts, floods, and hailstorms (Deressa, Hassan, and Ringler 2009).

Existing evidence also suggests that increased weather volatility associated with climate change is likely to have the largest effects on the consumption of poor rural households, which have the fewest assets, human resources, and social networks. A study from rural China finds, for example, that a 10 percent decline in income is associated with a 4 percent decline in consumption among poor households but only a 1 percent decline among nonhouseholds (Jalan and Ravallion 1999). Likewise, evidence suggests that households with greater asset holdings and members who are more educated are better able to protect themselves from the impacts of shocks and to recover afterward. In rural Nicaragua and El

Salvador, for example, recent studies show that households with higher levels of education and greater asset holdings are both less likely to fall into poverty during an aggregate shock and more successful at growing their incomes in the aftermath (Vakis, Kruger, and Mason 2006; Beneke de Sanelú and Shi 2004).

In sum, available evidence indicates that households' abilities to smooth consumption in the face of weather-related (and other) shocks are "real and significant but not complete" (Murdoch 1995). The increasing weather volatility associated with climate change is likely to further stretch the capacities of rural agricultural households to protect their consumption and increase the risks of their use of inefficient or costly mechanisms to respond to shocks. The evidence also indicates that the increased frequency and severity of weather shocks is likely to have the greatest impact among the poorest rural households.

### *Income Smoothing*

In light of their inability to fully smooth consumption in the face of income shocks, rural households dealing with decreasingly predictable returns to agriculture may turn to income-smoothing strategies, such as altering crop portfolios or diversifying sources of household income. In doing so, households often choose to reduce risk even at the cost of significant foregone returns.

One strategy adopted by households in the absence of formal mechanisms for insuring against production risk is to pursue low-risk, low-return agricultural portfolios. In a rural municipality in southern Peru, for example, farmers plant and cultivate many small, geographically dispersed and less profitable plots to reduce aggregate production variance (Goland 1993). Likewise, farmers in southern India whose production is vulnerable to rainfall variability are more likely to plant low-risk, traditional varieties of rice and castor rather than higher-risk, high-yield (more profitable) varieties (Morduch 1990). High levels of weather variability may also discourage farmers from adopting production-enhancing technologies.

There are significant costs associated with such forms of adaptation; conservative factor allocations may reduce profits and, thus, farmer incomes. In the Peruvian municipality described above, the field-scattering strategy was found to reduce farmers' net yields by about 7 percent (Goland 1993). In the case of southern rural India, microsimulation exercises showed that a one-standard deviation reduction in weather risk (rainfall variability) would increase the average profits for farmers at the bottom of the income distribution by nearly 35 percent (Binswanger and Rosenzweig 1993).

Another way that households adapt to increasingly unpredictable agroclimatic conditions is to diversify their sources of income. Indian households that are vulnerable to transitory shocks, including those induced by weather risk, are more likely than similar households to have a member employed in stable,

nonagricultural sectors, even when off-farm employment offers lower wages (Rosenzweig and Stark 1989). Studies from Guatemala, Thailand, and China, among other places, describe similar behavior (see, e.g., World Bank 2005a; Townsend 1995; Jalan and Ravallion 1999).

Increasing weather variability and shocks could therefore lead to movements away from agricultural labor markets. Although diversifying out of agriculture is often associated with income growth and poverty reduction (World Bank 2007), moving into nonagricultural employment in response to weather-related shocks is not without its risks. A recent study using panel data from Bangladesh analyzes the evolution of wages in the five years following a massive flood in 1998 (Mueller 2010). Although agricultural labor markets there experienced short-term negative consequences due to the flood, nonagricultural labor markets performed worse than agricultural labor markets in the medium term.

### *Human Capital Accumulation*

Human capital accumulation is also highly susceptible to unfavorable weather events. By reducing investment in education and health and by changing the incidence of certain diseases, weather shocks may jeopardize long-term household welfare, increasing susceptibility to future shocks and increasing the likelihood of the intergenerational transmission of poverty.

Households in rural India, for example, pull their children out of school to cope with various weather-related income shocks (Jacoby and Skoufias 1997). In Nicaragua, Hurricane Mitch negatively affected the progress of students in school (Ureta 2005). There is some evidence that such transitory coping mechanisms may solidify into permanent losses in human capital investment. Evidence from Mexico shows that children who are withdrawn from school during difficult times are one-third less likely to continue their studies than similar children whose schooling is not interrupted (de Janvry, Finan, Sadoulet, and Vakis 2006). This trend is particularly troubling in light of the fact that education—and female education, in particular—appears to dramatically increase households' adaptive capacity (see, for example, Blankespoor, Dasgupta, Laplante, and Wheeler 2010).

These weather-related declines in schooling in rural areas are often associated with increases in households' use of child labor. In Tanzania, for example, there is evidence that transitory income shocks created by accidental crop loss increased child labor, more so among poor than nonpoor households (Beegle, Dehejia, and Gatti 2003). In El Salvador, the fraction of children working outside the household following earthquakes in 2001 more than doubled, from 6.5 to 16.5 percent (Santos 2007). Child labor also increased in rural Nicaraguan households affected by the coffee crisis and in rural Nicaraguan households directly in the path of Hurricane Mitch (Vakis, Kruger and Mason 2006; Baez and Santos 2007).

