The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics

Justin Sandefur and Amanda Glassman

Abstract

Across multiple African countries, discrepancies between administrative data and independent household surveys suggest official statistics systematically exaggerate development progress. We provide evidence for two distinct explanations of these discrepancies. First, governments misreport to foreign donors, as in the case of a results-based aid program rewarding reported vaccination rates. Second, national governments are themselves misled by frontline service providers, as in the case of primary education, where official enrollment numbers diverged from survey estimates after funding shifted from user fees to per pupil government grants. Both syndromes highlight the need for incentive compatibility between data systems and funding rules.

JEL Codes: C83, E31, I15, I25, I32

Keywords: Africa, national statistics systems, household surveys, administrative data, immunization, school enrollment, EMIS, HMIS

Center for Global Development
www.cgdev.org

Working Paper 373
July 2014

The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics

Justin Sandefur
Center for Global Development

Amanda Glassman
Center for Global Development

Thanks to Morten Jerven, Oliver Morrissey, two anonymous referees, and the members of CGD's Data for Africa Development working group (www.cgdev.org/page/data-african-development) for comments and suggestions. Denizhan Duran, Sarah Dykstra, and Yuna Sakuma provided excellent research assistance. All viewpoints and remaining errors are ours. The data and code underlying all tables and graphs can be found on the CGD page on the Harvard Dataverse Network http://www.cgdev.org/dataverse.

CGD is grateful for contributions from the Omidyar Network, the UK Department for International Development, and The William and Flora Hewlett Foundation in support of this work.

Justin Sandefur and Amanda Glassman. 2014. "The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics." CGD Working Paper 373. Washington, DC: Center for Global Development.
http://www.cgdev.org/publication/political-economy-bad-data-evidence-african-survey-administrative-statistics-working

Center for Global Development
2055 L Street, NW
Fifth Floor
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

1 Introduction

There is a growing consensus among international observers that official statistics in many Sub-Saharan African countries are woefully inadequate and unreliable (Jerven, 2013), what Devarajan (2013) calls a "statistical tragedy". In response to this tragedy, the U.N. High Level Panel on post-2015 development goals has called for a "data revolution" to improve tracking of economic and social indicators in Africa and the rest of the developing world (United Nations, 2013). The agenda emerging around these discussions has tended to assume that more money and better technology will solve the problem, focusing on an expansion of survey data collection efforts, and a push for national governments to disseminate information under open data protocols (Caeyers, Chalmers, and De Weerdt, 2012; Demombynes, 2012).

Do these solutions address the underlying causes of bad statistics? Relatively less attention has been paid to understanding why national statistics systems became so badly broken in the first place, or to analyzing the perverse incentives which any data revolution in Sub-Saharan Africa must overcome. We attempt to fill this gap by documenting systematic discrepancies between data sources on key development indicators across a large sample of countries. By necessity, we focus on cases where multiple independent sources report statistics on ostensibly comparable development indicators.[1] For this we draw on cross-national data within Africa on primary school enrollment and vaccination rates taken from the Demographic and Health Surveys (DHS), and contrast it with data from education and health management information systems maintained by line ministries.[2]

The core hypothesis to be tested in this paper is that misrepresentation of national statistics does not occur merely by accident or due to a lack of analytical capacity – at least not always – but rather that systematic biases in administrative data systems stem from the incentives of data producers to overstate development progress. The administrative data we analyze are designed to be part of management information systems in health and

[1] This approach has parallels to the approach used by Jerven (2010) to examine discrepancies in GDP data for several African countries. While in that case Jerven calculates margins of error based on disagreement between multiple sources – all of which contain some element of signal and noise – for social statistics we have access to a "gold standard" measure in the form of independent household surveys, and can study not only the magnitude but also the direction of the bias in official data.

[2] It is important to acknowledge that household surveys – which we treat as our gold standard to assess errors in administrative data – have their own imperfections, including both sampling and non-sampling errors discussed by other papers in this special issue. Because our focus here is on bureaucratic and political incentives that are unique to administrative data systems, household surveys remain a very useful benchmark. Our argument ultimately requires only that errors in household surveys are *independent* of, not necessarily smaller than, the errors in administrative data systems.

education ministries. It should be no surprise, then, that misrepresentations in the data reflect the incentives provided by the governance and funding structures of these ministries, particularly in the low-income, highly aid-dependent countries which dominate our sample.[3]

The paper is organized around two interlinked principal-agent problems: in the first, national governments can be seen as agents of international aid donors and domestic constituencies; in the second, governments act as principals seeking to motivate civil servants to simultaneously provide public services and report truthful data on the same.

In the first case, an international aid donor (the principal) seeks to allocate resources between and evaluate the performance of a national government (the agent). The principal requires data to monitor agents' performance. Recognizing the risks inherent in "self-reported" official statistics, international donors invest heavily in the collection of survey data on households, farms, and enterprises. Notably, these surveys involve considerable foreign technical assistance paid for directly by donors. In the extreme case of the DHS data sponsored by the U.S. Agency for International Development (USAID) and other partners – on which much of the analysis below relies – the donor insists on a standardized questionnaire format in all countries, donor consultants train the data collectors and oversee fieldwork, and all or most raw data is sent back to the donor country for analysis and report writing.[4]

Donors can't always rely on such carefully controlled data products like the DHS though, and Section 3 shows the predictable results when donors link explicit performance incentives to administrative data series managed by national governments. In 2000, the Global Alliance for Vaccines and Immunization (GAVI) offered eligible African countries a fixed payment per additional child immunized against DTP3, based on national administrative data systems. Building on earlier analysis by (Lim, Stein, Charrow, and Murray, 2008), we show evidence that this policy induced upward bias in the reported level of DTP3 coverage amounting to a 5% overestimate of coverage rates across 41 African countries.

In short, pay-for-performance incentives by a donor directly undermined the integrity of administrative data systems. To invert the common concern with incentive schemes, "what

[3]It should be noted that our statistical methodology lends itself to focusing on averages, e.g., how the mean discrepancy between administrative and household data changed after a certain policy change. This masks considerable heterogeneity across countries that is difficult to study econometrically in samples of this size, but should not be taken to imply that Africa's "statistical tragedy" is a uniform phenomenon.

[4]While we focus on the role of external aid donors, national governments are ideally accountable first and foremost to national citizens. Statistical literacy in the region is low, and the role of economic data in public discourse is limited, but there are exceptions to this rule, particularly for the highly visible phenomenon of rising consumer prices. Sandefur (2013) illustrates the potential pitfalls of politically salient data series with the case of the measurement of inflation in consumer prices in a panel of African countries.

gets measured gets managed," in the case of statistics it appears that what gets managed gets systematically mis-measured, particularly where few checks and balances are in place.

In the case of immunization statistics, national governments mislead international donors and their citizens, whether by accident or design. Previous analysis of African statistics has focused on this dynamic in which central governments are the producers of unreliable statistics (Jerven, 2011). But in other cases national governments themselves are systematically misled, creating an important obstacle to domestic evidence-based policymaking.

In this second accountability relationship discussed in Section 4, national governments and line ministries (the principal) seek to allocate resources between and evaluate the performance of public servants such as nurses and teachers (the agents). By and large, the information the principal relies on in such settings comes from administrative data systems based on self-reports by the very agents being managed. The result is systematic misreporting, undermining the state's ability to manage public services, particularly in remote rural areas.

Section 4 illustrates this problem in primary school enrollment statistics. Comparing administrative and survey data across 46 surveys in 21 African countries, we find a bias toward over-reporting enrollment growth in administrative data. The average change in enrollment is roughly one-third higher (3.1 percentage points) in administrative than survey data – an optimistic bias which is completely absent in data outside Africa. Delving into the data from two of the worst offenders – Kenya and Rwanda – shows that the divergence of administrative and survey data series was concomitant with the shift from bottom-up finance of education via user fees to top-down finance through *per pupil* central government grants. This highlights the interdependence of public finance systems and the integrity of administrative data systems. Difference-in-differences regressions on the full sample confirm that the gap between administrative and survey of just 2.4 percentage points before countries abolished user fees grew significantly by roughly 10 percentage points afterward.

