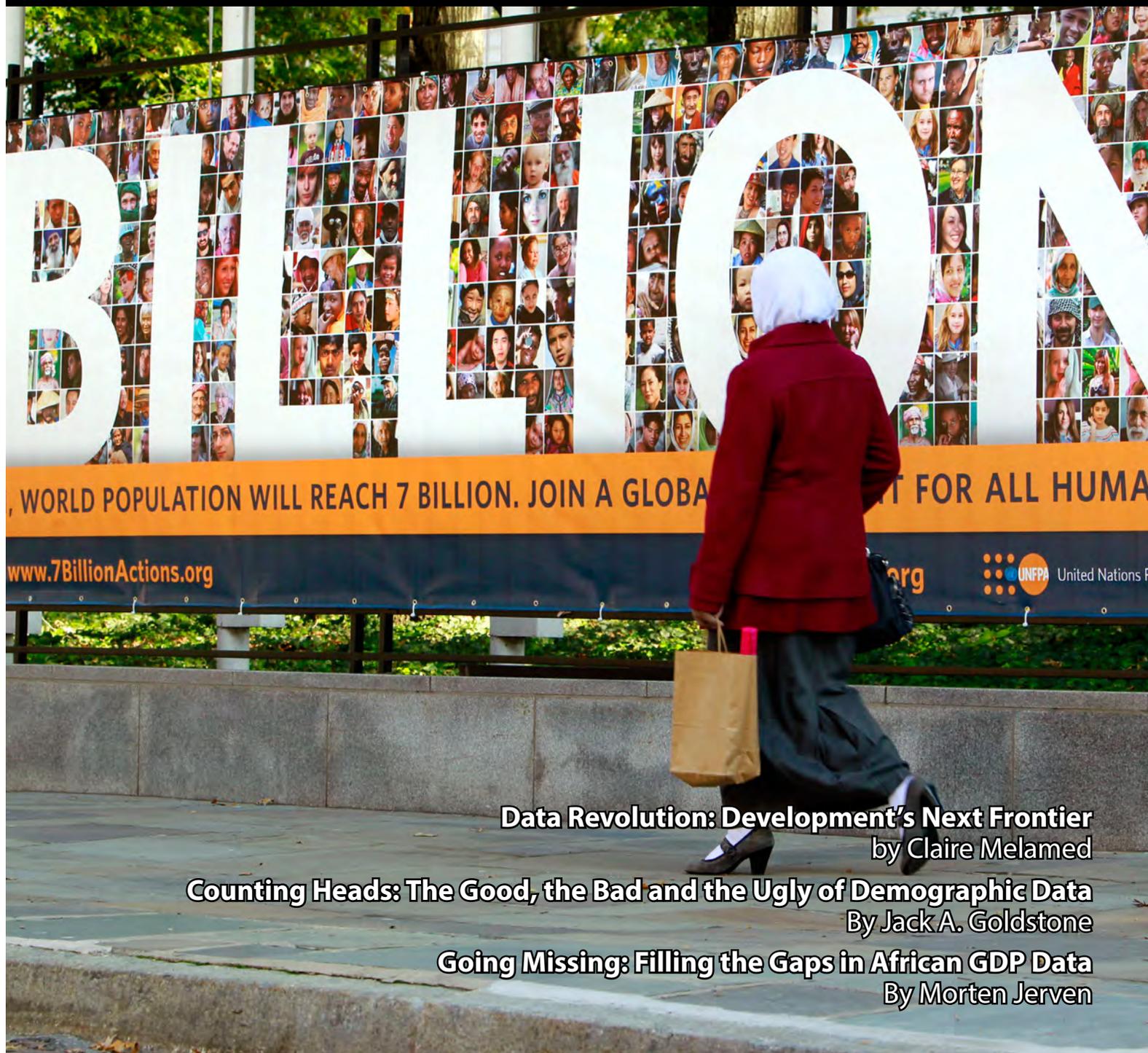


### From Bigger to Better: Data Quality in Aid and Development



**Data Revolution: Development's Next Frontier**  
by Claire Melamed

**Counting Heads: The Good, the Bad and the Ugly of Demographic Data**  
By Jack A. Goldstone

**Going Missing: Filling the Gaps in African GDP Data**  
By Morten Jerven

Compilation © 2014 World Politics Review LLC.

“Data Revolution: Development’s Next Frontier,” © 2014 Claire Melamed. Used under license.

“Counting Heads: The Good, the Bad and the Ugly of Demographic Data,” © 2014 Jack A. Goldstone. Used under license.

“Going Missing: Filling the Gaps in African GDP Data,” © 2014 Morten Jerven. Used under license.

First published in 2014 by World Politics Review

World Politics Review  
231 Front St, Suite 204  
Brooklyn, NY 11201  
[www.worldpoliticsreview.com](http://www.worldpoliticsreview.com)  
(202) 596-9771

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, email the publisher at [reprints@worldpoliticsreview.com](mailto:reprints@worldpoliticsreview.com).

Published in the United States of America

Cover image: Banner for a global campaign by the U.N. Population Fund UNFPA, New York, Nov. 1. 2011 (U.N. photo by Rick Bajornas).

DATA REVOLUTION: DEVELOPMENT'S NEXT FRONTIER BY CLAIRE MELAMED	4
COUNTING HEADS: THE GOOD, THE BAD AND THE UGLY OF DEMOGRAPHIC DATA BY JACK A. GOLDSTONE	9
GOING MISSING: FILLING THE GAPS IN AFRICAN GDP DATA BY MORTEN JERVEN	14
ABOUT WPR	19

# DATA REVOLUTION: DEVELOPMENT'S NEXT FRONTIER

BY CLAIRE MELAMED



Governments have been collecting data on their citizens almost from the first moment that they came into being. Data was needed to determine what was out there that could be extracted: The Egyptian Pharaohs conducted a census to find out the scale of the available labor force to build the pyramids, and in the Roman Empire, the five-yearly census was all about finding out who was available for military service and what wealth existed to be taxed. But governments have also used data to find out what people needed: The ancient Babylonians collected data from their citizens nearly 6,000 years ago in order to understand how much food was required to feed their population, and the Egyptian census was also used to work out how to divide the land after the flooding of the Nile.