Weather-related shocks can also have a detrimental impact on the nutritional status of poor rural women and children. Evidence from Ethiopia suggests that households do not smooth consumption across individuals within the household; rather, the nutritional status of women “bears the brunt of adverse shocks” (Dercon and Krishnan 2000). Similarly, Baez and Santos (2007) find that Hurricane Mitch increased the likelihood of early infant malnutrition by an alarming factor of four.

As in the case of education, there is evidence that weather-induced shocks to nutrition can have adverse long-term consequences on household welfare. In Zimbabwe, for example, a recent study of young children who became stunted as a result of a drought finds that those affected never fully recovered. Specifically, the study found lasting negative effects not only on affected children’s body weights but also on their school attainment and subsequent earnings (Alderman, Hoddinott, and Kinsey 2006).

Climate change may also affect health more directly by changing the incidence of diseases. Malaria and dengue, for example, which are mosquito-borne diseases, flourish in warm, humid conditions, depending on stagnant water (where mosquitoes breed) and wet air (in which adult mosquitoes can best survive). Cholera and diarrheal diseases are also water-borne. Several studies have shown that the prevalence of these diseases increased in the wake of heavy rainfall caused by El Niño in various years in Ecuador, Bolivia, Peru, Colombia, and Venezuela (Vos, Velasco, and Labastida 1999, WHO 1999, Bouma et al. 1997, Bouma and Dye 1997). Extremes of heat and cold are also linked to higher morbidity and mortality, mostly among the elderly (Rumel et al. 1993, Shumway, Azari, and Pawitan 1988, Bull and Morton 1978).

## Household Adaptation to Long-term Shifts in Temperature and Rainfall Patterns

Thus far, this paper has focused on how—and how well—rural households may be expected to respond to shocks associated with climate change: tropical storms, droughts, floods, and other extreme weather events. We now turn to gradual shifts in temperature and rainfall patterns. Discrete weather-related shocks and gradual shifts in temperature and precipitation levels associated with climate change are, of course, intimately related. The effective distinction that we draw here is that shocks represent deviations from a local mean, whereas shifts represent long-term changes in the mean.

For rural households, these two phenomena require different types of adaptive behaviors. Shocks require efforts to smooth consumption and income, whereas shifts require longer-term—and perhaps permanent—changes in households’

methods of income generation. Long-term climate shifts also complicate the task of insuring against weather shocks because traditional insurance mechanisms are designed to address variations in the mean of some variable of interest (e.g., agricultural productivity or income) as opposed to secular changes in the means of those variables over time.

The existing evidence on farmers' perceptions of climate change indicates that farmers often recognize long-term shifts in environmental conditions and that they attempt to respond to these shifts. In a recent survey conducted by IFPRI in Sub-Saharan Africa, more than half of the respondents stated that they perceived long-term increases in temperature and decreases in precipitation (Nhemachena and Hassan 2007). They also indicated that a lack of credit and a lack of information were significant barriers to adaptation.

There are a number of ways in which households might try to adapt to gradual, long-term changes in climactic conditions. Climate shifts may induce households to (i) adapt their use of agricultural inputs and technologies; (ii) diversify their sources of household income, either within agriculture or toward non-agricultural enterprises; or (iii) emigrate from affected regions, all in an attempt to compensate for the adverse effects of climate shifts on agricultural productivity, farm profitability, and household income.

Directly measuring how—and how effectively—households adapt to long-term climate shifts is difficult because doing so requires long panels linking household production behavior, productivity, and income to climate variables, data that largely do not exist at present. Most panel evidence captures household economic mobility and its determinants over relatively short periods of time, on the order of five years.<sup>3</sup> One recent study of a village in the Philippines uses a unique 30-year panel of household data to examine the factors that enable household economic mobility (Fuwa 2007). The study finds that in addition to economic growth (which is found to be an important driver of household economic mobility), education plays a critical role in upward economic mobility. Assets—specifically, land holdings—are found to be important in preventing downward mobility among rural farm households.<sup>4</sup> The study does not examine the effect of changing climate conditions, however.

Despite the paucity of long-term panel data, a large body of empirical literature on developing countries that examines adaptation to long-term economic changes can inform our consideration of how well rural households—particularly poorer households—will adapt to long-term climate shifts.

### *Adapting Agricultural Input Use and Adopting New Technologies*

Adopting new technologies is one way in which rural agricultural households can adapt to long-term climate shifts. A substantial body of empirical farm

studies, some dating back several decades, examines determinants of farm productivity and profitability in developing countries across Asia, Africa, and Latin America. The evidence indicates that not all households are equally capable of adapting their use of productive inputs or adopting new technologies in the face of changes in the economic environment. Studies consistently demonstrate the importance of farmers' education levels in facilitating higher farm productivity through several channels: (i) direct improvements in worker productivity; (ii) increased ability of workers and enterprise managers to make appropriate decisions about the most efficient mix of inputs and technologies in production; (iii) increased ability of decision makers to adapt to changing technologies, policies, and the external economic environment; and (iv) at sufficiently high levels of education, an ability to generate productivity-enhancing innovations.<sup>5</sup>

### *Diversifying Household Incomes*

Diversifying household income sources, both within agriculture and toward non-agricultural enterprises, is another form of adaptation that can potentially help households manage longer-term climate shifts. The empirical literature suggests that income diversification strategies are often at the heart of rural households' ability to escape poverty (World Bank 2007). However, as the earlier discussion on income smoothing suggests, not all households are equally capable of diversifying their income sources in an efficient, welfare-enhancing way.