This dual framework relating the reliability of statistics to the accountability relationships between donors (and citizens), national governments and frontline service providers clearly abstracts from certain nuances, as does any useful model. Household survey data are not *only* used by international donors as a tool to monitor aid recipients. Donor agencies also use survey data for research purposes, and recipient governments frequently cite survey reports in planning documents and incorporate survey data into the construction of macroeconomic aggregates like GDP which are key indicators in domestic policymaking.

Conversely, international donors are far from apathetic about administrative data systems, and indeed invest heavily in education and health management information systems in the region. Nevertheless, we believe the political economy dynamics suggested by this framework, however simplistic, help make some sense of the seemingly chaotic data discrepancies documented in the paper.

Seen through the lens of this framework, the agenda for a data revolution in African economic and social statistics clearly must extend beyond simply conducting more household surveys to help donors circumvent inaccurate official statistics – to avoid, as we label it, being fooled by the state. If donors are genuinely interested in promoting an evidence-based policymaking process, they must assist government to avoid being systematically fooled itself by administrative data systems built on perverse incentives. Aid resources must be directed in a way that is complementary to, rather than a substitute for, statistical systems that serve governments' needs.

2 Seeing like a donor versus seeing like a state

The different needs of donors and government present trade-offs between the comparability, size, scope, and frequency of data collection. Given that donors finance a large share of spending on statistics, these differing needs can imply that national statistical systems arent built to produce accurate data disaggregated for use by domestic policymakers and citizens. In stylized form, this creates a choice between (i) small-sample, technically sophisticated, possibly multi-sector, infrequent surveys designed to facilitate sophisticated research and comparisons with other countries[5], and (ii) large sample surveys or administrative data sets providing regional or district level statistics on relatively fewer key indicators at higher frequency, designed for comparisons across time and space within a single country.[6]

International aid donors must make allocation decisions across countries, and in many cases they are bound to work solely with national governments as their clients. Due to

[5]In the first category, several internationally comparable survey programs are very active in sub-Saharan Africa since 2000; the Demographic and Health Surveys (DHS), LSMS, Global Findex and Multiple Indicator Cluster Surveys (MICS) are main examples.

[6]We ignore, for present purposes, a third category which is likely quite large but not directly relevant to our analysis here: one-off survey projects that ignore both time-series and cross-country comparisons, and often produce redundant data due to a lack of coordination between donors or government agencies. At times such one-off, bespoke surveys may be fully justified for, say, specific project evaluation purposes, but again this falls outside our scope here.

this focus, donors' preferences often (but far from always) skew toward statistics based on standardized international methodologies and homogenized questionnaire formats.[7] At times this desire for international comparability is directly at odds with comparability over time within a single country. A second key implication of donors' concern with international comparisons is less attention to domestic, sub-national comparisons. Household survey data reflects this preference.

Consider the case of primary education in Kenya. The DHS provides comparable information across countries and time on both health and schooling outcomes and is designed to provide provincial estimates, with most analysis focusing on a single set of national or rural and urban statistics. At the time of the last DHS, Kenya had eight provinces. This allowed at least limited correspondence between survey estimates and units of political accountability. In neighboring Tanzania, the mainland was at the time of the last survey divided into twenty-one regions, but the survey reported results for just seven aggregate zones corresponding to no known political division. To stress the point, we might say that the structure of the DHS meets the needs of a donor sitting in Washington, allowing them to evaluate, say, Kenyan and Tanzanian performance in primary schooling on a comparable basis.

But national governments need to make sub-national resource allocation decisions. To be useful, data is often required at relatively low levels of disaggregation that coincide with units of political accountability. Ministries of Health require up-to-date information on clinics' stocks to manage their supply chain; Ministries of Education require school-level enrollment statistics to make efficient staffing decisions. In Kenya, the education ministry obtains this information from the country's Education Management Information System (EMIS), three times a year, for all twenty-thousand government schools in the country.[8]

[7]Note again that this is an aspiration for some, not all, donors and often not the reality. An independent evaluation of the Integrated Household Survey Network (Thomson, Eele, and Schmieding, 2013) notes that "IHSN has not been effective in improving the coordination of surveys" and "[s]ome international organisations have committed considerable amounts of their own resources to developing survey formats and questions which are not always mutually compatible." Furthermore, "[t]he duplication of both actors and products in the data management field as well as feedback from countries obtained during the evaluation indicate that effective coordination remains a serious challenge." We would endorse the conclusion of the evaluation, i.e., "[w]ith resources for global public goods in statistics often in very limited supply, what is done is largely driven by agency-specific budgets and decision-making processes."

[8]Although we argue that household surveys can be of limited utility to domestic policymakers and citizens, we recognize that international standards and comparability also serve these groups - where standards don't already exist, or as a way to enhance quality of estimates and shield statistical activities from political interference, or to provide a framework for comparing, assessing and learning experiences globally. Indeed, household surveys also allow us to assess the quality of more disaggregated administrative data that is the main subject of this paper.

Arguably, citizens' interests are better-aligned with the preferences of their own governments EMIS than donors in this case. In order for citizens to exert bottom-up accountability of public service providers, they require data in an extremely disaggregated form. Kenya's national trends in literacy rates are likely of limited interest to citizens of rural Wajir, but the local primary school's performance on national exams relative to neighboring villages may be of acute interest. Thus, appropriately, the Kenyan government's "open data portal" provides access to the disaggregated administrative data EMIS system. Unfortunately, as we show in Section 4, the reliability of this data is questionable.

Kenya may not be unique in this respect, as we show, but the situation is far from uniform across the region. While quality measures in statistics are few and far between, and indeed measuring discrepancies is a key contribution of this paper, the World Bank's Bulletin Board of Statistical Capacity provides some indication.[9] On average, Sub-Saharan Africa scores below all other regions, with an overall score of 58 compared to a global average of 64. But the variance within Africa is enormous, ranging from the very bottom of the rankings (Somalia with a score of 24) to Malawi with a score of 79 (just above the 75th percentile globally). These numbers should be treated with caution though: while it is no surprise to see Mauritius score highly (76), it is perhaps surprising to see it tied with Nigeria, where Jerven (2013) documents considerable challenges in data reliability.

Viewed from the perspective of national governments' data needs for policy planning and management purposes, household surveys perform poorly in terms of geographic representativeness and frequency. Many surveys provide only national estimates, offering little guidance to domestic policymakers allocating resources and attention between sub-national units. Few surveys are able to provide statistics at the district or equivalent level, and many are unable to provide even regional or provincial estimates. Furthermore, surveys of household income, health outcomes, agricultural production, and other key indicators typically occur only once every several years, often with long lags between data collection and the publication of statistics.

The overwhelming strength of household surveys in Sub-Saharan Africa, however, is that they provide information that is likely to be more reliable as it is better documented and collected with much higher levels of international technical assistance and oversight. While this section has focused on outlining the theoretical advantages of administrative data sources, the following two sections turn to the core task of the paper, documenting the deep

[9] http://data.worldbank.org/data-catalog/bulletin-board-on-statistical-capacity

flaws with administrative data systems in practice and diagnosing the causes of these ills. In the conclusion we turn to the question of whether administrative and survey data could be better integrated for purposes of cross-validation to prevent or correct the discrepancies we document.

3 Fooled by the state: Immunization rates across 41 countries

In this section we turn to role of national statistics in holding national governments accountable to international donors, as well as their own citizens. Relative to other government functions, the production of national statistics in sub-Saharan Africa is highly dependent on foreign aid. [10] Donors demand statistics for a variety of purposes including, but not limited to, the allocation of aid resources across countries, and the evaluation of specific programs as well as recipient governments' overall economic management. We study these dynamics in the case of an explicit donor incentive scheme to promote vaccinations.

The health sector provides an important case study of the tension between more reliable, smaller sample, less frequent survey data and the high-frequency administrative data with limited quality controls which purports to cover the entirety of the population. Like EMIS databases in education examined below, many countries' health management information systems (HMIS) databases rely on self-reported information from clinic and hospital staff, which aggregated up by district and regional health officers, each with potentially perverse incentives.