As the activities of governments have become more complex, and people's expectations have grown, more and more data has been collected. Early censuses just recorded numbers of people, with basic information about wealth and maybe occupation. It's all more complicated now—the most recent U.K. census, held in 2011, had 56 different questions. And that's not even including data collected by the health service, the education department or the myriad other public and private bodies who collect and use information about individuals.

A lot is known about some people, and that information is used by governments and other institutions to design, implement and evaluate policy on their behalf. But very little is known about others—and the decisions that are made for and about them are likely to be much worse as a result.

The state of data on some of the world's people is so bad that even facts that appear incontrovertible are highly questionable. [According to the latest figures](#), maternal mortality rates in Africa dropped from 740 deaths per 100,000 births in 2000 to 500 deaths per 100,000 births in 2010. Good news. But look more closely and you find that actual data from civil registration systems on maternal mortality is only available for 16 percent of all the world's births. The vast majority of what we think we know about maternal mortality is the product of estimates—basically statistical guesswork—high-quality and serious stuff, yes, but still guesswork.

For the purposes of official U.N. statistics, maternal mortality [is modeled using three variables](#)—GDP, fertility rates and the national probability of having a skilled birth attendant present. And data on all of these is in turn highly unreliable—famously, Ghana's GDP went up by 60 percent overnight after a recalculation; fertility rates are hard to measure in countries without national systems for registering births; and the figures on birth attendants are derived from survey data, which is incomplete, given the fact that only 28 out of 49 African countries [have had a survey](#) in the past seven years.

So while the “official” number of maternal deaths per 100,000 live births in sub-Saharan Africa in 2000 was 500, it could in fact be as low as 400 or as high as 750. In other words, there might not have been any fall in the maternal mortality rate at all. In fact, given the uncertainty in both the 2000 and the 2010 estimates, the maternal mortality rate could even have gone up.

Does it matter if governments, service providers and people in general don't have data? Some pretty impressive things have been done in the world of development with terrible data—smallpox, for example, was eradicated despite a lack of a census or basic health information in many countries. Progress is possible without data.

But this is true only in some areas, and only up to a point. The range and complexity of what governments are expected to do make life almost impossible without good data. Say a government wants to run a health service and allocate resources to it in a way that helps to reduce inequalities so that the sickest get more resources. In the U.K., which has plentiful data, the formula for allocating health care financing is based on the population in a given region, on the cost of supplying services in different areas and on a range of demographic factors relevant to health needs and health inequalities such as age, death rates, HIV rates, indicators of poverty and deprivation and even, for a time, the size of the homeless population in a given area. In addition to data on inputs, data on outcomes is also collected regularly so that the government can judge quite quickly if its attempts to improve people's health are working or not.

Much of that information is simply not available to policymakers in large numbers of countries. Malaria, for example, is a leading cause of death and ill health in poor countries, so for allocating health resources in any country, information about where malaria is most prevalent is crucial, and progress on tackling malaria would be a key indicator with which to judge the success of health policies. But there **is not enough data to be certain** about either the incidence of malaria or trends in tackling it in countries that account for 85 percent of all (estimated) deaths from the disease. Malaria is just one example: Infection rates for key diseases such as HIV and AIDS are often extrapolated from the numbers among particular population groups—such as pregnant women visiting clinics, where HIV tests are mandatory. But these can be wrong—the actual rates of infection in Ethiopia, for example, **turned out to be about half the estimated rates** once a national survey was carried out.

Bad data is a recipe for bad decisions. Without data, governments and other institutions and individuals trying to make policies or run programs in any sector are operating almost blindly. They can't know what inputs are needed where. And they can't know if the things they are doing are actually working, as there's no way of telling if things are getting better or not. Of course, governments will still make decisions, and some of these decisions will turn out well. But without data it's harder for effective governments to know what to do, and it's harder for people to hold ineffective governments to account without the evidence that data provides of poor outcomes and bad or corrupt decision-making.

The pressures for improving data quality are growing. The creation of globally agreed sets of indicators to monitor human progress, most famously the Millennium Development Goals agreed on in 2000, created new incentives to invest in a limited range of indicators to measure progress. Increasingly, the public is demanding better data too: As democratic government becomes the norm in most countries in Africa, Asia and Latin America, people expect data so they can judge the success of the politicians they vote for, and politicians need data to know how to satisfy people's expectations and be re-elected. Many countries have strong campaigns calling for greater transparency of both outcome and, crucially, input data, so that citizens can track government spending and the results it is producing.

And as the general public in many donor countries becomes more skeptical about the positive effect of the aid they pay for, the onus is on governments to come up with the figures to prove that aid programs are doing something good. Within the development sector, pressure for better evaluations, for trials of new interventions to judge their effects and for a much more rigorous approach to demonstrating effectiveness, has led to much more attention being paid to data collection by those in charge of implementing projects—be they from official agencies, nongovernmental organizations or governments. The Millennium Villages project, a multimillion-dollar development project led by the economist Jeffrey Sachs, has become embroiled in controversy due to poor evaluation procedures and lack of data on control groups against which to measure success or failure.

The general concern for and complaints about data quality have crystallized in a commitment to action in recent months, with the call from the U.N.'s High-Level Panel on the Post-2015 Development Agenda for a “data revolution” and a global partnership on data to provide resources for improvements in data quality and quantity worldwide. This call has grabbed the imagination of official agencies, NGOs and governments worldwide, and there is real potential—at last—for resources and political will to improve what some commentators call the “statistical tragedy” of poor data in poor countries.

The resources involved are huge. Properly staffing and resourcing statistical offices in some countries would itself be a needed first step, which someone would have to pay for. Beyond that, the U.N. estimates the average cost per person of a national census is \$4.60, although costs vary hugely—the cost per person of the most recent census in India was just 50 cents, while in the U.S. it was \$42. The costs of a single census for the whole population of sub-Saharan Africa, at 910.4 million people, would therefore be somewhere between \$450 million, if the costs are closer to those of the Indian census, and \$4 billion, if the costs are closer to the world average. And such a census would have to be repeated every 10 years to produce useful and usable data. Add to that regular household surveys, which should take place around every five years, at around \$1 million to \$2 million each, to collect more in-depth information on trends for key demographic, social and economic issues; other data-collection exercises using mobile technologies, big data and so on; plus the cost of processing data into a form that is usable by governments, other institutions and, crucially, individual citizens; and the resources needed are really quite formidable.

