



www.elsevier.com/locate/worlddev



CrossMark

World Development Vol. 80, pp. 48–60, 2016  
0305-750X/© 2015 The Authors. Published by Elsevier Ltd.  
This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

<http://dx.doi.org/10.1016/j.worlddev.2015.11.017>

# Reassessing Tax and Development Research: A New Dataset, New Findings, and Lessons for Research

WILSON PRICHARD\*

University of Toronto, Canada  
Institute of Development Studies, UK

**Summary.** — There is growing concern with the weaknesses of economic statistics relating to developing countries, and the risks that poor data have generated misleading research findings and poor policy advice. Cross-country tax data offer a striking example, with existing datasets frequently highly incomplete, analytically imprecise, plagued by errors, and sharply lacking in transparency. This paper introduces the new Government Revenue Dataset from the International Centre for Tax and Development, which provides a more reliable, transparent, and comprehensive basis for cross-national research. This new dataset has initially been used to re-examine major questions about the relationships between tax and aid, elections, economic growth, and democratization. The results deepen some previous conclusions and call others seriously into question—notably the assertion that aid dependence consistently undermines domestic revenue efforts. Above all, the research demonstrates the value of the new dataset, the broader sensitivity of many results to changes in data quality and coverage, and the consequent importance of expanded attention to, and investments in, data quality.  
© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Key words* — taxation, cross-country data, aid, resource curse, growth, political budget cycles

## 1. INTRODUCTION

Concern about the low quality of many economic statistics relating to developing countries has increased greatly in recent years (Jerven, 2013a). Key research findings and policy prescriptions may rest on fundamentally problematic foundations, driving a misleading view of development challenges and, at worst, misguided development policy. The landscape of development research is littered with studies relying on highly incomplete or questionable data, frequently with little acknowledgment of these limitations. Despite some progress, an earlier argument from Herrera and Kapur (2007) appears to still hold significant truth: “Inattentiveness to data quality is, unfortunately, business as usual. . . there are serious weaknesses in many datasets used in cross-country regressions currently in vogue. . . the data sets, problematic or not, become acceptable by repetition. . . There is a certain irony in the fact that a lot of work is devoted to improving methods, but that work on methods does not necessarily translate into improved everyday use of data” (p. 366, 382, 383).

These data weaknesses are particularly acute in relation to data on government revenues. Figures are readily available from the IMF Government Finance Statistics (GFS), and from a growing array of other official international sources. They are widely used in cross-country statistical analysis. However, they suffer major limitations: very extensive missing observations; unaccounted differences between alternative sources; conflicting and ambiguous treatment of revenues from natural resource extraction; and low transparency. In addition, the most common measure of overall revenue collection performance is the ratio of tax collection to GDP. But the GDP data series are themselves sometimes unreliable. In response, an increasing number of researchers have constructed their own revenue datasets on an *ad hoc* basis, often by going back to individual country-level sources and/or by merging data from multiple overlapping international sources. This had achieved some important gains. But it has often also

carried major costs: It has reduced the scope for comparison of results and replication, while many *ad hoc* datasets have themselves suffered from significant errors. Both problems have often been exacerbated by a troubling lack of transparency.

This article correspondingly introduces the new International Centre for Tax and Development Government Revenue Dataset (ICTD GRD). It is the outcome of a four-year process<sup>1</sup> of analyzing data from all available international data sources, along with IMF country reports, developing a standard system for classifying that data, and combining data from mutually compatible sources into a single research dataset. It achieves demonstrably large gains in both completeness and analytical accuracy; and, critically, is both publicly available and transparent in its construction.

Nor are these abstract gains. While improved data quality is desirable in and of itself, it also offers the promise that improved data will drive more reliable research findings. With this in mind, the remainder of the paper summarizes a first round of research findings employing the ICTD GRD, which collectively illustrate that improved data do, in fact, generate new and more robust results. These studies have sought to replicate earlier results in several widely explored research areas: (a) tax and aid, (b) tax revenue, non-tax revenue and democracy, (c) electoral cycles and taxation, and (d) taxation

\*Thanks to the International Centre for Tax and Development for funding this project, to Alex Cobham, Andrew Goodall, Fariya Mohiuddin, Vanessa van den Boogaard, Ava-Dayna Sefa and Kyle McNabb for invaluable contributions to constructing the dataset, to Paola Salardi, Paul Segal, Oliver Morrissey, Paul Clist, Samantha Torrance, Kyle McNabb and Philippe Le-May Boucher for leading the initial studies employing the new data, and to Mick Keen, Claudia Dziobek, Mike Seiferling, Jocelyn Pierre and others from the IMF and OECD who offered insight and support over the course of the project.  
Final revision accepted: November 6, 2015.

and growth. The results are striking. In two cases access to the new data adds substantial depth to earlier findings. In the other two cases employing the new data yields substantially new findings, to the point of calling earlier findings into question. More simply, the results illustrate, in stark terms, the value of ICTD GRD specifically, but also the broader importance for development research of much greater attention to data quality.