There are a number of reasons why HMIS and survey sources may disagree. Numerators in administrative data can be inaccurate due to incomplete reporting, reporting on doses distributed rather than administered, repeat vaccination or omission of the private sector and non-governmental organizations. Denominators can be inaccurate due to migration, inaccurate or outdated census estimates or projections, inaccurate or incomplete vital registration systems, among others. Indeed, Brown (2012) notes that denominators are frequently estimated by program managers in each country for the WHO's Expanded Program on Immunization, based on counts or estimates by local program staff or health

[10] McQueston (2013) shows that in Ethiopia (2007-2008) and Malawi (2007-2011) over 80% of funding for the national statistics office came from direct donor support, while national statistics offices in Tanzania (2008-2014) and Kenya (2008-2009) received 36% and 54% from aid donors, respectively.

workers. Finally, in countries where immunization card distribution, retention and utilization are suboptimal and mothers report vaccination coverage from memory to enumerators, survey-based coverage estimates can also be biased, particularly for multi-dose vaccines that can be under-reported (WHO and UNICEF, 2012).

One additional layer of perverse incentives for accurate reporting in the health context is the policy of aid donors to link payments to progress on HMIS indicators. Starting in 2000, the Global Alliance on Vaccines and Immunizations (GAVI) offered low-income countries cash incentives for every additional child immunized with the third dose of the vaccine against diphtheria, tetanus, and pertussis (abbreviated as DTP3) based on HMIS reports. Lim, Stein, Charrow, and Murray (2008) compare survey-based DTP3 immunization rates and their growth over time with HMIS or administrative rates reported to the WHO and UNICEF, finding that survey-based coverage has increased more slowly or not at all when compared to administrative reports.

We extend this analysis with a particular focus on establishing the causal role of the GAVI performance incentives on misreporting in a sample of African countries. Our primary innovation is the use of "placebo" regressions that test for an effect of GAVI on a vaccination (i.e., measles) for which GAVI created no incentive to misreport. This allows us to construct a more plausible counter-factual for what would have happened in the absence of the GAVI ISS scheme, and thus go beyond noting the growing discrepancies as in past analyses, to identifying the impact of GAVI through the creation of perverse incentives. [11] Interestingly, we are also able to document a significant decline in misreporting since publication of the Lim, Stein, Charrow, and Murray (2008) study, which suggest GAVI was able to partially rectify the perverse incentives it helped create – though largely by reducing payouts under the incentive scheme.

3.1 Vaccination data

We focus on two childhood vaccinations – DTP3, which was the target of the incentive scheme, and the vaccine against measles, which was not included in the scheme. Data is drawn from two sources: household surveys and administrative data on vaccination rates.

[11] The implicit counterfactual used by Lim, Stein, Charrow, and Murray (2008) is simply the past, i.e., that without GAVI ISS, the *absolute* gap between survey and administrative data in the 2000s would have looked exactly like the 1990s, despite the rapid growth in immunization and health aid over this period, as well as improvements in both survey and administrative data coverage.

Our benchmark is the vaccination rate for a given disease in a given country and year, based on harmonized household survey data, i.e., the Demographic and Health Survey (DHS) indicators. The DHS data is sponsored by the United States Agency for International Development (USAID), and is collected with a high level of technical assistance from USAID contractors based in the U.S. For African countries participating in the DHS, international technical assistance frequently covers sample design and selection, enumerator training, questionnaire development, and all data cleaning and analysis. From the perspective of country ownership, capacity building, and sustainability, this heavy-handed approach to foreign technical assistance may be regrettable. For the narrow purposes of this study however, DHS provides us with a more reliable, fairly independent benchmark against which to judge administrative data.

From 1990 to 2011, the pooled DHS data set contains 181 surveys spanning 70 countries with data on both DTP3 and measles coverage, of which 93 surveys were conducted across 41 African countries. We focus on this sample of African countries here, though note below where results differ from the broader global sample. For the 44 African surveys in the period 1990-2000 (i.e., pre GAVI), the average rate of DTP3 vaccination in our sample was 59% and 63% for measles. For the 49 African surveys from 2001-2011, this rose to 70% for DTP and 72% for measles. In both periods, there was enormous heterogeneity across countries, with the lowest performer recording vaccination rates below 25% for both diseases (i.e., Chad) and the highest performers above 90% coverage for both diseases (e.g., Rwanda and Malawi).

To test hypotheses about mis-reporting, we compare this household survey data to administrative data as recorded by the WHO. The primary data series we use consists of the estimates of vaccination coverage posted online by the WHO.[12] Based on personal communication with WHO staff, our understanding is that from 2000 onward (but not before) these figures do not reflect countries' raw reports, but have been modified by the WHO – often with the specific goal of ensuring greater consistency with survey data. Clearly, any such correction would cause us to understate the effect of GAVI on misreporting. To overcome this problem, we extract the original, uncorrected "official" vaccination rates submitted by countries in their country reports.[13] We rely on the WHO web data as our primary data source, but also test the robustness of our results to using the raw, uncorrected reports from

[12]These data are available for download here: http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tscoveragedtp3.html.

[13]These official country reports are available only in PDF format, which we converted to machine-readable data. The PDFs are available here: http://apps.who.int/immunization_monitoring/globalsummary/wucoveragecountrylist.html.

countries which we have extracted from the PDFs. On the whole, we find discrepancies between survey and administrative data are larger when using the uncorrected reports, but the sample of PDFs from which we draw is quite small ($N = 42$) to make confident inferences.

Finally, a number of differences between our sample and that used by Lim, Stein, Charrow, and Murray (2008) bear emphasis. We rely on a single set of harmonized household surveys believed *ex ante* to be as comparable as possible, both across countries and – more importantly, given our estimation strategy – over time within countries. In contrast, Lim, Stein, Charrow, and Murray (2008) use a variety of survey sources including vaccination rates based on maternal self-reports and drop observations they define as outliers ex post, though there is no sign this affects results. Additionally, Lim, Stein, Charrow, and Murray (2008) rely on fairly ambitious imputation of survey data for years in which it is not available, expanding their apparent sample size several fold and inducing a risk of overly confident statistical inferences. We eschew imputation, preferring to restrict ourselves to the most reliable available survey data for the years in which it is available.

3.2 Estimation and results: GAVI and vaccination rates

Our estimation strategy amounts to a quadruple-difference estimate of the effect of GAVI's incentive scheme, comparing (i) changes over time (ii) in survey versus administrative data, (iii) before and after 2000 (iv) for DTP3 versus measles.

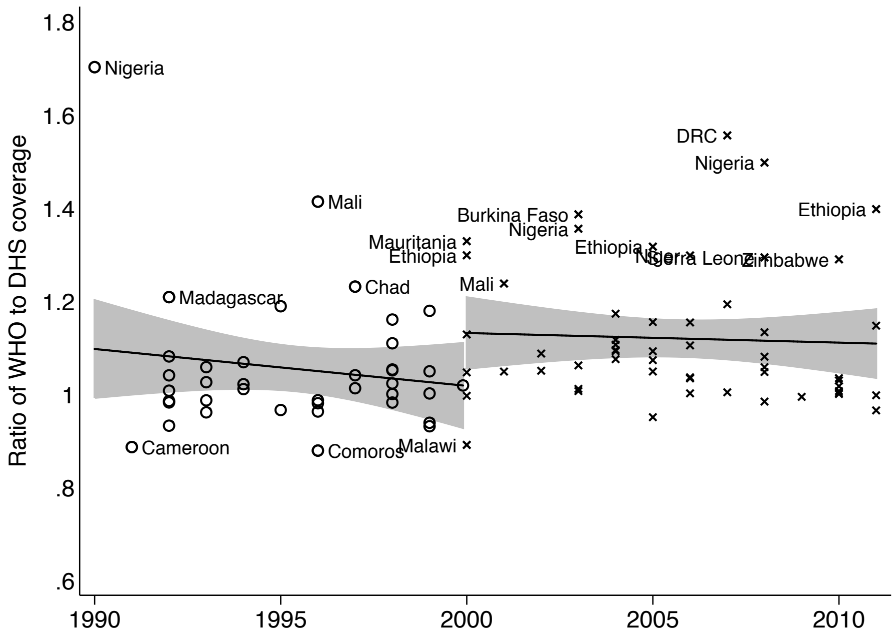
Let c index countries, d diseases (DTP3 or measles), and t years. We regress changes over time within the same country in vaccination rates as measured in administrative data, ΔV_{cdt}^{WHO}, on the equivalent change in survey data, ΔV_{cdt}^{DHS}.