Raising the money would be just the start. Spending it would also produce huge challenges. First among these would be prioritization—what is the data that’s really important to collect? There’s a lot of agreement on basic demographic data that countries should have—information about numbers of people, births and deaths, incomes and assets, health and education levels and so on. But beyond that, there’s an almost infinite range of data that would be very useful—but each extra piece would add to the costs of collection. If, as we should probably assume will happen, a new set of global goals on sustainable development are agreed to in 2015, there will be new requirements for data to measure progress on them. This will include things that have never been measured in a comprehensive way in many countries: food waste, for example, or rates of domestic violence, both of which are likely to feature among the data necessary to measure progress on the new goals, and for which there are nothing approaching credible figures for most countries. A global initiative on data would have to involve some compromise between a globally agreed set of core data and data that reflects different national priorities. But it’s inevitable that this will mean, in very many countries, collecting much more information than is available now.

The scale of ambition, however, must not be confined to data on objective indicators like income or education levels. The time has passed when it’s acceptable for a government or a researcher or an NGO to tell people that their lives are improving, without also asking them what they think. A global initiative on data would have to find out from the world’s citizens what their priorities are, what they feel about their lives and how they perceive the changes happening around them. Thanks to the work of opinion polling companies like Gallup and Ipsos over decades, there’s a huge amount of expertise on how to do this in a rigorous and credible way—and this is the moment to do it.

It’s not just about adding more and more questions to censuses or surveys. There are also some things that surveys might not be able to do, for which new approaches would be needed. Sensitive issues might require different ways of collecting data: Measuring crime rates or illegal activity, for example, through face-to-face surveys might be difficult if people are reluctant to speak openly—anonymous data collection through mobile phones might be more effective.

Sampling might also need to be rethought. A typical global opinion poll, such as Gallup’s World Values Survey, asks questions of somewhere between 1,000 and 2,000 carefully selected people to represent a whole country. Often, this works well. But part of “better” data is also data that allows

for more effective disaggregation of populations—if the commitment of a post-2015 agreement is to end poverty, more information is needed on who is poor to understand why and what to do about it. If, for example, poverty in a given country is overwhelmingly concentrated among a small ethnic group, or among people with a mental illness or a physical disability, then a sampled survey might not pick up a sufficient number of people in that group to really get a picture of their needs and the problems they face. Something more deliberate might be needed to provide the information required to track progress on poverty.

Another part of better data might involve data that's collected more often, or at specific times. In a country where employment is precarious, short-term and very informal, you might need to collect data on employment and wages from a sample of the population every month to really get a good picture of what's going on in the labor market. And in an agricultural economy in which all the money comes in at once at harvest time, there's a premium on asking the questions at the right time. If the researcher happens to come in 11 months after the harvest asking questions about income, people will have forgotten what they earned, and the quality of the data is likely to be worse than it would be if the same questions were asked at the right moment.

So better data means more data, but also data that's collected in different ways and from some specific groups of people. There is, rightly, a huge amount of excitement about the potential for new technologies to help meet some of these challenges. Mobile phones and the Internet can help enormously with coverage and with the more frequent collection of data at a low cost. But, as with all apparent “magic bullets” in the development sector, these technologies are not, in fact, magic bullets.

[Experience with MY World](#), a global survey on people's priorities run by the Overseas Development Institute and the U.N., suggests caution. Though most of the respondents are asked their views through face-to-face surveys or on the Internet, there has also been experimentation with different ways of conducting the survey by mobile phone. Of the more than quarter of a million responses collected via mobile phones, seven out of 10 come from men. This is probably a problem that can be dealt with, but it does illustrate that, even with shiny new technologies, the old rules about representativeness and rigor still apply.

Of course, collecting more data is all very well, and it would be a big step forward from the current situation. But it's nothing unless it is presented in a form that people can actually use. Data is the raw material for information; it's not the information itself. Making data useful would also involve a huge processing and dissemination job to turn the new data into products that could be used by people, organizations and governments to monitor progress and improve decision-making and practice.

The barriers to all of this are political as well as financial and technical. Some facts are known but kept hidden for political reasons by governments seeking to avoid too much scrutiny of their decisions. Increasingly, citizens and NGOs are demanding more data on inputs. How much are governments, official agencies and NGOs spending on different projects or sectors? Where is the money going and to whom? What money is being earned in the country, by what individuals and companies, and what tax is being paid on it? Most of this information is known by somebody—making it public is more about tackling the politics than about technical issues having to do with collection and analysis.

Not all new data will be popular either. Data can make governments or other bodies look bad if outcomes are worse than people think they are. This isn't a new problem: Data from one of the earliest modern censuses, carried out by the Swedish government in 1749, was kept secret because, to the government's surprise, the population turned out to be smaller than expected—something of an embarrassment and a military risk. A data revolution will need constant vigilance and monitoring to make sure that bad as well as good news goes public.

Researchers have been lamenting bad data for decades. It's almost an iron law of development conferences that once two or more researchers are gathered in a room, they will start to talk about bad data. But most people have been unaware of just how bad the situation is—and of the consequences of knowing so very much less than we think we do.

There is, finally, now an opportunity to do something about this sorry situation. It would take a wholehearted engagement from all the usual players—the World Bank, the U.N., NGOs, national statistical offices and academics, plus newcomers like mobile phone companies, opinion pollsters and companies that collect data for commercial purposes. Big funders would need to be involved, and also citizens themselves, to give their views on the indicators and the data that matter to them.