Recent evidence using panel data from Nicaragua shows that family education levels, wealth, control of physical and financial assets (specifically, land and credit), more diversified income portfolios in the initial period, and greater access and proximity to markets (and urban centers of nonagricultural activity) are all positively associated with households' economic mobility, income diversification, and income growth (Vakis, Kruger, and Mason 2006). The findings from Nicaragua are consistent with those from other recent panel studies. In El Salvador, for example, households with higher levels of education, greater asset holdings, and greater proximity to markets experienced faster real consumption growth over the period analyzed (Beneke de Sanelú and Shi 2004).

### *Migration*

A third form of adaptation that can potentially help households manage longer-term climate shifts is migration away from affected regions, either to other rural areas or to urban areas. Two studies on Brazil find that weather-related factors currently induce emigration from affected areas (Assuncao and Chein Feres 2008, Mueller and Osgood 2009).

Consistent with the evidence on input use, technology adoption, and income diversification, migration appears to be positively associated with people's education levels and income. In addition, the propensity to migrate appears to be positively associated with initial income levels. A recent study from Bangladesh suggests that the poorest households may be least likely to migrate because returns on migration are uncertain and because, relative to wealthier households, the consequences of a negative outcome are severe. Paying for transport to the city and then failing to obtain a job could mean starvation (Bryan, Chowdhury, and Mobarak 2011).

Similarly, analyses of Brazilian and Mexican census data indicate that illiterate people are least likely to migrate, whereas those with higher education have a higher probability of migrating than do other rural inhabitants (World Bank 2007). Higher schooling levels are also associated with higher probabilities of migration from rural areas in China (Du, Park, and Wang 2005).

## Directions for Public Policy

The fact that some rural households, particularly poorer ones, lack sufficient capacity to adapt to climate change suggests an important role for public policy. To date, climate change studies that focus on household adaptation have tended to take one of two approaches to strengthening household adaptation policies. The first approach, common in agency reports, is to establish a comprehensive agenda for action across sectors, often without attention to prioritization across sectors or activities (see, e.g., IPCC 2007). The second, more recent discussion has focused largely on what has been termed a no-regrets approach to fostering household adaptation (see Heltberg, Siegel, and Jorgensen 2009; de la Torre, Fajnzylber and Nash 2009; and World Bank 2009).

The no-regrets approach emphasizes policies and investments that are part of a sound development and poverty-reduction strategy for developing countries, actions that can generate positive net benefits “under all future scenarios of climate change and impacts” (Heltberg, Siegel, and Jorgensen 2009). Such an approach might include, for example, efforts to increase educational attainment among members of rural households, improving poor households' control of economic assets (such as land and livestock), deepening of rural financial markets, and improving rural households' access to markets through further investments in infrastructure.

The evidence presented here suggests that even in the face of considerable uncertainty regarding climate change—or perhaps precisely because of the uncertainties—there may be occasion for more specific attempts to strengthen rural households' adaptation capacity. We highlight three such areas here: (i) systems of information

for adaptation, (ii) climate-adapted risk-management instruments (such as indexed production insurance), and (iii) social safety nets that are targetable and scalable on the basis of weather events.

### *Information for Adaptation*

Because signal detection and the evaluation of the expected consequences are critical to households' ability to respond to climate change, developing systems of information for adaptation will may improve households' ability to adapt to both weather shocks and long-term climate shifts.

Indeed, better information on expected rainfall and seasonal temperatures, coupled with appropriate agricultural extension advice, can help farm households to make more appropriate decisions about crop mix, irrigation, the timing of planting, and other ex ante production options. Since the major ENSO event of 1982/1983, various governments have attempted to affect the behavior of agricultural households by producing and disseminating weather forecast information (Magrin et al. 2007). In the northeastern Brazilian state of Ceará, for example, a government agency engaged poor rural households in the Planting Time Program. Households would receive high-yield seeds in exchange for a portion of their crop; the agency would distribute the seeds at a time deemed most suitable for planting, given seasonal weather forecasts. The agency also encouraged households to plant drought-resistant crops before the low-rainfall season of 1991. Various assessments considered this intervention highly successful (Orlove and Tosteson 1999). Likewise, in Mexico, ENSO forecasts have been used to adjust crop mix (Magrin et al. 2007).

The development of effective information systems for adaptation will face important implementation challenges. As Orlove and Tosteson (1999) note, the institutions that produce climate forecasts do not traditionally interact with the institutional end users of those forecasts (such as food security planning agencies), much less with households or individuals. Establishing relationships between producers and users of climate forecasts can be complex. The Brazilian agency mentioned above, for example, struggled with sharp public criticism after issuing incorrect forecasts in the mid-1990s.

In this context, Patt and Gwata (2002) identify six factors that are considered critical if climate forecasts are to be effective in strengthening rural households' ability to adapt to climate change. To induce the desired effect on farmer behavior, climate information needs to be (1) considered credible, (2) communicated by agents who are considered legitimate (that is, not beholden to special interests), (3) calculated at appropriately detailed (that is, local) scales, (4) expressed in terms that the audience can understand, (5) appropriately timed, and (6) relevant to the decisions that farmers actively make. In this sense, the process of developing

effective interventions to strengthen household adaptive capacity requires the active participation of the households themselves (Schröter, Polsky, and Patt 2005).