$$\begin{aligned} \Delta V_{cdt}^{WHO} &= \beta_0 + \beta_1 \Delta V_{cdt}^{DHS} + \beta_2 I[t \geq 2000] + \beta_3 I[t \geq 2008] + \beta_4 I[d = DTP3] \\ &\quad + \beta_5 I[t \geq 2000] \times I[d = DTP3] + \beta_6 I[t \geq 2008] \times I[d = DTP3] + \varepsilon_{cdt} \quad (1) \end{aligned}$$

We alternate between estimating equation (1) in levels rather than differences, and running separate regressions for DTP3 and measles. Here the coefficient of interest is β_2, which measures whether discrepancies increased after the GAVI scheme was introduced (i.e., when $I[t \geq 2000]$ takes a value of one). We then pool the data so that each country-year reports two observations. In the pooled specification, the parameter of primary interest is β_4, the coefficient on the interaction between a dummy variable for observations in 2000 or later

Figure 1: Vaccination rates: WHO vs DHS

(a) DTP



(b) Measles

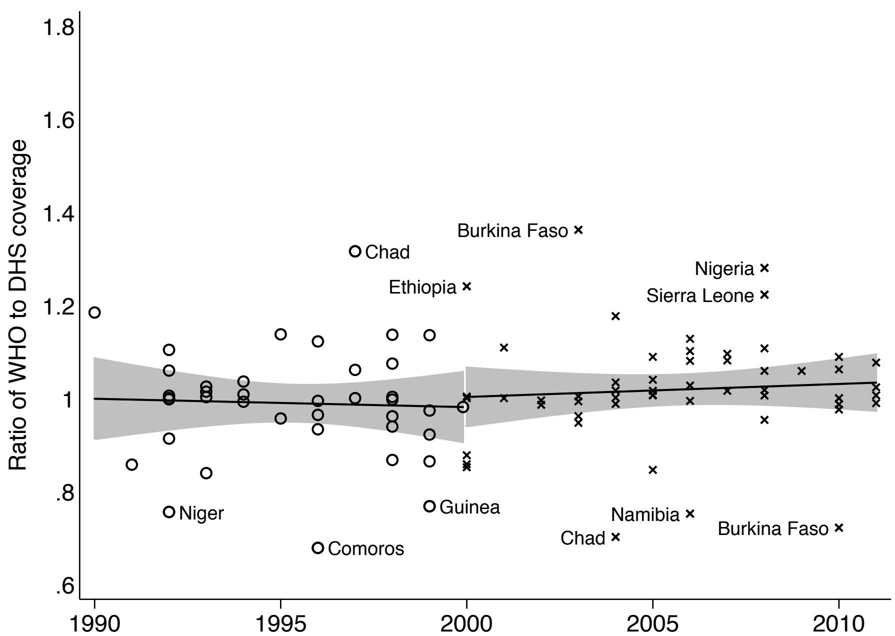


Table 1: Immunization rates: Regression results

	DTP3			Measles			Pooled		
	Level		Change	Level		Change	Level		Change
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year $\geq$ 2000	11.03	5.55	5.84	7.75	2.25	2.60	7.75	2.48	1.75
	(3.99)***	(1.30)***	(2.38)**	(3.42)**	(1.66)	(3.07)	(3.43)**	(1.65)	(2.89)
Year $\geq$ 2008	11.61	-.62	-6.15	12.16	-.64	-3.93	12.16	-.12	-4.75
	(3.69)***	(2.16)	(3.59)*	(3.34)***	(2.75)	(3.47)	(3.35)***	(2.76)	(4.01)
DHS rate		.87	.95		.94	.66		.90	.82
		(.03)***	(.08)***		(.04)***	(.19)***		(.03)***	(.12)***
DTP3							-.24	2.35	-2.39
							(.99)	(.85)***	(1.79)
(Year $\geq$ 2000) $\times$ DTP3							3.28	2.90	4.67
							(1.54)**	(1.12)***	(2.27)**
(Year $\geq$ 2008) $\times$ DTP3							-.55	-.88	-1.12
							(3.16)	(3.12)	(3.85)
Const.	61.11	9.38	-2.29	61.34	3.06	.23	61.34	5.44	.31
	(4.08)***	(2.16)***	(2.30)	(3.90)***	(2.47)	(2.28)	(3.91)***	(2.03)***	(2.28)
Observations	93	93	57	93	93	57	186	186	114
Countries	41	41	26	41	41	26	41	41	26

The sample is restricted to African countries. Each column shows a separate regression. The dependent variable is the immunization rate reported in administrative data via the WHO. The unit of observation is a country-year for a given disease. Each regression in columns 1-6 includes one observation per country-year, while the pooled regressions in columns 7-9 include two observations per country-year: one for DTP3 and one for measles. "Year $\geq$ 2000" is a dummy variable for the years 2000 and beyond; DTP3 is a dummy variable that takes the value of one for DTP3 rates and zero for measles rates. For the first two columns under each disease, both the dependent variable and the DHS rate are measured in levels; in the third column under each disease both of these variables are measured in changes (i.e., first differences between survey rounds). All standard errors are clustered at the country level. Asterisks (*,**,***) denote coefficients that are significantly different from zero at the 10, 5 and 1% levels, respectively.

and a dummy variable for observations of DTP3 rather than measles vaccination ($I[t \geq 2000] \times I[d = DTP3]$). Our estimate of $\hat{\beta}_5 > 0$ measures the degree to which the discrepancy between administrative and survey data grew more rapidly over time for DTP3 vaccinations relative to measles vaccinations after the onset of the GAVI scheme in 2000. The β_3 and β_6 terms are included to test whether any observed discrepancies have improved in more recent years.

Starting with the simplest single-difference comparison, Table 1 reports regression estimates of DTP3 and measles immunization coverage before and after the onset of GAVI incentive scheme (essentially, estimating β_2 in isolation). The dependent variable in each case is the immunization rate from the administrative HMIS reported to the WHO. Without controlling for other variables, average DTP3 rates in our sample of 93 observations spanning 41 countries increased by 11 percentage points from 2000 onward (column 1), while measles coverage increased by roughly 8 percentage points (column 4). Column 7 pools the data from both diseases to compute the difference-in-differences (the shift after 2000 in administrative data for DTP3 relative to measles) and shows that this divergence in immunization rates of 11 versus 8 points is statistically significant at the 5% level.