Improving data might seem like a geeky and somewhat marginal pursuit, compared with the weight of need and injustice that we know exists in the world. But, as Napoleon is alleged to have said, “War is 90 percent information,” and that applies to the war on poverty as much as to any military conflict. A data revolution might not be what most people think of as a real revolution. But it's sure to be revolutionary. □

*Dr. Claire Melamed is the head of the Growth, Poverty and Inequality program at the Overseas Development Institute, and leads the institute's work on the post-2015 global agenda. Prior to working at ODI, she worked for 10 years in a number of different U.K. development NGOs as well as for the United Nations in Mozambique. She also taught at the University of London and the Open University. Her recent ODI reports on the post-2015 development agenda can be found [here](#).*

*Photo: An employee of the Southern Sudan Commission for Census interviews residents of Juba, April 22, 2008 (U.N. photo by Tim McKulka).*

# COUNTING HEADS: THE GOOD, THE BAD AND THE UGLY OF DEMOGRAPHIC DATA

BY JACK A. GOLDSTONE



Let us start with a simple question: What was the total fertility rate (TFR) for a given developing country in a given year? The total fertility rate—or the number of children, on average, a woman in a given population is expected to have in her lifetime—is a crucial indicator, as it, in conjunction with mortality and migration, determines the future size, growth rate and age structure of a population. It is a much more important and informative measure than the crude birth rate, which is simply a snapshot of the total number of births in a population divided by the size of the population in a given year. The crude birth rate can vary quite a bit from year to year, and is affected by the age distribution of the population—a younger population will usually have a higher crude birth rate, because there are more fertile women in the population, even if the number of children most women will actually bear is falling. By contrast, the total fertility rate is calculated by finding the birth rate for women at each age and totaling up the number of children, on average, a woman will have as she moves through her entire span of childbearing years, assuming she survives that entire span.

To calculate the TFR, one simply needs to have census or survey data that provides a count or estimate of the number of women at each age in the society, and whether or not they gave birth that year. So it is a straightforward exercise in counting women and births. Or so it would seem.

Unfortunately, there is nothing straightforward or simple about it. Even if one had a census that reached every household, people tend to be inaccurate in giving their age; women in particular tend to take a year or so off their true age, or round down to the nearest half-decade or round number. Moreover, while TFR should include all live births, women sometimes do not report births if the baby has died in the interval between the start of the year and the date of the survey or census. Add to this the normal problem of taking a census or doing a survey in less-developed societies: The less common that formal employment and permanent residence are within the society, the more difficulty there is in ensuring that a census or survey reaches or is representative of the entire population. So the basic numbers of women in each age group and their births are subject to miscounting.

The result is that even for a basic demographic indicator like TFR for a given year—which is usually based on annual or biannual demographic surveys, as censuses often take place only at intervals of a decade or more—it is difficult to pin down the “true” rate. Take, for example, Ethiopia, where the reported level of total fertility ranges from 4.59, according to the United Nations Population Division, to 6.07, according to the CIA World Factbook, with the U.S. Census giving an “in-between” number of 5.55. For demographic data, this is an enormous range. It means a difference of more than 30 percent in the estimated size of the next generation in Ethiopia.

But these discrepancies are fairly common for low-income nations with weak formal bureaucratic administrations. For Nigeria, the discrepancy among data sources with regard to fertility is also more than one child per woman—but in the opposite direction, with the U.N. giving a TFR far higher than that of the CIA. So the differences are clearly not due to a systematic bias always pushing in the same direction.

Nor are there systematic discrepancies between international and national sources. For Yemen, the U.N. and the CIA again differ considerably, while the U.S. Census Bureau reports exactly the same fertility as the CIA; but for the Democratic Republic of the Congo, the U.N. and CIA are close, while the U.S. Census Bureau differs from the CIA-reported fertility by 0.73, a difference of almost a whole child. For more-developed countries with well-established population surveys and well-qualified health and population services, the reported data is much more convergent. Yet for many of the fastest-growing and largest countries in the world, the available demographic data is remarkably uncertain and inconsistent.

These problems are compounded by the inevitable process of forecasting population changes into the future. Five years ago, the United Nations' fertility data for African countries seemed to indicate a steady decline in fertility; assuming this decline would continue, the U.N. population division offered a "medium" projection for Africa's 2050 population of just less than 2 billion people. Last year, however, new data showed that the fertility decline had stalled. This forced the U.N. demographers to modify their medium projection for Africa's 2050 population to 2.4 billion people; it now seems likely that the population will be 20 percent larger at mid-century than was thought just a few years ago.

In sum, on some very basic items, demographic data is fraught with uncertainty and errors. Our data is only as good as the ability of surveys and state and regional governments to accurately count their populations—and not just to count people, especially women, but also to accurately identify their ages and recent births. As long as a large part of the world's population is living in states that, for reasons of politics or capabilities, do not deliver accurate counts, our ability to obtain accurate data will be limited, as will be our ability to provide accurate projections for the future.

However, the situation is not entirely bad. Let me highlight two trends that we are confident are underway that are helpful as well as a few areas where we really do not know much at all.

## THE GOOD

By "good," I mean demographic trends for which we have good quality data, and that are favorable trends. Two such trends seem clear. The first is regarding mortality—all across the world, even in the poorest nations, death rates are going down, and life expectancy has been extended by decades over just the past half-century. Even in sub-Saharan Africa, life expectancy at birth for both sexes combined has risen by 20 years since 1950, rising from 36.2 years—a level common in medieval Europe—to 56.0 years. India has made even greater progress, with life expectancy rising by 30 years in the same period, from 36.2 to 66.3. China has enjoyed similar gains, with life expectancy rising from 44.6 in 1950 to 75.2 in 2010. The same trend is visible, from a higher starting point, in rich countries. Moreover, these gains have not stopped coming: Most rich countries still saw life expectancies rise by at least two years in just the past decade (2000-2010); if that continues, "normal" life expectancy in the richer nations will be approaching 100 by the end of the century.

The second trend regards fertility in Asia. In 1968, when Paul Ehrlich published "[The Population Bomb](#)," warning of runaway population growth, the main anxiety was about Asia. Latin America and Africa at that time had relatively small populations, especially relative to their potential arable land; the real concern was high population growth rates in the historically densely populated rice-farming regions of Asia. Yet in fact, due to a combination of government policies and voluntary family limitation, fertility has fallen rapidly all across Asia—from 5.76 in 1960 to 2.19 in 2010 for the region as a whole.

