Data poverty makes it harder to fix real poverty. That’s why the UN should push countries to gather and share data.

By Rohini Pande and Florian Blum  July 20

In September 2000, all United Nations member states agreed to adopt the Millennium Development Goals, eight targets to guide global development until 2015. It was an unprecedented global consensus about how to solve the world’s biggest problems.

Those goals expire this year. The final report tells of tremendous progress on some targets. Extreme poverty has more than halved, the gender gap in primary education has largely disappeared, and the spread of HIV/AIDS has slowed.

But critics have long pointed out that the data on MDG progress is patchy and impossible to corroborate. World Bank researchers report that only 77 of the 155 countries they studied collect reliable data on poverty. The UN itself issued a 2014 report stating that data deprivation “can lead to the denial of basic rights, and for the planet, to continued environmental degradation,” and calling for a “data revolution.”

As researchers working in South Asia – a region that remains home to the highest number of people living in extreme poverty – we have witnessed the pervasive effects of data deprivation on the ground. For instance, we recently asked tax administrators in Pakistan about work challenges, assuming that underreporting would be the crux of their woes, as is usually the case. Instead these administrators described the hardest challenge as identifying how much tax a citizen has already paid. Absent usable data systems, they cannot verify sums collected and so require taxpayers to document taxes paid – a daunting task in a system dominated by cash payments and withholding schemes. Since most taxpayers fail to provide such evidence, the tax administration is unable to undertake basic tasks such as providing refunds, and therefore cannot implement a fair and effective tax system.

The UN is currently creating a list of Sustainable Development Goals (SDGs), intended to take the place of the MDG, which will shape how governments and NGOs allocate an estimated $2.5 trillion of aid over the next 15 years. The UN released a preliminary draft on June 2 and will finalize the SDGs in September. Goal 17 vaguely addresses data deprivation, by stating an aim to build on existing initiatives and “support statistical capacity building in developing countries.”
But a country’s capacity to produce and use statistics does not only require investment in the infrastructure needed to collect, collate, and open up their administrative data to the public, particularly researchers. It also depends on political economy considerations, which thus far have been left out of the conversation.

The political economy of data collection

The centerpiece of the Millennium Development Goals was halving global poverty, which is usually measured by household surveys. These are costly to conduct, so historically, richer countries have richer data on poor people. World Bank data shows that between 1977 and 2012, high-income countries, on average, collected data on poverty every 2 years, 3.5 times as often as low-income countries. There are exceptions, India being one of them. (The horizontal axis maps countries by income quintile.)

But India’s data exceptionalism doesn’t extend to measuring whether its citizens are successfully investing in acquiring the tools to escape poverty—for instance, by getting an education. Arguably, this missing information has limited India’s ability to pull the right policy levers to reduce poverty more quickly.

A look at India’s history suggests that domestic politics influenced its decisions to collect data—or fail to do so.

In 1950, very soon after independence, India made the decision to invest in comprehensive household expenditure surveys. At the time, most Indians believed that their country had been impoverished by colonialism. That made documenting poverty a politically palatable choice. And once data was coming in regularly, enough groups had a stake in the Indian poverty debate that pulling back was not an option.

But in 2000, OECD started offering all countries a detailed assessment of educational attainment among school children called the Programme for International Student Assessment, or PISA. By then India could no longer blame former colonial powers if its children weren’t learning—which means it had something to lose by collecting that data. Still, it tested the waters in 2009 by having two high-performing states take part. When these showcase states were ranked 72 and 73 of regions tested – beating out only Kyrgyzstan – India exited PISA.

A keystone of the political economy of reform literature is Dani Rodrik and Raquel Fernandez’s finding that reforms that benefit the majority may fail if it is easier to identify the reform losers than the potential winners. Given Indian teachers’ and educational administrators’ knowledge of the dire reality, and the uncertainty surrounding the benefits of achievement tests for the majority of India’s schoolchildren, it is easy to understand why India opted out.

Yet, reforms do occur – often aided by data, not in spite of it. More recently, Rodrik argued in the Journal of Economic Perspectives that ideas pioneered by powerful policy entrepreneurs can sometimes outmaneuver vested interests.
In 2000, Brazil’s President Cardoso – who had campaigned to combat Brazil’s rising economic inequality, in part by initiating school reforms – saw PISA as an opportunity to justify his agenda. Brazil’s dismal initial PISA results helped Cardoso implement a comprehensive system of measuring education results and school quality. Over the next decade, Brazil’s PISA ranking improved dramatically, from second to last in science and last in math out of 40 countries in 2003 to 58th and 59th, respectively, out of 65 countries in 2012 – that’s the 78th and 80th percentile. In math, Brazil is the country with the largest performance gain since 2003.

Data and researchers: If you build it, will they come?

In a recent training session conducted under the aegis of our research group, Evidence for Policy Design’s collaboration with the UK’s Department for International Development (DFID)’s Building Capacity to Use Research Evidence program, we asked Indian administrators what held them back from making public data that has already been collected. The Indian government responded was that if researchers wanted to see the data, they could file a request under the Freedom of Information Act.

Over the last half-century, Freedom of Information Acts have spread from Northern Europe to more than 95 countries, rich and poor. That spread has granted citizens, researchers and the press access to a wide range of government records. We often hear bureaucrats in India and beyond make the argument that, given that access – as well as the abundance of household surveys and open data initiatives – they have little reason to invest limited government resources in collecting, collating and publishing data. Moreover, if you add the possibility that policymakers can suffer adverse consequences when researchers analyze their programs, then the impetus to resist reform becomes strong.

But a set of figures provided via a Freedom of Information request offers a keyhole through which a researcher can view a small slice of the data a government collects. That does not compare to giving researchers hands-on access to entire administrative datasets. Denmark facilitated access to de-identified social security data from one central hub in 2005, and extended data access in 2008. The associated increase in the number of papers written about Denmark has been dramatic.

This has mattered for how policy works for the Danish public. In a 2011 paper, Kleven et al. used administrative data to show that tax evasion was much higher among self-reporters compared with those whose earnings were recorded by third parties. This led the Danish government to implement a major tax reform, expanding third-party reporting to financial institutions. And that will likely lead to fairer implementation of the tax code and more money in the coffers to fund public programs. Closer to home, access to U.S. Internal Revenue Services administrative files recently led Harvard researchers to provide definitive evidence on the importance of neighborhood quality for lifelong opportunities enjoyed by Americans.
But even as the arc of history bends toward open data, we’ve seen countries move in the opposite direction – again, acting on political motivations. In March of this year, the Tanzanian parliament, facing an October election, introduced a law requiring any published data to be endorsed by the National Bureau of Statistics, essentially giving the government the ability to bury reports by the press and research by scholars. While it is hard to imagine this law gaining traction, the very fact that it passed suggests that some in power wish to keep a tight hold on data as elections approach. Furthermore, even among those wealthy countries whose administrative datasets could feed hugely beneficial research, many (e.g. Japan) have largely restricted access.

Governments respond to external as well as internal political pressure. In September, the United Nations should not just announce the final list of Sustainable Development Goals, but also create clear mechanisms to recognize and reward countries that build open data systems and spark research and public discourse, as we have seen Brazil and Denmark do. Countries that use their statistics agencies as clearinghouses should be lauded, while others that use theirs as censors should be shamed.

*Rohini Pande, a professor of public policy at Harvard Kennedy School, co-directs the Evidence for Policy Design Initiative (@EPoDHarvard). Florian Blum is an Economics PhD student at the London School of Economics and studies the determinants of state capacity in developing countries.*