Controlling for the coverage rate reported by the DHS puts the focus directly on discrepancies between administrative and survey data. When including this control, DTP3 coverage rates in the administrative data rose by 5.6 percentage points after 2000 (column 2, statistically significant at the 1% level) while measles coverage rose by only 2.3 points (column 4, insignificant). Column 8 shows that this triple-difference is again significant (now at the 5% level). Finally, we look at changes over time immunization rates before and after 2000 to produce a quadruple-difference comparison. As see in column 9, discrepancies in administrative data on DTP3 coverage accelerated significantly (by 4.7 percentage points, significant at the 10% level) relative to measles rates.[14]

So far we have focused on establishing a casual effect of GAVI's ISS program on misreporting of DTP3 data (i.e., the post-2000 dummies). While we would argue that comparison to a control group (in this case, measles vaccinations) is necessary to draw causal inferences,

[14] As noted in the previous sub-section, all of these results can be replicated using the data on the WHO website, rather than the raw reports submitted by countries to the WHO in PDF format which we have used here. When we re-run the regressions with the web data, the findings are strikingly robust. The size of the discrpeancies is slightly smaller for both diseases, but the divergence in discrepancies between DTP3 and measles after 2000 is almost identical; i.e., the point estimates in columns 1-6 are mostly smaller but retain statistical significance, while those in columns 7-9 are virtually unchanged. For brevity, we do not reproduce an alternate version of the table here. Results are avaialble upon request.

the existence of discrepancies was highlighted prominently several years ago by Lim, Stein, Charrow, and Murray (2008). An obvious question then arises: did awareness of these problems, at least since 2008, lead to any improvements? We can answer this question using the coefficient on the dummy variable for 2008 and thereafter (as well as it's interaction with DTP3). Both DTP3 and measles coverage continued to grow after 2008, as seen in columns 1 and 3. The good news is that it appears this growth in coverage was genuine, in the sense that administrative and survey sources agree. Furthermore, there is suggestive evidence that the discrepancies induced by the GAVI ISS program have ameliorated since 2008. Column 3 shows a larger and marginally significant (at the 10% level) reduction in DTP3 discrepancies since 2008 compared to a smaller and insignificant improvement in measles coverage in column 6. [15]

The last point suggests that GAVI, the WHO, and national governments may have successfully responded to the Lim, Stein, Charrow, and Murray (2008) findings. How was this achieved? The short is answer is that GAVI disbursements through the ISS program declined rapidly from 2008 onward.[16] From a peak of $90 million disbursed in 2007, the spending rate declined to roughly $10 million per annum from 2009 to 2011 and fell even further in subsequent years. So while misreporting may have been ameliorated, this was achieved by curtailing payment-for-performance payments. An important outstanding question is whether reflected (a) better monitoring of data quality, leading to fewer payments, or (b) a quiet scale-back of the program in the face of irresolvable data issues. The former would suggest such data problems are fixable with proper due diligence. The latter would suggest that the perverse incentives created by a payment for performance scheme were corrected only by effectively removing the scheme. As of now, we cannot distinguish these hypotheses.

To summarize, in the case of both DTP3 and measles, we find over- and under-reporting of vaccination coverage by administrative sources. On average, over-reporting of DTP3 coverage was about 5% prior to the GAVI scheme, and non-existent for measles. But this changed over time. When comparing trends in discrepancies, DTP3 discrepancies accelerated when GAVI introduced its ISS incentives in the early 2000s, and this acceleration significantly outpaced similar trends for measles vaccination. This analysis confirms and updates the

[15] This relative improvement for DTP3 compared to measles since 2008 is not itself significant, as seen in columns 7 through 9. Sample sizes since 2008 may simply be too small to detect real changes. A more optimistic reframing of the same regression result would note that since 2008 there is no significant tendency for discrepancies in DTP3 administrative data relative to measles. That is, while we have no definitive proof the problem as been fixed since 2008, nor is there definitive proof it still exists.

[16] http://www.gavialliance.org/results/disbursements/

findings of Lim, Stein, Charrow, and Murray (2008); without greater verification of self-reported administrative data, financial incentives from donors may affect the accuracy of data used by the vaccination program by making it vulerable to political interference. This finding is not a general feature of survey versus administrative data (as we rely on changes over time) and is not a general feature of periods where vaccination rates are increasing rapidly, as we find misreporting that is specific to the diseases incentivized by GAVI.

This lack of accuracy has knock-on effects for the impact of programs. GAVI, for example, uses coverage data to forecast vaccine procurement. While over-procuring inexpensive vaccines such as measles (at 3 cents a dose) does not imply large additional costs or major trade-offs with other health system priorities, newer vaccines donated by the GAVI Alliance cost about US$3.50 per dose, and require several doses. As a result, every vaccine purchased that is not used due to inaccuracy in vaccination coverage implies significant expense and opportunity cost, both in lives and money.

Not all, or perhaps even most, of the discrepancies in HMIS data are the result of the incentives to misreport provided by the GAVI ISS program. The issue of weak state capacity to monitor front-line service providers discussed in Section 4 is likely crucial here as well. We note this to caution against interpreting the analysis here to imply that health systems suffer from one type of malaise, while education systems suffer from another. We do not believe this is the case: rather the health data provides the opportunity to identify one specific problem, and the education data another.[17]

4 Fooling the state: School enrollment across 21 countries

Having focused in the previous section on the role of national governments as suppliers of statistics, we return to the challenges facing national governments as users of statistics for evidence-based policymaking.

[17]Note also that vaccination coverage is only one of many essential indicators of health system performance that are affected by weak institutional capacity to collect and analyze data. For example, even in countries where vital registration systems are almost complete in terms of coverage, the quality of reporting remains an important problem. In South Africa, where 89% of adult deaths are reported via the vital registration system, a death certificate audit found errors in 45% of all records (Yudkin, Burger, Bradshaw, Groenewald, Ward, and Volmink, 2009; Nojilana, Groenewald, Bradshaw, and Reagon, 2009; Bradshaw, Groenewald, Bourne, Mahomed, Nojilana, Daniels, and Nixon, 2006; Burger, Van der Merwe, and Volmink, 2007). A 2009 study found that 43 out of 46 countries in the WHO/AFRO region had no population-level data on cause of death (Mathers, Boerma, and Fat, 2009).

We discuss the pitfalls of the principal-agent relationship between central governments and front-line service providers scattered across the country. Rather than managing a single agent, line ministries require data on thousands of schools, clinics, police stations, water points, and road maintenance activities across the country. Given many African states' historical weakness in terms of their ability to exert control over or provide government services to remote populations (Herbst, 2000), it is perhaps unsurprising that they struggle to collect reliable data on these same activities.

We draw lessons from the education sector, but similar examples exist in the agriculture and other sectors. What these cases share in common is that administrative data systems are based on self-reports by low-level public servants. The resulting biases are disproportionately in the direction one might expect given the inherent conflicts of interest in data collection, and they help point to lessons about (a) how public finance systems and administrative data systems must be designed in tandem to avoid compromising the integrity of the evidence base for policymaking, and (b) how surveys could be designed to complement and correct rather than substitute for administrative data sets.

4.1 Enrollment data

We compare two independent sources of information on primary school enrollment: administrative records and household surveys. Administrative records are drawn primarily from the Education Monitoring and Information System (EMIS) databases sponsored by UNESCO and maintained by Ministries of Education throughout the region. EMIS data is typically compiled from reports submitted by school officials and aggregated up. We compare these records to survey-based estimates of school enrollment, focusing both on levels at a point in time and trends over time. Surveys data are taken from Demographic and Health Surveys (DHS) sponsored by DHS and collected by national statistics offices, usually in collaboration with Ministries of Health.

The full sample of the 21 country-year periods for which comparable administrative and survey data are available is listed in Table 2. Fifteen of the 21 spells show discrepancies in the direction of greater optimism in the administrative data relative to household surveys. This tendency appears to be particularly pronounced in sub-Saharan Africa: the average gap between enrollment growth in administrative versus survey data (i.e., the degree of over-optimism in administrative data) was 3.1 percentage points in the African sample, but was slightly in the pessimistic direction at -0.8 for the 15 observations available from non-African countries.

There are multiple reasons why EMIS records may exhibit systematic biases. The first is underreporting of private schools. There is evidence from household surveys of a rapid increase in private schooling in at least some countries (Bold, Kimenyi, Mwabu, and Sandefur, 2011a), and even where theoretically required to report to EMIS, unregistered schools may have little incentive to do so. The second, potentially more damaging bias stems from the incentives for public school officials to report accurately. The abolition of school fees for primary education in much of the region has brought a shift, in many cases, to a system of central government grants linked to the head-count of pupils. In Tanzania, for instance, enrollment rates in the EMIS database suggest the country is on the verge of reaching the Millennium Development Goal of universal primary enrollment. Yet household survey estimates show that 1 in 6 children between ages 7 and 13 are not in school (Morisset and Wane, 2012).

We explore the second hypothesis by examining two cases from Table 2 which exhibit large discrepancies and where survey data spans the abolition of user fees in primary education: Kenya and Rwanda.