Although it is difficult to quantify the expected gains associated with developing information systems for adaptation, several studies have attempted to estimate how past information dissemination programs have increased net returns to agricultural production. Estimates range from 5 percent for Argentine soybeans (Magrin and Travasso 2001) to 20 percent for Mexican maize (Jones 2001). Attention to effective communication strategies and the effective use of new communications technologies may serve to increase net returns associated with the provision of information for adaptation. Efforts to disseminate weather and other information relevant to production via cell phone text messages, a method that is increasingly viable given the high and rising cell phone penetration rates in poor rural areas, could help to expand outreach in some countries while lowering the costs of disseminating information. Such approaches are already being used to deliver market price information to farmers in developing countries.<sup>6</sup>

### *Climate-adapted Production Insurance*

As discussed above, changes in average temperature and rainfall patterns over time will require adaptation of the traditional insurance model to address the fact that mean climate (and agronomic) conditions change in a systematic way over time.<sup>7</sup> More frequent and more intense weather shocks may also warrant a closer link between climate variables and production insurance, perhaps building on a small but a growing body of experience with weather-indexed insurance.

Indeed, weather-indexed insurance may hold some potential for mitigating the negative effects of climate change. In an attempt to address the moral hazard and adverse selection issues that plague traditional crop insurance, some insurers now offer products in which payouts are linked to the occurrence of a weather event, such as low rainfall, which can be observed without a costly loss assessment procedure. Larson, Anderson, and Varangis (2004) provide a useful review of weather-indexed insurance and related risk management tools. One difficulty, they note, is that these products require reliable climate-related information and therefore significant weather station infrastructure. Basis risk (that is, the risk that index-based payouts will correspond only poorly with actual losses) may also constitute a significant barrier to the purchase of index insurance products (e.g., Clarke 2011), although informal risk sharing that protects farmers against individual-level idiosyncratic shocks may mitigate the problem of basis risk (Mobarak and Rosenzweig 2012).

Experience with weather-indexed insurance remains relatively limited, but several recent applications are being monitored with interest to determine whether they may have broader applicability in developing countries. For

example, weather indices are currently being used for agricultural insurance in Ontario and Alberta, Canada. In Argentina, in light of the positive correlation between rainfall and milk yields, a rainfall insurance contract is being used by a milk-producing cooperative. In Mexico, the agricultural reinsurance company AGROSEMEX recently used weather derivatives to protect part of its crop reinsurance portfolio exposed to weather risks.

Index insurers cannot begin with a few clients and then expand (as in microfinance). Because of the covariate nature of weather shocks, index insurers must begin their operations with many clients in numerous areas (Morduch 2006). In this context, the potential of weather-index insurance to strengthen adaptive capacity may be constrained by smallholder farmers' demand for such instruments, although recent work from India suggests that the availability of informal insurance may increase the demand for index insurance under certain circumstances (Mobarak and Rosenzweig 2012). Several pilot projects, in which information was provided and smallholder farmers' demand for insurance was "aggregated" by local producer organizations and cooperatives, such as those in El Salvador and Nicaragua, may provide lessons on how best to link such risk mechanisms to those who need them (Varangis et al. 2003).

### *Social Safety Nets*

Safety nets represent another important component of a risk management strategy for climate change and are well suited to fortifying households' ability to adapt to both shocks and shifts. In the face of recent shocks (related both to weather and economic shocks), it has become common to call for flexible safety net programs that can respond to affected households' needs without compromising governments' long-term fiscal positions. Indeed, a number of developing countries around the world have adapted their safety net programs to help poor and crisis-affected households deal with the recent food, fuel, and global financial crises.

One example of successful flexibility along these lines is the Honduras Social Investment Fund (FHIS). Credited with limiting the rise of poverty in the wake of Hurricane Mitch (which killed thousands of Hondurans, left one million homeless and inflicted damage equivalent to two-thirds of GDP), the FHIS was a public program created in 1990 to finance small-scale investments in poor communities. Originally conceived as a way to address the social impacts of structural adjustment policies, the FHIS nimbly transformed into an emergency response program after Mitch devastated the country in 1998. Within 100 days of the hurricane, the program approved \$40 million for 2,100 community projects; by the end of 1999, the FHIS had financed 3,400 projects, four times the number financed in a comparable prehurricane period (Warren 2003).<sup>8</sup>

There have been more recent experiences with safety net programs helping to protect the poor from shocks, including weather shocks. In 2000, Nicaragua established a conditional cash transfer program called *Red de Protección Social* (RPS). Although this program was developed to address structural poverty and to foster human capital development among the poor, the RPS was also effective in reducing the vulnerability of poor rural households to income shocks. During a sharp downturn in world coffee prices from 2000 to 2003, consumption among coffee-growing households that were RPS beneficiaries declined by only 2 percent, whereas consumption among coffee-growing households who were not beneficiaries declined by 30 percent (Vakis, Kruger, and Mason 2006). Building on the success of the RPS, the government of Nicaragua introduced a cash-transfer pilot program in a drought-prone region. After a mudslide in eastern Nicaragua, the program quickly responded by transferring resources to all affected households.<sup>9</sup> An evaluation of the pilot program found that the income and consumption of beneficiaries was significantly more resilient to droughts, price shocks, and health shocks than that of nonbeneficiary families.