The result of this decline in fertility is that most Asian countries have enjoyed a "demographic dividend" in recent decades. That is, with sharply falling birth rates, the age structure of the population shifted to having a larger proportion in the prime working age of 20-59 and fewer dependent

children. Investments in feeding, clothing and schooling children thus stabilized or dropped while investments in job training and capital to employ the labor force could be readily increased. These changes have helped drive the economic boom of recent decades across Asia; the population bomb in Asia has thus largely been defused.

In addition, Latin America, the Middle East and North Africa are all following Asia's lead. All across the developing world, with the exception of sub-Saharan Africa, fertility is declining and populations should stabilize before the end of the century.

## THE BAD

By “bad” I do not mean bad demographic trends (more on those below). Rather, I mean demographic trends and issues for which we simply have bad data—inaccurate, incomplete or missing—so that we are handicapped in making sensible policy decisions.

Perhaps the most egregious use of bad data has been in the field of education. While education may not seem like part of demography, in fact, progress in education has been measured in most developing countries using a simple demographic metric: a count of the number of children enrolled in primary, secondary and tertiary education. However, as Lant Pritchett has documented in his powerful new book [“The Rebirth of Education: Schooling Ain't Learning,”](#) simply counting the number of children enrolled in schools is a wholly invalid measure of the number of people gaining a meaningful education. Far too many children are enrolled in schools where teachers do not show up, where physical facilities are inadequate, and where textbooks and other essential materials are lacking. Counting students as “enrolled” in primary school means nothing if they have not learned to read and to be competent in basic math by the time they complete sixth grade.

Similarly, counting students as “enrolled” in secondary and tertiary education is grossly misleading if those students are not gaining useful skills that will enable them to compete for jobs. Today, because of much higher fertility in developing countries, roughly 90 percent of children below age 15 worldwide—1.7 billion out of a total of 1.9 billion—are growing up in less-developed societies. The demographic data shows that more than 90 percent of those children are enrolled in or have completed primary school. Yet the unfortunate reality, as Pritchett documents, is that this does not mean that 90 percent of those children are getting or have completed a primary education. In fact, we have hardly any sound data on the educational accomplishment of billions of school children around the world. While we have excellent tests of learning for children in the more-developed nations of the OECD and China, these children make up less than a quarter of those who will form the world's labor force in 2040. Thus we are moving into the middle decades of this century largely in the dark about what will be the skill levels and work capacities of the majority of the global labor force.

A second major area in which we are handicapped is the lack of good quality demographic data from societies that have significant ethnic and religious divisions. In countries divided along these lines, censuses have been subject to political manipulation and abuse for many decades. In Nigeria, for example, the census administration deliberately decided not to request information on ethnicity and religious affiliation on the most recent census in 2006; therefore no such data exists. In the absence of such information, the data collected was simply a count of the population, collected and aggregated by state, then at the national level. However, many claim that the state-level data has been grossly manipulated to preserve a demographic advantage for the northern states over those of the south. Prior censuses showing the southern population overtaking that of the north had been dismissed by the government as faulty; more recent censuses have kept the northern advantage but shown suspiciously low population growth rates in the south, along with questionably high growth rates in the north. The demographic information we have for what seems to be the most populous and important country in Africa thus is very likely flawed; whether that can be remedied depends mainly on political will.

The same is true for many other countries in Africa, the Middle East and South Asia. We do not know the precise ethnic and religious makeup of most divided countries; all we have are disputed estimates and the politically motivated claims of various parties. Governments often publish “official” figures, but these are usually not grounded in widely accepted and validated census data. These distortions often make it difficult to accurately gain even more basic data—such as the total fertility rate in Ethiopia, to return to our earlier example—without a relatively high measure of uncertainty.

## AND THE UGLY

I should also note that there are demographic trends that are well established by good data, but which are worrisome. These “ugly” trends include the rapid aging of many countries due to extremely low fertility (1.6 or lower), and the huge youth surges underway in countries where fertility remains stuck at very high levels (5.0 or higher).

The two dominant demographic trends in rich countries are people living longer and having fewer children. This has been undermining crucial worker-to-retiree ratios. Modern pension plans operate by taxing workers to pay retirees’ pensions, and they work well with a ratio of workers to pensioners of at least 3-to-1. However, in Japan and Germany, the ratio of those aged 20-64 to those aged 65 and older has already dropped below this level: It is 2.9 in Germany and 2.6 in Japan. Moreover, it will fall even further, in Germany to 1.9 by 2030 and just 1.5 in 2050; in Japan to 1.7 by 2030 and 1.3 in 2050.

This is not simply a one-off result of baby-boomer retirements. Because people are living longer and having fewer children, we are moving to a new long-term equilibrium population age structure. If people adopt a permanent level of 2.0 children per family—or even far fewer, as in Germany and Japan—eventually the age pyramid will become a vertical line, with the same number of people in every age bracket. And by 2100, life expectancy in most rich countries is expected to reach 100. At that point, the ratio of those aged 20-64 to those aged 65 and older will stabilize at 1.3 for all such societies. Even if people retired at age 70, the worker-to-retiree ratio would still be only 1.66, making current systems of pension provision unsustainable. In sum, rich countries will need to adopt a new model of retirement with new ways to finance pensions before the century is out.

A second worrisome trend is the opposite tendency in the countries of sub-Saharan Africa, where fertility remains quite high. In the 1980s and 1990s, fertility in Africa started to fall, dropping from 7 or 8 children per mother to 5 or 6. It was widely assumed this trend would continue and accelerate, leading to fertility declines like those seen in Asia and Latin America.

Yet in fact, while there is still some uncertainty about the rate of change, it is clear that hardly anywhere in tropical Africa have national fertility levels fallen below 5. For the less-developed regions of the world as a whole, the U.N. estimates that TFR fell from 5.93 in 1970 to 2.69 in 2010, a decline of almost 3 children per family; while in sub-Saharan Africa, TFR over the same years declined only from 6.66 to 5.39, or about 1 child per family.