Kenya abolished user fees in government primary schools beginning with the 2003 school year. Figure 2a shows the trend in net primary education spanning this reform as reported by the Ministry of Education's (MOE) administrative data, as well as two independent household survey data sources: two rounds of the DHS conducted in 2003 and 2008, and two successive surveys conducted by the Kenyan National Bureau of Statistics (KNBS) in 1997 and 2006.[18] The striking feature 2a is the steady upward trend in enrollment, including a sharp jump with the introduction of free primary education, juxtaposed with absolutely no change in enrollment measured by either household survey.

Rwanda, which also abolished user fees for primary education in 2003, presents a similar, if slightly less stark picture in Figure 2b. Administrative data from the Ministry of Education (MINEDUC) shows steady enrollment growth spanning the abolition of fees. The DHS rounds from 2000 to 2005 confirm this general growth trend, but the 2005 to 2010 rounds of

[18]These are the 1997 Welfare Monitoring Survey and the 2006 Kenyan Integrated Household Budget Survey, which measured comparable indicators of primary school enrollment. For further discussion of these surveys and the effect of Kenya's free primary education policy on enrollment, see Bold, Kimenyi, Mwabu, and Sandefur (2011b).

the DHS show a very modest increase from 85 to 87% net enrollment, while the administrative data shows a huge leap over the same period from 82 to 99%.

How comparable are the data in these Kenyan and Rwandan examples? Are surveys and administrative data sets measuring the same thing? Note that while administrative statistics are typically defined in terms of net and gross enrollment, the DHS fields two distinct question types, both of which focus on school *attendance*. In theory, attendance rates may be higher or lower than enrollment rates, for reasons above and beyond the main hypotheses of this paper: for instance, children may enroll but rarely attend; on the other hand, Ministry of Education enrollment statistics may omit attendance at private and non-recognized schools which is reported in survey data. In practice, Kenya reported DHS primary attendance rates well above official enrollment numbers prior to the introduction of free primary education as seen in Figure 2a. Beyond Kenya though, on average the discrepancy goes the other way: DHS attendance rates are about 13% lower than official enrollment rates in our regression sampe.

The relevant question for our analysis is how these differences in definitions might affect the evolution of enrollment and attendance rates over time, before and after the abolition of user fees. One clear pattern observed in Kenya after FPE was a large increase in private school enrollment (Bold, Kimenyi, Mwabu, and Sandefur, 2011b; Lucas and Mbiti, 2012). This would, if anything, tend to increase survey relative to official statistics, causing us to understate any effect of FPE in exacerbating data discrepancies. It is also possible that fee abolition led to enrollment of pupils with little intention of actually attending school, leading to a genuine and widening discrepancy between enrollment and attendance rates. This alternative hypothesis should be kept in mind when interpreting our results below. We note however, that while the mechanism is different, the public finance implications of this story are very similar: a change in the way schools are funded leads to discrepancies in data between how many pupils are counted and how many are actually in school.

Note that we focus our analysis (and report all numbers in the tables) using net enrollment and attendance rates, which provide the best available apples-and-apples comparison across data sources and time periods. We should note, however, that there are often anecdotal reports of a surge in gross enrollment relative to net enrollment after fee abolition, as older children enter school who were previously excluded. This should in no way affect our results, but merits exploration.[19] These anecdotes find limited corroboration in administrative data from the WDI which include both gross and net numbers: in our sample of African countries with data spanning the abolition of fees, average net enrollment before and after FPE rose by 41% or about 30 percentage points (from 74% to 104%) while net enrollment rose by 44% or about 24 percentage points (from 54% to 78%).

Unfortunately, it is not possible to construct comparable gross attendance rates spanning fee abolition from the DHS data. While more recent DHS data also includes both a gross and net rate, older DHS data sets only report rates by age group. We use these older numbers to construct a rough estimate of net attendance using the 6 to 10 age range (valid assuming no children in this age range attend secondary school, otherwise an over-estimate) but are unable to construct a gross figure using these DHS data.

To test whether the patterns observed in Kenya and Rwanda represent a systematic pattern in the data, we draw on the full sample of African countries for which comparable administrative and survey data is available. This includes 46 spells spanning 21 countries, of which 35 spells are drawn from the 14 countries in Table 2 with more than one data point over time. All of these countries, with the exception of Guinea, previously imposed user fees for primary education and repealed those fees during the period covered here. We record the year of fee abolition for each country in the final column of Table 2 , based on information reported in Riddell (2014).[20] This staggered abolition of user fees across countries creates the quasi-experiment we exploit in the following sub-section.

4.2 Estimation and results: User fees and net enrollment statistics

Using a methodology similar to the vaccination regressions in Section 3, we regress changes in primary school enrollment from administrative data (ΔE_{ct}^{WDI}, as reported in the World Bank's World Development Indicators) on changes in enrollment in survey data (ΔE_{ct}^{DHS}).

$$\Delta E_{ct}^{WDI} = \gamma_0 + \gamma_1 \Delta E_{ct}^{DHS} + \gamma_2 \Delta FPE_{ct} + \gamma_3 t + \upsilon_{cdt} \tag{2}$$

[19] An exception, whereby gross enrollment trends could be driving our results, would be a case in which (a) gross and net enrollment are conflated in the data, (b) this conflation is unique to administrative data and not found in the survey data, and (b) worsens over time after FPE. We know of no reason to suspect this is happening, but cannot rule it out as a potential mechanism to explain the data discrepancies we document.

[20] For countries that abolished user fees after the publication date of Riddell (2014), we rely on a variety of U.N., NGO, and media reports, e.g., UNICEF (2012) for Burkina Faso, Brandt (2013) for Namibia, and Scan News (2013) for Nigeria.

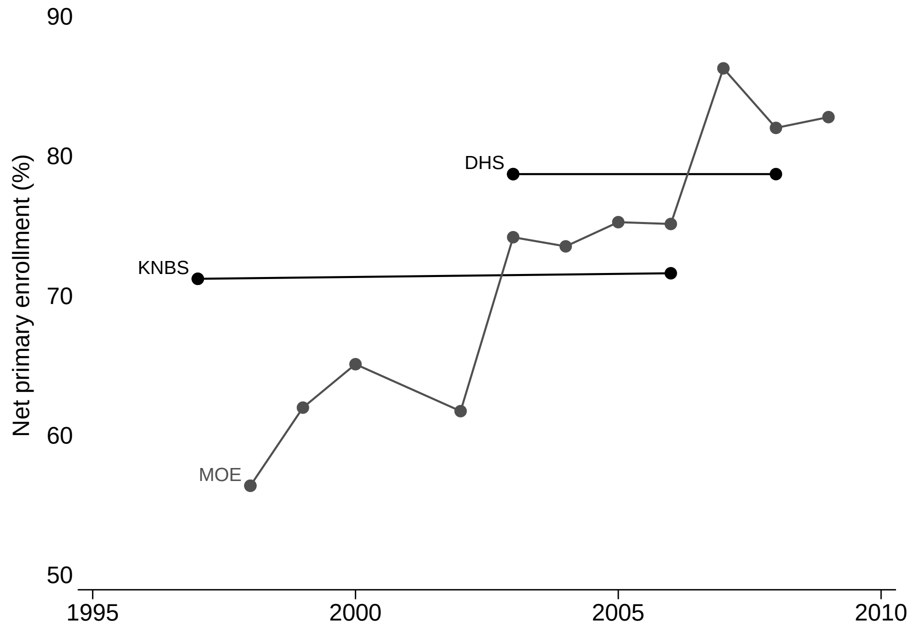
Table 2: Changes in primary school net enrollment