One key feature of the Nicaragua pilot was targeting the program, *ex ante*, to geographic regions of the country that faced elevated levels of weather-related risk (specifically drought risk, in this case).<sup>10</sup> Indeed, as with climate-adapted insurance mechanisms, safety net programs designed to protect households from increasingly frequent and severe weather shocks must address the increased information-related challenges associated with targeting on the basis of weather risk. Recent advances in mapping technology, combined with increasingly precise weather-risk forecasts, can help policy makers to address these issues.

## Conclusion

Under the best of circumstances, agricultural livelihoods are difficult to sustain and improve. Even in the friendliest of agronomic environments, small-scale farming can be incredibly challenging and risky. For those rural households in developing countries that depend on agriculture, climate change adds yet another dimension of difficulty to the problem of maintaining and improving welfare. The same is true, albeit via different mechanisms, for rural households whose income does not come primarily from agricultural activity.

This paper has argued that climate change poses two distinct, if interrelated, sets of challenges for poor rural households: challenges related to the increasing frequency and severity of weather *shocks* and those related to long-term *shifts* in temperature, rainfall patterns, water availability, and other environmental factors. Although much is unknown about how poor rural households will adapt to these shocks and shifts, a substantial body of empirical evidence from developing

countries provides a reasonably robust sense of the strengths and limitations of available adaptation mechanisms.

An important and encouraging finding from the literature is that rural households—even poor ones—possess a number of strategies for mitigating the effects of shocks associated with climate change. They may attempt to smooth consumption in the face of income shocks (using credit, insurance, savings, and other methods), and they may attempt to smooth income (by adjusting agricultural portfolios or diversifying sources of income, for example). The evidence indicates, however, that households' abilities to smooth consumption and income are less than perfect. Moreover, some strategies, such as employing child labor, come with clear risks.

Longer-term climate shifts create other challenges to rural households' livelihoods and call for other types of adaptation responses. Rural agricultural households confronted with significant environmental and agronomic changes due to climate change will need to adjust their livelihood strategies if they are to maintain or increase their incomes. Adapting to climate shifts may require farming households to adjust their cropping choices and alter their use of productive inputs and technologies. Farmer households may also seek to diversify their income sources or to migrate to areas less affected by long-term changes in temperatures and precipitation.

Understanding the factors that affect rural households' capacity to adapt to shocks and shifts will be important in designing effective interventions to strengthen their adaptation capacity. Here, the evidence provides some consistent messages: policies and programs that help to increase households' human capital and improve their access to physical assets, financial capital, and markets can play important roles in strengthening their capacity to adapt to climate change. In this sense, many interventions that can fortify rural households' adaptive capacity are also part of an effective long-term development and poverty-reduction strategy for developing countries. Hence, there is an appeal to a no-regrets approach that focuses on interventions with positive net benefits regardless of how the shocks and shifts associated with climate change unfold.

Yet, perhaps precisely because of the inherent uncertainties of climate change, an intense focus on the adaptive capacity of rural households warrants consideration. We have discussed three potential focus areas. In addition, investing in improved household data collection and analysis, including the development of longer-term panels of data that can facilitate direct observation and analysis of households' and policy makers' responses to climate-related shocks and shifts, will ultimately improve our understanding of adaptation to climate change in rural areas.

## Notes

\*Andrew D. Mason is with the Poverty Reduction and Economic Management Department, East Asia and Pacific Region, The World Bank, Washington, DC 20433. Javier Baez is with the Poverty Reduction and Economic Management Department, Latin America and Caribbean Region, The World Bank, Washington DC 20433. Email: [jbaez@worldbank.org](mailto:jbaez@worldbank.org). Dorothy Kronick is at Stanford University. Email: [dkronick@stanford.edu](mailto:dkronick@stanford.edu). An earlier version of this article was written as a background paper to the World Bank study for Latin America and the Caribbean, *Low Carbon, High Growth: Latin American Responses to Climate Change*. The authors gratefully acknowledge support and feedback from Augusto de la Torre, Pablo Fajnzylber, and John Nash, the lead authors of that report, as well as Helena Ribe from the Bank's Human Development Department for Latin America and the Caribbean. We also thank three anonymous referees for their comments. The corresponding author is Andrew D. Mason; email address: [amason@worldbank.org](mailto:amason@worldbank.org). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, the Executive Directors of the World Bank, or the governments that they represent.

1. Although climate change creates benefits and opportunities for rural households in regions where expected shifts may actually improve local agronomic conditions, we focus here on household adaptation to expected adverse shocks and shifts.

2. In the context of climate change, behavioral factors will likely be linked to economic factors through human capital in the sense that people's perceptions of risk and, even, of their own adaptive capacity are a function of their ability to detect and evaluate the actual (or likely) impacts of a weather/climatic event. Although observing and evaluating the impact of shocks may be relatively straightforward, assessing the risks associated with shifts can be much more complex.

3. Overall, panel studies support the notion that education, assets, access to finance, and access to markets (to different degrees in different places) are important for household income growth and the movement of the poor out of poverty (see [Dercon and Shapiro 2007](#) for a review of evidence). [Dercon and Shapiro \(2007\)](#) also note that risk and shocks serve to inhibit household income growth and mobility. In that sense, households' vulnerability to shocks can affect their ability to adapt to long-term climate shifts.