These stubbornly high fertility rates in tropical Africa will produce stunningly high growth rates—close to 3 percent per year for the region as a whole, with rates closer to 4 percent in some countries, such as Niger. If these fertility rates continue on the track they have been on in recent decades, the projected population totals for many African states by mid-century are alarming: Ethiopia at nearly 200 million; Nigeria at substantially more than 400 million; Uganda and Tanzania at more than 100 million each. Sub-Saharan Africa as a whole could well grow from around 900 million people today to more than 2 billion by 2050, with more than half the population still under age 25.

While these population totals will tax these countries’ ability to provide food and jobs, a greater concern is their political impact. Much recent research has shown that countries with surging

youth populations have a much more difficult time sustaining political stability and transitioning to democracy than countries with higher mean ages. This is particularly true when the country is reaching a point of rapid transition of the population from mostly rural to substantially (30 percent or more) urban. We have recently seen this phenomenon play out in northern Africa and the Middle East; it is now increasingly a risk in tropical Africa, especially given the educational deficits noted above.

## CONCLUSION

Overall, demographic data is perhaps the most useful data available in the social sciences for making medium-term social projections. After all, we know how many people will be in their 20s, 30s, 40s and older 20 years from now—all of them have already been born, and we only need to count them and adjust for expected mortality to obtain highly accurate estimates. Knowing how many people are living in cities and going to school is also incredibly useful, as long as we realize that we also need to know whether they are finding jobs and getting valuable educations before we can interpret what those numbers imply.

Unfortunately, there is still much demographic data that we need to know but cannot know—what is happening to fertility in rural villages in Africa that are hard to reach and whose populations are hard to count accurately; what shifts are occurring in ethnic and religious identification in contentious countries; and whether the extremely depressed levels of fertility seen in some rich countries, such as Japan, South Korea and Germany, where women have on average fewer than 1.5 children each, will remain the norm—in which case pension financing systems will crash even sooner—or whether fertility will rebound soon to the 2.0 level required to sustain a constant-sized population.

In sum, political scientists and policymakers should always consult demographers when planning for the future, but they should be aware that counting alone is not knowledge, and that projections are just that—not certain predictions. Demographic data is embedded in social systems, and thus for better or worse is embedded in politics as well. □

*Jack A. Goldstone is Hazel Professor of Public Policy at George Mason University, and director of the Research Laboratory on Political Demography and Social Macro-Dynamics at the Presidential Russian Academy of National Economy and Public Administration, Moscow. His book “Political Demography” was published in 2012 by Oxford University Press.*

*Photo: Primary school class, Harar, Ethiopia, June 12, 2003 (U.N. photo by Eskinder Debebe).*

# GOING MISSING: FILLING THE GAPS IN AFRICAN GDP DATA

BY MORTEN JERVEN



On Nov. 5, 2010, Ghana Statistical Service, the country's government statistics office, announced that it was revising its gross domestic product (GDP) estimates upward, and as a result Ghana's GDP per capita almost doubled. The country was upgraded in an instant from a low- to lower-middle-income country. A sense of bewilderment and confusion arose in the development community. When did Ghana really become a middle-income country? What about comparisons with other countries? Shanta Devarajan, the World Bank's chief economist for Africa, struck a dramatic tone in an address to a conference organized by Statistics South Africa, calling the state of data collection on the continent "[Africa's statistical tragedy](#)."

The real tragedy was that we did not know how little we knew about income and growth in Africa. Observers and analysts have for too long taken the numbers at face value, but there is a large gap between the economic realities on the African continent and the numbers that purport to describe it.

Since the publication of my book "[Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It](#)," a healthy debate on the meaning of development statistics in the African context has taken place. Predictably, the emphasis in the media has been on the politics of the numbers. Inspired by the old phrase "lies, damned lies and statistics," commentators and headlines have focused on the dark forces that tamper with numbers and consciously mislead the public discourse on development Africa.

My main message is much simpler: We know less than we would like to think we do about growth and development in Africa based on the official numbers. The problem starts with the basic input: information. The fact of the matter is that the great majority of economic transactions in the rural agricultural sector and in medium- and small-scale urban businesses go unrecorded. Ghana's huge jump in GDP in 2010 is a symptom of this gap between what we think we know and what we actually know about income and growth in Africa.

## WHAT HAPPENED IN GHANA?

How could a country like Ghana be among the poorest in the world one day, and find itself among aspiring middle-income countries the next? To answer this question, one needs to understand some basic facts about national accounting, which yields the aggregate we refer to as GDP. GDP is calculated as a sum of the "value added" in the production of goods and services. The reality is that all GDP measures are an approximation, but in most African economies the statistical offices simply do not have all the information, time or resources needed to generate a new aggregate each year, and there is only very limited data available. Usually, the statistical office picks a "benchmark year" when it has more information on the economy than is normally available, such as data from a household, agricultural or industrial survey. The information from these survey instruments is added to other administrative data to form a new GDP estimate. This total is then weighted by sectors, and thereafter other indicators and proxies are used to calculate or guess at new annual figures. One variation of the usual method is to assume that food production grows in line with rural population growth, that the informal and unrecorded urban sector grows in line with the recorded service sectors and that construction

grows in line with cement imports. The estimates are made using a bit of qualified guesswork and brave assumptions.

This means that the benchmark estimation is important. Sectors that were important in the benchmark year will continue to appear important in following years, even if structural changes have occurred subsequently, while sectors that were unimportant or didn't exist in the benchmark year will barely have an impact on GDP. The data sources and the use of proxies are also set in the benchmark year, so even when new sources of information become available, national accountants may be unable to add them. When the benchmark year is out of date, the GDP series becomes unreliable. The IMF statistical division recommends a change of base year every five years—in Ghana's case, the benchmark year prior to the 2010 revision had been 1993. Quite obviously, the structure of Ghana's economy has changed radically since then. In 2010, the GDP of Ghana was recalculated using new data sources using 2006 as a benchmark, which is what caused the big jump in GDP. It turned out that since 1993, almost half of the Ghanaian economy had gone missing from the official count.

Now, Ghana is one of most closely watched economies on the African continent, yet the result of the revision took most if not all observers by great surprise. What then should we make of the reported numbers from other African economies?