Country	Years		Admin. data			Survey data				
	Start	End	Start	End	Δ	Start	End	Δ	Gap	FPE
Kenya	1998	2003	56.4	74.2	17.8	82.3	78.7	-3.6	21.4	2003
Rwanda	2005	2010	81.9	98.7	16.9	85.4	87.3	1.9	15	2003
Ethiopia	2000	2005	40.3	61.9	21.6	30.2	42.2	12	9.6	2002
Cameroon	1991	2011	70	93.8	23.7	63.7	78.3	14.6	9.1	1999
Burkina Faso	1993	1999	26.9	33.3	6.4	26.6	25	-1.6	8	2007
Kenya	2003	2008	74.2	82	7.8	78.7	78.7	0	7.8	2003
Benin	1996	2006	62	87.1	25.1	41.8	60.1	18.3	6.8	2006
Burkina Faso	2003	2010	36.5	58.1	21.5	28	44.4	16.4	5.1	2007
Eritrea	1995	2002	26.5	43.2	16.7	36.6	50.3	13.7	3	2005
Niger	1992	2006	22.3	43.2	20.9	13.6	32.1	18.5	2.4	2009
Ethiopia	2005	2011	61.9	86.5	24.6	42.2	64.5	22.3	2.3	2002
Guinea	1999	2005	43.2	68.3	25.1	21.6	45	23.4	1.7	
Senegal	2005	2010	72.2	75.5	3.3	52	54.3	2.3	1	2001
Namibia	1992	2000	82.6	88.1	5.5	76.5	81.3	4.8	0.7	2013
Burkina Faso	1999	2003	33.3	36.5	3.2	25	28	3	0.2	2007
Tanzania	1999	2004	49.3	86.2	36.9	35	73.1	38.1	-1.2	2001
Tanzania	1992	1996	50.6	48.7	-1.9	26.2	27.3	1.1	-3	2001
Nigeria	1999	2003	61.3	65.6	4.3	56.7	64.2	7.5	-3.2	1999
Nigeria	2003	2008	65.6	58.8	-6.8	64.2	62	-2.2	-4.6	1999
Tanzania	1996	1999	48.7	49.3	0.6	27.3	35	7.7	-7.1	2001
Lesotho	2004	2009	73.9	71.9	-2	80.9	88.8	7.9	-9.9	2000
Ave.: Africa					12.9			9.8	3.1	
Ave.: Other					3.8			4.5	-0.8	

Table reports the starting and ending rates (%) of net primary enrollment in the administrative data reported in the WDI data base and attendance rates in the DHS survey data, and their respective changes over time. The "gap" measures the difference between the rise in the admin data and the rise in the survey data. The FPE column lists the date that the country removed user fees for public primary education.

Figure 2: Trends in net primary enrollment: See text for sources.

(a) Kenya



(b) Rwanda

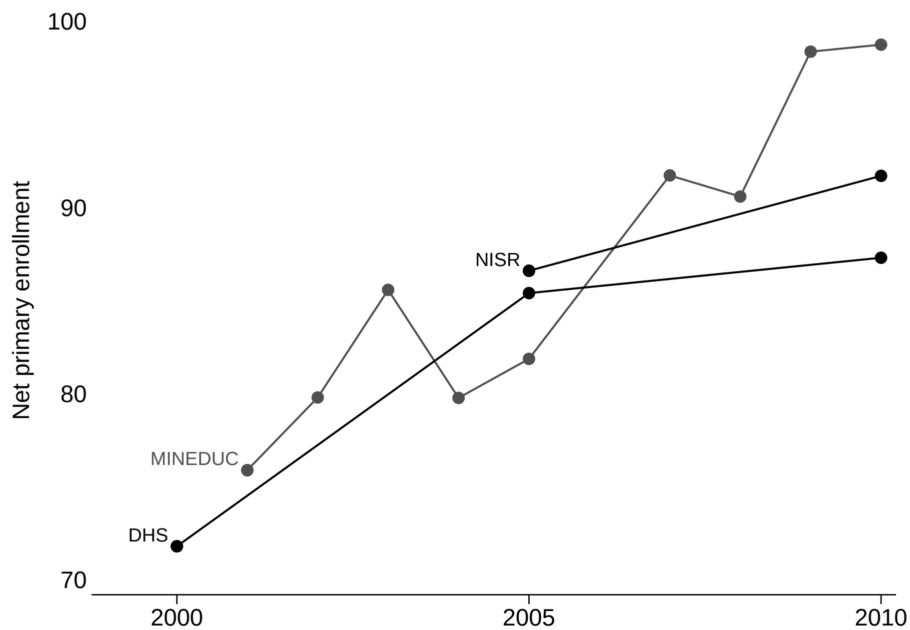


Table 3: Primary school net enrollment rates: Regression results

	DHS		WDI			
	Level	Change	Level	Change	Level	Change
	(1)	(2)	(3)	(4)	(5)	(6)
Free primary educ (dummy)	15.81	6.46	15.30	14.89	4.93	10.25
	(7.48)**	(5.66)	(6.79)**	(4.90)***	(4.82)	(3.59)***
Time trend	1.38	.34	.96	.49	.06	.25
	(.66)**	(.43)	(.53)*	(.43)	(.45)	(.42)
DHS enrollment (%)					.66	.72
					(.12)***	(.18)***
Const.	42.40	4.79	56.81	5.84	29.00	2.40
	(7.71)***	(2.25)**	(7.79)***	(2.75)**	(7.14)***	(2.77)
Obs.	46	21	46	21	46	21
Countries	21	14	21	14	21	14

The sample is restricted to African countries. Each column shows a separate regression. The dependent variable is the primary school enrollment rate as measured by either the WDI (primarily administrative data) or DHS (survey data). The unit of observation is a country-year. In columns labeled 'Change', the dependent variable as well as both the FPE dummy and the DHS enrollment variable are measured in first-differences, as shown in equation (2). The time trend is a year variable set to zero in 2000. All standard errors are clustered at the country level. Asterisks (*,** ,***) denote coefficients that are significantly different from zero at the 10, 5 and 1% levels, respectively.

We estimate equation (2) both in levels and in changes. In both cases, the parameter of interest is γ_2, which measures the extent to which the gap between administrative and survey-based enrollment rates diverged after the introduction of free primary education.

As a preliminary step, columns 1 and 2 of Table 3 use the DHS enrollment rates as the dependent variable, showing that enrollment rates were nearly 16 percentage points higher for observations in our sample after fees were abolished, though when comparing like with like (i.e., in the first-difference specification in column 2), this change falls to just 6 percentage points and is statistically insignificant. Columns 3 and 4 repeat this exercise for the WDI administrative data, which rose by 15% on average and remained at 15% when comparing changes within the same country, significant at the 1% level in both cases. Finally, columns 5 and 6 estimate the full specification shown in equation (2), with the administrative data as the dependent variable and the survey data as a control. Our estimate of $\hat{\gamma}_2$ is 4 percentage points (statistically insignificant) in the levels specification in column 5, but rises to 10.3% (significant at the 1% level) when comparing changes within the same country in column 6. In all cases, standard errors are clustered at the country level.

As an additional placebo test, we can run the same regression from Table 3 using secondary instead of primary data. If our hypothesis is correct, and discrepancies are driven by the incentives to misreport created by the abolition of user fees at the primary level, we should see no effect at the secondary level. Unfortunately, data is extremely limited, as many countries do not report secondary net enrollment administrative statistics in the WDI. For the handful of countries with comparable data, it is notable that discrepancies between administrative and survey data *declined* by about 8%, while they increased at the primary level for this same set of countries by about 15%.[21]

To summarize, we find that administrative data sources claim enrollment increases that are more than twice as fast after FPE than comparable survey sources – a differential of roughly ten percentage points.

While the patterns in Kenya and Rwanda, and even from the econometric results for the broader sample of countries, are far from definitive proof of a causal chain from top-down funding for education to the production of unreliable enrollment statistics, they are suggestive of an important link between funding systems and data systems. EMIS in much of the region appear to have been designed with little regard for the incentives faced by teachers and education officials to report accurately. With the onset of free primary education, these incentives changed radically, and the data systems have failed to keep up. The resolution to overcoming these perverse incentives is not immediately obvious, but in Section 5 we discuss possible ideas for how survey and administrative data collection efforts could be better integrated so that the former could act as a check on misreporting in the latter.

5 Discussion and conclusion

While recognizing the different uses and timing of administrative and survey data, our analyses of the discrepancies between administrative data and household survey-based estimates in education and health suggest that - in some African countries – there are significant inaccuracies in the data being published by national and international agencies. These inaccuracies appear to be due in part to perverse incentives created by connecting data to financial incentives without checks and balances, and to competing priorities and differential funding associated with donor support.