4. Interestingly, among smallholder farmers, larger landholdings are found to inhibit movement toward nonagricultural employment, perhaps by strengthening these farmers' ties or commitment to farming activities.

5. Among the studies presenting empirical findings on the relationship between farmer education and agricultural productivity, including the effect of education on input use and technology adoption, see [Lockheed et al. 1980](#); [Jamison and Lau 1982](#); [Mooch and Addou 1994](#); [Hussain and Byerlee 1995](#); and [World Bank 2007](#).

6. For example, seminomadic herders in rural Mongolia can now subscribe to a service through which they can obtain market price data on key commodities in the nearest district towns through their cell phones.

7. It should be noted that this, by itself, argues for better, more accessible, information on local shifts in temperature and rainfall levels and may generate incentives for the production and dissemination of such information, independent of the perceived demand among rural households.

8. Projects prioritized clearing debris and repairing or rebuilding water lines, sanitation systems, roads, bridges, health centers, and schools, thus hastening national recovery and generating approximately 100,000 person-months of employment in the three months following the crisis. The decentralized structure and institutional flexibility of the FHIS facilitated its brisk and influential response. Building on strong preexisting partnerships with municipalities and communities, FHIS directors established 11 temporary regional offices and quickly delegated resources and responsibilities.

Directors reduced the number of steps in the subproject cycle from 50 to 8, established safeguards to ensure accountability and transparency, and effectively accessed International Development Association financing.

9. The transfers took place within three months of the event; by the sixth month after the event, RPS had enrolled all affected households as beneficiaries.

10. Beneficiaries were selected from among RPS enrollees using a proxy-means test.

## References

- Alderman, H., J. Hoddinott, and B. Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers* 58 (3): 450–74.
- Assuncao, J., and F. Chein Feres. 2008. "Climate Migration." Background Paper for World Bank Flagship Report on Climate Change in Latin America, *Low Carbon, High Growth: Latin American Responses to Climate Change*, World Bank Latin American and Caribbean Studies.
- Baez, J., and I. Santos. 2007. "Children's Vulnerability to Weather Shocks: A Natural Disaster as a Natural Experiment." Ph.D. diss. Syracuse University, Maxwell School of Public Affairs, Syracuse, NY, and Harvard University, Kennedy School of Government, Cambridge, MA.
- Beegle, K., R. Dehejia, and R. Gatti. 2003. "Child Labor, Crops Shocks and Credit Constraints." NBER Working Paper 10088. National Bureau of Economic Research, Cambridge, MA.
- Beneke de Sanfeliu, M., and M. Shi. 2004. *Dinámica Del Ingreso Rural En El Salvador*. 1st ed. San Salvador, El Salvador: Fundación Salvadoreña para el Desarrollo Económico y Social, FUSADES, 2004.
- Binswanger, H., and M. Rozensweig. 1993. "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *The Economic Journal* 103:56–78.
- Blankespoor, B., S. Dasgupta, B. Laplante, and D. Wheeler. 2010. "The Economics of Adaptation to Extreme Weather Events in Developing Countries." Working Paper 199. Center for Global Development, Washington, DC.
- Bouma, M., and C. Dye. 1997. "Cycles of Malaria Associated with El Niño in Venezuela." *Journal of the American Medical Association* 278:1772–74.
- Bouma, M., G. Poveda, W. Rojas, D. Chavasse, M. Quinones, J. Cox, and J. Patz. 1997. "Predicting High-Risk Years for Malaria in Colombia Using Parameters of El Niño Southern Oscillation." *Tropical Medicine and International Health* 2:1122–27.
- Bryan, G. 2010. "Ambiguity and Insurance." Ph.D. diss. Yale University, Department of Economics, New Haven, CT. (<http://www.asb.unsw.edu.au/schools/economics/Documents/G.%20Bryan%20-%20Ambiguity%20and%20Insurance.pdf>).
- Bryan, G., S. Chowdhury, and A. M. Mobarak. "Seasonal Migration and Risk Aversion." Discussion Paper DP8739. Yale University, Department of Economics, New Haven, CT. (<http://faculty.som.yale.edu/mushfiqmobarak/migration.pdf>).
- Bull, G.M., and J. Morton. 1978. "Environment, Temperature, and Death Rates." *Age and Aging* 7 (4): 210–24.
- Clarke, Daniel. 2011. "A Theory of Rational Demand for Index Insurance." Discussion Paper Series 572. Department of Economics, Oxford University, Oxford.
- De Janvry, A., F. Finan, E. Sadoulet, and R. Vakis. 2006. "Can Conditional Cash Transfers Serve as Safety Nets in Keeping Children at School and from Working when Exposed to Shocks?" *Journal of Development Economics* 79 (2): 349–73.