One country whose numbers warrant consideration is Nigeria. Like Ghana, Nigeria has recently announced a political goal of reaching middle-income status. Since the fall of 2011, news of an upcoming revision of GDP statistics underway in Abuja has circulated widely, but the revision is not yet complete. It has also been reported that the revision may lead to a doubling of per capita income in Nigeria. The current benchmark year for the national accounts in Nigeria is 1990, meaning Nigeria's GDP estimates are even more outdated than Ghana's were prior to that country's revision, and therefore it is very likely that Nigeria's upward revisions will be as large as expected.

These GDP revisions are of game-changing proportions, and have direct implications for what we think of African economic development. The revision in Ghana certainly invites a rethink about the relationship between numbers and reality, but the Nigerian revision promises to change the economic picture of sub-Saharan Africa quite significantly; indeed, it will amount to a redrawing of the economic map of the region. According to the most recent World Development Indicators, the total GDP of Nigeria in current U.S. dollars was more than \$200 billion in 2010. Nigerian GDP, according to the unrevised numbers, already accounts for 18 percent of the total GDP of sub-Saharan Africa, which was measured to be about \$1.2 trillion in 2011. Assuming Nigeria's revised GDP is double the current figure, as media reports and private communication from the International Monetary Fund (IMF) and Nigerian Bureau of Statistics indicate is likely, the revision by itself will mean that GDP for the whole region will increase by more than 15 percent. The value of the increase adds up to no less than 40 economies of Malawi's size. The knowledge that there are currently 40 "Malawis" unaccounted for in the Nigerian economy should raise a few eyebrows, and leaves statements about trends in growth and poverty on the African continent in deep uncertainty.

## HOW GOOD ARE THE NUMBERS?

My book presents a study of the production and use of African economic development statistics. I emphasize that this is not just a matter of technical accuracy—the arbitrariness of the quantification process produces observations with very large errors and high levels of uncertainty. At the same time, this "numbers game" has taken on a dangerously misleading air of accuracy, and the resulting figures are used to make critical decisions that allocate scarce resources. International development actors are making judgments based on erroneous statistics. Governments are not able to make informed decisions because existing data are too weak or the data they need do not exist.

In "Poor Numbers," I surveyed the status of GDP statistics in sub-Saharan Africa, and in particular collected information on the methods and data used to compile national accounts. Of 37 countries

I surveyed in 2011, I showed that only 10 countries had a benchmark year that was less than 10 years old. I further showed that seven countries had a base year that was more than two decades old, and that there were only 6 countries that followed the IMF's advice to have a benchmark year that was 5 years old or newer, meaning 2006 at the time.

In response to this survey, the African Development Bank (AfDB) [commissioned a study](#), published in June 2013, that provided information on the same variables. The AfDB attempted to get a response from all 54 member countries, and received a response from 44 of them. In the survey of base years, the AfDB reports results from 34 countries, of which 9 countries met the five-year rule with a base year of 2007 or later. Another report that set out to replicate my study of GDP statistics in "Poor Numbers" was published in the [IMF's 2013 Regional Economic Outlook for Sub-Saharan Africa](#) in May. According to that survey of 45 countries, only four countries met the so-called five-year rule.

It is symptomatic of the state of knowledge on African statistics that two key players like the IMF and African Development Bank published conflicting metadata on African statistics within a month of one another. Despite these discrepancies and disagreements on the number of very recently updated GDP estimates, both reports confirm that many countries use very outdated base years in calculating their GDP statistics. The African Development Bank reports that 19 countries have base years older than 10 years old, including eight with base years more than 20 years old. In the IMF's larger sample, one finds 28 countries with base years more than 10 years old, while 13 countries are still using base years more than 20 years old.

But metadata such as benchmark years is just scratching the surface. A country's GDP estimates are only as good as the data on which they are based. As the African Development Bank reports, there is a lack of data on industrial production, and very few countries conduct regular surveys of agriculture. And while half of African countries have undertaken surveys focusing on the informal sector over the past decade, regular annual data is still based on guesses and very partial information on the economy.

### **STATISTICAL CAPACITY IN AFRICA: RISE AND FALL AND RISE?**

In "Poor Numbers," I describe how the statistical capacity of African states was greatly expanded in the late colonial and early postcolonial period, but was greatly impaired during the economic crisis of the 1970s. Statistical offices were neglected in the decades of liberal policy reform—the period of "structural adjustment" in the 1980s and 1990s— that followed. In retrospect it may be puzzling that the IMF and the World Bank embarked on growth-oriented reforms without ensuring that there were reasonable baseline estimates that could plausibly establish whether the economies were actually growing or stagnating. For statistical offices, structural adjustment meant having to account for more with less: Informal and unrecorded markets were growing, while public spending was curtailed. As a result, our knowledge about the economic effects of structural adjustment is limited. More generally, the economic growth time series—or the cumulative record of annual growth—between 1960 and today for African economies does not appropriately capture changes in economic development.

First, the decline in economic growth in Africa in the 1980s was overstated. Second, for many economies, such as Tanzania and Zambia, the upward swing in the 1990s was overstated. The marked improvement we see in the GDP time series in the mid-1990s was driven by expanding the estimates for the informal sector—thus it was statistical growth, not real growth.

Third, and similarly, a lot of the recent rapid growth we are now recording in Africa is in fact statistical growth deriving from adding the informal sector and the service sector to the old estimates. Certainly, much of the recent apparent growth derives from appropriately recorded growth in external trade, but exactly how this growth relates to the domestic economy, and more generally to economic development such as poverty reduction, remains guesswork.

The contrast between the statistical offices and the central banks in the region is striking. While statistical offices are located in rundown offices, often with limited computer facilities, the central banks of African countries are located in new high-rise buildings with modern facilities. Positions in the central banks command higher salaries and prestige than those in the statistical offices, and central bank employees are in a better position both symbolically and physically to provide timely and useful advice to policymakers. A similar lift is needed for statistical offices.

The current development agenda is set by the Millennium Development Goals of the United Nations. This has led to some statistical capacity-building in a number of countries, while in others there have been perverse effects when statistical capacity is diverted to data collection that serves the monitoring of particular donor targets. The Millennium Development Goals have generally meant that there has been more data available for measurement of social development, but the data needed for economic governance is either not supplied or of very questionable quality.