[21] The countries included are Burkina Faso, Eritrea, Guinea, Kenya, and Lesotho. Note that while both the increase in discrepancies at primary level and decrease at secondary level are statistically significant, the regression in changes (as opposed to levels) includes only seven data points, thus we do not report tables here. Full results are available upon request.

Further, in spite of international declarations to support statistical capacity in Busan and a concerted effort by the World Bank's IDA and Paris21 to support national statistical strategies, indices prepared by the World Bank and UNECA suggest that performance has not improved much over time, in large part because statistical agencies in the region – particularly those in Anglophone Africa – lack functional independence, fail to attract and retain high-quality staff, depend on external funders for the majority of their spending and experience significant volatility and unpredictability in their year-to-year budgets. Plans are often divorced from budget realities, thus forcing NSOs to prioritize "paying customers" rather than national priorities and core statistical activities as articulated in country developed plans.

Together, these inaccuracies, perverse incentives and lack of functional independence mean that public and private investment decisions based on poor data can be deeply flawed, with major implications for well-being and public expenditure efficiency. Further, pressure to open data can result in "garbage in, garbage out" if measures are not taken to strengthen underlying source data.

There are clear opportunities to redesign accountability relationships in ways that can serve all kinds of data clients, whether national governments, donors or citizens. Getting administrative and survey data to "speak to each other" is one strategy, where household surveys can be used to provide regular checks on administrative data at different levels. This analysis has shown that a system to identify perverse incentives currently operating within various sectors including the NSO and line ministries' statistics units could provide a valuable measure of the statistical capacity that matters, and can suggest alternative policy or measurement arrangements. This would be one way to deliberately assess the extent to which administrative data contains avoidable biases and to understand where the introduction of additional checks and balances would be needed to correct the accuracy issues within that data.

Both of the main case studies examined in this paper involved data misrepresentations that resulted from pay-for-performance initiatives (i.e., the GAVI ISS or capitation grants to primary schools). Such schemes appear to be spreading in the region, both due to the widespread roll-back of user fees and to donor enthusiasm for results-based aid. For instance, Huillery and Seban (2013) notes 18 ongoing impact evaluations of payment-for-performance in the health sector within African countries. In future, new pay-for-performance initiatives must seek to avoid the pitfalls documented here. One potential example of how to do this is provided by pay-for-performance agreements between national governments and providers under the World Bank's Health Results Innovation Trust Fund (HRITF). The HRITF uses survey data in small, local samples to cross-validate and improve the accuracy of the administrative data used to report performance. HRITF verifies data reported by participating facilities using a representative sample, visited unannounced, while also including penalties for over-reporting measured using the micro-household survey. A clear and relatively rapid jump in the accuracy of self-reported data on quantity of services delivered has been observed. In Cameroon, for example, independent verification of administrative data helped reduce over-reporting of outpatient consultations by over 90% in less than a year (World Bank, 2013). Still, there remains much to learn about the optimal strategy for measuring and verifying data quality.

These kinds of mutually reinforcing administrative plus survey data verification arrangements will make all kinds of results-based funding work better, whether between a national government and subnational providers, or between donors and recipient country governments.

References

Bold, T., M. Kimenyi, G. Mwabu, and J. Sandefur (2011a): "The High Return to Private Schooling in a Low-Income Country," Center for Global Development Working Paper No 279.

——— (2011b): "Why Did Abolishing Fees Not Increase Public School Enrollment in Kenya?," Center for Global Development Working Paper Series, 271.

Bradshaw, D., P. Groenewald, D. E. Bourne, H. Mahomed, B. Nojilana, J. Daniels, and J. Nixon (2006): "Making COD statistics useful for public health at local level in the city of Cape Town," Bulletin of the World Health Organization, 84(3), 211–217.

Brandt, E. (2013): "Namibia: Free Primary Education Starts," http://allafrica.com/stories/201301090299.html, Online; accessed 2013-08-08.

Burger, E. H., L. Van der Merwe, and J. Volmink (2007): "Errors in the completion of the death notification form," South African Medical Journal, 97(11), 1077–81.

Caeyers, B., N. Chalmers, and J. De Weerdt (2012): "Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment," Journal of Development Economics, 98(1), 19–33.

Demombynes, G. (2012): "Opening Up Microdata Access in Africa," World Bank 'Development Impact' Blog, October 1, 2012.

Devarajan, S. (2013): "Africa's Statistical Tragedy," Review of Income and Wealth.

Herbst, J. (2000): States and power in Africa: comparative lessons in authority and control. Princeton University Press.

Huillery, E., and J. Seban (2013): "Performance-Based Financing for Health: Experimental Evidence from the Democratic Republic of Congo," .

Jerven, M. (2010): "Random Growth in Africa? Lessons from an Evaluation of the Growth Evidence on Botswana, Kenya, Tanzania and Zambia, 1965-1995," The Journal of Development Studies, 46(2), 274–294.

——— (2011): "Users and producers of African income: Measuring the progress of African economies," African affairs, 110(439), 169–190.

——— (2013): Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It. Cornell University Press.

Lim, S., D. Stein, A. Charrow, and C. Murray (2008): "Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage," The Lancet, 372(9655), 2031–2046.

Lucas, A. M., and I. M. Mbiti (2012): "Access, sorting, and achievement: the short-run effects of free primary education in Kenya," American Economic Journal: Applied Economics, 4(4), 226–253.

Mathers, C. D., T. Boerma, and D. M. Fat (2009): "Global and regional causes of death," British Medical Bulletin, 92(1), 7–32.

McQueston, K. (2013): "Autonomy, Independence, and Capacity of National Statistics Offices," Background Paper for the Data for African Development Working Group: Center for Global Development, Washington D.C., and the African Population and Health Research Council, Nairobi.

Morisset, J., and W. Wane (2012): "Tanzania: Let's think together," World Bank, accessible at http://blogs.worldbank.org/africacan/tanzania-lets-think-together/. Last accessed on September 10, 2012.

Nojilana, B., P. Groenewald, D. Bradshaw, and G. Reagon (2009): "Quality of cause of death certification at an academic hospital in Cape Town, South Africa," SAMJ: South African Medical Journal, 99(9), 648–652.

Riddell, A. (2014): "The introduction of free primary education in sub-Saharan Africa," UNESCO background paper prepared for the Education for All Global Monitoring Report 2003/4 http://unesdoc.unesco.org/images/0014/001469/146914e.pdf, Online; accessed 2013-08-08.

Sandefur, J. (2013): "Africa Rising? Using Micro Surveys to Correct Macro Time Series," Center for Global Development, mimeo.

Scan News (2013): "Niger Spends N15bn On Free Education," http://scannewsnigeria.com/news/niger-spends-n15bn-on-free-education/, Online; accessed 2013-08-08.

Thomson, A., G. Eele, and F. Schmieding (2013): "Independent Evaluation of the International Household Survey Network (IHSN) and Accelerated Data Program (ADP) – Final report," .

UNICEF (2012): "Schools for Burkina Faso: Investing in the Future," http://www.schoolsforafrica.com/results/Burkina%20Faso%20Story.pdf, Online; accessed 2013-08-08.

United Nations (2013): "Communiqué," Meeting of the HighLevel Panel of Eminent Persons on the Post2015 Development Agenda in Bali, Indonesia, 27 March 2013.

WHO and UNICEF (2012): "Users Reference to Country Reports of WHO and UNICEF Estimates of National Infant Immunization Coverage," WHO and UNICEF working group for monitoring immunization coverage.

World Bank (2013): "Using Results-Based Financing to Achieve Maternal and Child Health - Progress Report 2013," http://rbfhealth.org/progressreport2013, Online; accessed 2014-05-01.

Yudkin, P. L., E. H. Burger, D. Bradshaw, P. Groenewald, A. M. Ward, and J. Volmink (2009): "Deaths caused by HIV disease under-reported in South Africa," AIDS, 23(12), 1600–1602.