- De la Torre, A., P. Fajnzylber, and J. Nash. 2009. *Low Carbon, High Growth: Latin American Responses to Climate Change*. World Bank, Washington DC.
- Dercon, S., and P. Krishnan. 2000. "In Sickness and in Health: Risk-Sharing Within Households in Rural Ethiopia." *Journal of Political Economy* 108 (41): 688–727.
- Dercon, S., and J.S. Shapiro. 2007. "Moving on, Staying Behind, Getting Lost: Lessons on Poverty Mobility from Longitudinal Data." In D. Narayan, and P. Petesch, eds., *Moving Out of Poverty: Cross-Disciplinary Perspectives on Mobility*, pp. 77–126. Washington DC and New York: A co-publication of the World Bank and Palgrave Macmillan.
- Deressa, T., R. Hassan, and C. Ringler. 2009. "Household Vulnerability to Climate Change: The Case of Farmers in the Nile Basin of Ethiopia." Discussion Paper 00935. International Food Policy Research Institute, Washington, DC.
- Du, Y., A. Park, and S. Wang. 2005. "Migration and Rural Poverty in China." *Journal of Comparative Economics* 33:688–709.
- Fuwa, N. 2007. "Pathways Out of Rural Poverty: A Case Study in Socio-economic Mobility in the Rural Philippines." *Cambridge Journal of Economics* 31:123–44.
- Goland, C. 1993, "Field Scattering as Agricultural Risk Management: A Case Study from Cuyo Cuyo, Department of Puno, Peru." *Mountain Research and Development* 13 (4): 317–38.
- Grothmann, T., and A. Patt. 2005. "Adaptive capacity and human cognition: The process of individual adaptation to climate change." *Global Environmental Change* 15:199–213.
- Heltberg, R., P. Siegel, and S. Jorgensen. 2009. "Addressing Human Vulnerability to Climate Change: Toward a 'No Regrets' Approach." *Global Environmental Change* 19:89–99.
- Hussain, S.S., and D. Byerlee. 1995. "Education and Farm Productivity in Post-Green Revolution Asia." In G.H. Peters, and D. Headley, eds., *Agricultural Competitiveness, Market Forces, and Policy Choices*. Proceedings of the 22<sup>nd</sup> Conference of the International Agricultural Economics Association, Aldershot, UK: Dartmouth Publishers.
- Intergovernmental Panel on Climate Change (IPCC). 2007. *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. eds. M.L. Parry, O.E. Canziani, J.P. Palutikof, P.J. van der Linden, and C.E. Hanson. Cambridge, UK: Cambridge University Press.
- Jacoby, H., and E. Skoufias. 1997. "Risk, Financial Markets, and Human Capital in a Developing Country." *Review of Economic Studies* 64 (3): 311–35.
- Jalan, J., and M. Ravallion. 1999. "Are the Poor Less Well Insured? Evidence on Vulnerability to Income Risk in Rural China." *Journal of Development Economics* 58 (1): 61–81.
- Jamison, D. T., and L. Lau. 1982. *Farmer Education and Farm Efficiency*. Baltimore: Johns Hopkins University Press.
- Jones, J. 2001. "Comparative Assessment of Agricultural Uses of ENSO-based Climate Forecasts in Argentina, Costa Rica and Mexico: Project Summary." IAI Initial Science Program – III Project, University of Florida, Gainesville, FL.
- Kandlikar, M., and J. Ribsey. 2000. "Agricultural Impacts of Climate Change: If Adaptation is the Answer, What is the Question?" *Climate Change* 45:529–39.
- Kochar, A. 1999. "Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India." *The Review of Economics and Statistics* 81 (1): 50–61.
- Larson, D., J. Anderson, and P. Varangis. 2004. "Policies on Managing Risk in Agricultural Markets." *The World Bank Research Observer* 19 (2): 199–230.
- Lockheed, M.E., T. Jamison, and L. J. Lau. 1980. "Farmer Education and Farm Efficiency: A Survey." *Economic Development and Cultural Change* 29 (1): 37–76.