It is of course true that there has been a growth in output of numbers from statistical offices in response to the growing demand. But the progress has been uneven. There has been a clear shift in priority away from the collection of some of the basic data needed in the compilation of national accounts, and [a shift toward social statistics](#).

Moreover, the funds that are made available to statistical offices are generally ad hoc funds that support data collection for a donor-funded project. In practice, many statistical offices operate as data-collection agencies for hire, not as offices that provide objective information needed for day-to-day politics or policy planning. This means that donors distort data production rather than expanding statistical capacity. While resources for manpower and infrastructure are stretched thin. The problem here is lack of coordination: Many countries have national strategies for improving their statistical capabilities, but, more often than not, donors break with these plans' priorities and demand the data they need, thus adding to the fragility of statistical offices under increasing pressure.

## WHAT TO DO ABOUT IT?

Since the publication of “Poor Numbers” the debate has continued, and many useful and practical steps toward the production of more reliable data have been suggested. But recent events have also shown that an open and transparent debate on the reliability of statistics may be difficult to conduct.

Discussing economic statistics and GDP estimates of African economies is clearly important, but it's also sensitive. [In Bill Gates' review](#) of “Poor Numbers,” he wrote, “It is clear to me that we need to devote greater resources to getting basic GDP numbers right. As Jerven argues, national statistics offices across Africa need more support so that they can obtain and report timelier and more accurate data.” At face value, this should be wonderful news for statistical offices, but that was not how it was initially received. Some members of the African statistical community [responded by trying to block me from attending conferences](#) and [trying to control or stifle the debate](#) on African statistics. Now that some of the dust has settled, [there are good indications](#) that partners will unite on a common agenda for improving the data needed for economic governance in Africa.

Improving the data collection in Africa is not only a question of the amount of funding. It requires more than simply increased resources—the main problem is the incentives and the political economy surrounding the provision of statistics as well as the global governance of the demand for data. There are a lot of diverging interests among international organizations, and sometimes gaining support for one organization's particular program or agenda comes before actually doing the job of improving measurements.

Those who thought that there was a data revolution taking place, with the world united toward achieving a better measurement of development, might be disappointed. Lack of firm facts leaves

convenient room for negotiation of the numbers when it is needed. Currently the data, if they are available, are not timely or of the quality needed, and that hampers any serious agenda for economic governance or development planning. But as anyone who has read “Poor Numbers” will know, the big story in the book is governance by ignorance.

As I have pointed out, this stands in striking contrast to the demand for data in the development community. The most extreme version of “evidence-based policy” comes from those who [suggest that we tie financial rewards](#) directly to statistical evidence of success. The trend is that donors are increasingly demanding monitoring and data in return for funding. With the huge gaps in data combined with clear incentives to distort the results, the risk is that the outcome will be “policy-driven evidence” rather than evidence-based policy.

At present there is no coherent global strategy for improving the provision of data for development. This is related to the general problem of accountability in development. Sometimes ignorance is bliss, for both the donor community and local political leaders. Putting a coherent global strategy for statistical capacity building in place is important, and such a standard must be geared toward solving local problems. For the business community and investment banks, the availability of reliable statistics [is crucial for future growth](#) in Africa.

It is indeed a real tragedy that the statistical capacities of sub-Saharan African economies are in such a poor state, but it is a reality: African development statistics tell us less than we would like to think they do about income, poverty and growth in sub-Saharan Africa, even as governments, international organizations and independent analysts need these development statistics to track and monitor efforts at improving living conditions on the African continent. One of the most urgent challenges in African economic development is to devise a strategy for improving statistical capacity, as the current system causes more confusion than enlightenment. □

*Morten Jerven is associate professor at Simon Fraser University in Vancouver, Canada. An economic historian, he has published widely on patterns of African economic development, including a recent book, “Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It,” published by Cornell University Press in 2013. Jerven’s second book, “Economic Growth and Measurement Reconsidered in Botswana, Kenya, Tanzania, and Zambia, 1965-1995,” will be published by Oxford University Press in 2014.*

*Photo: Freedom Square, Accra, Ghana, June 26, 2011 (photo by Wikimedia user ryansworld licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license).*

**Want more? [Subscribe to World Politics Review](#) now to get unlimited access.**

# WORLD POLITICS REVIEW

<http://www.worldpoliticsreview.com>

**WORLD POLITICS REVIEW** was created for foreign policy professionals who have a world of online breaking news at their fingertips.

Our tightly focused coverage gives context to the key trends and events in international affairs. Follow our front-page headlines to keep ahead of current events. Explore our extensive database to understand the trends behind the news.

Published entirely online, WPR's original analysis is written by a network of leading analysts, journalists and scholars. Our nonpartisan, nonideological articles cover diplomacy, military affairs, energy, economics and other subjects that make geopolitics tick. WPR's accessibly written analysis speaks to engaged readers who need to know what's driving policy in Beijing, Brasilia and New Delhi as well as in Washington, London and Brussels.

[Subscribe now](#) or learn more at [www.worldpoliticsreview.com](http://www.worldpoliticsreview.com).

## The WPR Service

- Our daily Media Roundup links to the stories you must follow to stay informed. Online and in your inbox.
- Our daily Leading Indicators service highlights emerging stories before they become mainstream news. Online and in your inbox.
- Daily Briefings and Columns written by our global network of expert contributors.
- In-depth, full-length reports provide comprehensive examinations of single subjects. These include our Features (24 a year), Special Reports (12 a year) and Strategic Posture Reviews (six a year).
- Daily Trend Lines items provide quick-hitting expert viewpoints on important issues of the moment, including our Global Insider interviews.
- WPR's Media Roundup search allows subscribers to search for individual topics across dozens of international news sources.
- Our extensive archive contains more than 4,000 (and growing) pieces of original analysis.

## Questions?

Contact us at [subscriptions@worldpoliticsreview.com](mailto:subscriptions@worldpoliticsreview.com) or +1.202.596.9771.

Corporations, schools, non-profit organizations, government agencies and other institutions can get site-wide access. Get more information about our institutional subscriptions at [about.worldpoliticsreview.com/wpr-ebSCO](http://about.worldpoliticsreview.com/wpr-ebSCO)