- Magrin, G., C. Gay Garcia, D. Cruz Choque, J. C. Gimenez, A. R. Moreno, G. J. Nagy, C. Nobre, and A. Villamizar. 2007. "Latin America." In M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden, C.E. Hanson, eds., *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 582–615. Cambridge, UK: Cambridge University Press.
- Magrin, G., and M. Travasso. 2001. "Economic Value of ENSO-based Climatic Forecasts in the Agricultural Sector of Argentina." 2nd International Symposium Modelling Cropping Systems, European Society of Agronomy (ESA), Florence, Italy: 139–40.
- Mobarak, A. Mushfiq, and Mark Rosenzweig. 2012. "Selling Formal Insurance to the Informally Insured." Working Paper 97. Yale University, Department of Economics, New Haven, CT. (<http://faculty.som.yale.edu/mushfiqmobarak/insurance.pdf>).
- Mooch, P.R., and H. Addou. 1994. "Agricultural Productivity and Education." In T. Husen, and N. Postlethwaite, eds., *International Encyclopedia of Education* Vol. 1. Oxford: Pergamon Press.
- Morduch, J. 1995. "Income Smoothing and Consumption Smoothing." *Journal of Economic Literature* 37 (4): 1569–1614. In S. Dercon, ed., *Insurance Against Poverty*. Oxford: Oxford University Press.
- . 2006. "Micro-insurance: The next revolution?" In A. Banerjee, R. Bénabou, and D. Mookherjee, eds., *Understanding Poverty*, pp. 337–356. Oxford: Oxford University Press.
- Mueller, V. 2010. "Short- and Long-Term Effects of the 1998 Bangladesh Flood on Rural Wages." Discussion Paper 00956. International Food Policy Research Institute, Washington, DC.
- Mueller, V., and D. Osgood. 2009. "Long-Term Consequences of Short-term Precipitation Shocks: Evidence from Brazilian Migrant Households." *Agricultural Economics* 40:573–586.
- National Institute of Environmental Health Sciences. 1999. "Environmental Impacts of Hurricane Mitch." *Environmental Health Perspectives* 107 (3): A139–40.
- Nhemachena, C., and R. Hassan. 2007. "Micro-Level Analysis of Farmers' Adaptation to Climate Change in Southern Africa." Discussion Paper 00714. International Food Policy Research Institute, Washington, DC.
- Organisation for Economic Cooperation and Development (OECD; E. Levina and D. Tirpak). 2006. "Key Adaptation Concepts and Terms." Draft paper, Agenda Document 1, OECD/IEA Project for the Annex 1 Expert Group on the UNFCCC.
- Orlove, B., and J. Tosteson. 1999. "The Application of Seasonal to Interannual Climate Forecasts Based on El Niño-Southern Oscillation (ENSO) Events: Lessons from Australia, Brazil, Ethiopia, Peru, and Zimbabwe." Working Paper 99-3. Berkeley Workshop on Environmental Politics, University of California, Berkeley.
- Patt, A., and C. Gwata. 2002. "Effective seasonal climate forecast applications: Examining constraints for subsistence farmers in Zimbabwe." *Global Environmental Change* 12:185–95.
- Paxson, C. 1992. "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." *American Economic Review* 82 (1): 15–34.
- Rosenzweig, M., and O. Stark. 1989. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." *Journal of Political Economy* 97 (4): 905–26.
- Rumel, D., L. Riedel, M. Latorre, and B. Duncan. 1993. "Infarto do Miocardo e Acidente Vascular Cerebral Associados a Alta Temperatura e Monóxido de Carbono em Área Metropolitana do Sudeste do Brasil." *Revista de Saude Publica* 27 (1): 15–22.
- Santos, I. 2007. "Risk-Sharing and the Role of Inter-Household Transfers after a Natural Disaster. Evidence from Hurricane Mitch in Nicaragua." Ph.D. diss. Kennedy School of Government, Harvard University, Cambridge, MA.
- Schröter, D., C. Polisky, and A. Patt. 2005. "Assessing Vulnerabilities to the Effects of Global Change: An Eight-Step Approach." *Mitigation and Adaptation Strategies for Global Change* 10:573–96.

- Shumway, R., A. Azari, and Y. Pawitan. 1988. "Modeling Mortality Fluctuations in Los Angeles as Functions of Pollution and Weather Effects." *Environmental Research* 45:224–41.
- Timmins, C. 2003. "Health and Climate Change in the Developing World: The Impact of Global Warming on Brazilian Morbidity Patterns." Working Paper. Duke University, Durham, NC.
- Townsend, R. 1994. "Risk and Insurance in Village India." *Econometrica* 62 (3): 539–91.
- . 1995. "Financial Systems in Northern Thai Villages." *The Quarterly Journal of Economics* 110 (4): 1011–46.
- Turner, B.L., R.E. Kasperson, P.A. Matson, J.J. McCarthy, R.W. Corell, L. Christensen, N. Eckley, J.X. Kasperson, A. Luers, M.L. Martello, C. Polsky, A. Pulsipher, and A. Schiller. 2003. "A Framework for Vulnerability Analysis in Sustainability Science." *Proceedings of the National Academy of Sciences, United States of America* 100:8074–9.
- Ureta, M. 2005. "Hurricane Mitch, Family Budgets and Schooling in Nicaragua." Working Paper. Department of Economics, Texas A&M University, College Station, Texas.
- Vakis, R., D. Kruger, and A. Mason. 2006. "Shocks and Coffee: Lessons from Nicaragua." In *Income Stabilization in Agriculture: The Role of Public Policies*, Proceedings of the 86<sup>th</sup> EAAE Seminar, Anacapri, Italy: 171–231.
- Varangis, P., P. Siegel, D. Giovannucci, and B. Lewin. 2003. "Dealing with the Coffee Crisis in Central America: Impacts and Strategies." World Bank Policy Research Working Paper 2993. World Bank, Washington, DC.
- Vos, R., M. Velasco, and E. Labastida. 1999. "Economic and Social Effects of El Nino in Ecuador, 1997-1998." Inter-American Development Bank, Sustainable Development Department, Technical Paper Series, Washington, DC.
- Warren, D. 2003. "Coping with a Natural Disaster: Hurricane Mitch and the Honduran Social Investment Fund." *Spectrum Magazine*, Fall: 20–21.
- World Bank. 2005a. "Shocks and Social Protection: Lessons from the Central American Coffee Crisis." Central America Department, Latin America and Caribbean Region, Washington, DC.
- . 2005b. "DR-CAFTA: Challenges and Opportunities for Central America." Central America Department and Office of the Chief Economist, Latin America and Caribbean Region, Washington, DC.
- . 2007. "World Development Report 2008: Agriculture for Development." World Bank, Washington, DC.
- . 2010. "World Development Report 2010: Development and Climate Change." World Bank, Washington, DC.
- World Health Organization (WHO). 1999. "El Niño and Health – Protection of the Human Environment." Protection of the Human Environment: Task Force on Climate and Health, WHO/SPE/HE/99.4, Geneva.

**THE WORLD BANK**

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

OXFORD  OPEN



**SFI**<sup>®</sup>

Certified Sourcing

[www.sfiprogram.org](http://www.sfiprogram.org)  
SFI-00453

ISBN 978-0-19-967964-5



9 780199 679645