The paper proceeds as follows. The first three sections highlight the limitations of earlier data and the construction of the ICTD GRD. The fourth section provides an extended overview of the initial set of studies employing the new data, highlighting the new research results—and policy implications—that result. The final section notes some remaining limitations of the ICTD GRD, and reflects on the lessons that it provides for efforts to strengthen development data more broadly.

## 2. THE LIMITS OF INTERNATIONAL SOURCES

Researchers have long been troubled by the weaknesses of cross-country revenue data. The most widely used source prior to the public release of the ICTD GRD was the IMF GFS, which assembles official revenue and expenditure data from governments across the world. However, sharp limitations of the IMF GFS for research purposes have been increasingly recognized, with even IMF researchers turning elsewhere. The most important limitations of the GFS for research purposes fall into three categories: Inadequate data coverage, inconsistent treatment of natural resource revenues, and inconsistent GDP data employed to calculate tax-to-GDP ratios.<sup>2</sup>

With respect to *data coverage*, the IMF GFS suffers from extensive missing data, as data coverage amounts to only roughly 65 per of potential country-year observations for the period 1990–2010, and declines further for lower-income countries and in earlier years. This missing data are non-random, and often concentrated in countries experiencing instability, weak governance, or a poor relationship with international organizations, and thus risks generating misleading results.<sup>3</sup>

Additional concerns are raised by *inconsistent and analytically problematic treatment of revenues from the exploitation of non-renewable natural resources*. In general, countries record corporate taxes paid by private sector resource companies as taxes, while recording royalties, export taxes, profit sharing, the profits of state-owned enterprises, and similar revenues, as non-tax revenue. However, while these distinctions may be accurate from a pure accounting standpoint, they do not appear to be strictly adhered to; and, critically, they obscure the analytical distinctions relevant to most academic research.

Researchers are most often interested in the substantive distinction between what Moore (1998) terms “earned” and “unearned” income. “Earned income” refers broadly to non-resources taxes, which are raised on a relatively broad base, are not a payment for a specific service and generally require a measure of negotiation with the population as well as the construction of complex tax bureaucracies. By contrast, “unearned income” refers to natural resource revenues and other non-tax revenues that come from comparatively captive and concentrated sources, making collection relatively low cost and independent of local populations. When researchers speak of “tax revenue” they are most often interested in *non-resource* tax revenue. However, the IMF GFS, like most other sources, does not allow a consistent distinction between resource and non-resource sources of tax revenue.

Individual country examples serve to highlight these distinctions. When employing data from the IMF GFS, Angola reports tax collection ranging from 30% to 50% of GDP depending on the year—among the highest in the world. However, while accurate in an accounting sense, this is deeply misleading for much research: *non-resource* tax revenue amounts to around 5% of GDP in Angola—among the lowest in the world. This type of discrepancy is not uncommon across resource-rich states. In Iran prior to 1990, petroleum revenues were recorded as tax revenue in the IMF GFS. Post-1990, the same petroleum revenues are recorded as non-tax revenue. While the IMF warns explicitly against merging these pre- and post-1990 data series, some researchers have done so. In all such cases there is a major risk of highly misleading data generating unreliable research results.

Finally, many studies that rely on the IMF GFS—as well as other sources—have been undermined by *inconsistent GDP series employed to calculate tax ratios*. The root of these problems lies particularly, though not exclusively, in the irregular rebasing of GDP calculations in much of the developing world, which has resulted in the dramatic underestimation of GDP in many countries prior to rebasing. This has resulted in dramatic increases in recorded GDP—sometimes amounting to 50% or more—following rebasing exercises. Rebasing in Ghana in 2010 resulted in a 60% “increase” in GDP, while rebasing in 2014 in Zambia and Nigeria resulted in “increases” of 25% and 90%, respectively.<sup>4</sup> These episodes highlight the importance of regular rebasing, but also a much greater risk: international sources reporting GDP for the same country, but using different base years, can result in huge jumps in GDP from one year to the next (Jerven, 2013b), and correspondingly sharp (and entirely illusory) declines in tax-to-GDP ratios as a result.

This has been true, for example, of the IMF International Finance Statistics (IFS). At the time of writing, combining tax data from the IMF GFS with GDP data from the IMF IFS (as has been common) yields a tax to GDP ratio for Ghana of over 20% in 2005 (pre-rebasing) and less than 13% in 2006 (post-rebasing). In Iran GDP similarly became progressively underestimated over the course of the 1980s, leading the tax to GDP ratio reflected by international statistics to rise from 26% in 1982 to 119% in 1989. Meanwhile, data for Iran in 1990 witness two shifts: reliance on an updated and more realistic GDP series, and a move from recording petroleum revenue as tax revenue to recording it as non-tax revenue. The result: a fall in the recorded tax to GDP ratio from 119% to less than 6% in a single year.

These are, of course, extreme examples, but they are representative of more widespread, though less extreme, issues. They are, as importantly, indicative of insufficient attention to data quality in much existing research, and of the need for transparency of research datasets and their construction in order to weed out such potential problems.

## 3. RESEARCHERS' DATASETS AND THEIR LIMITATIONS

The problems noted so far have not gone entirely unnoticed. In response, researchers have increasingly relied either on regional sources, which often offer more complete data within a smaller subset of countries, or on *ad hoc* researcher datasets, which merge different data sources, clean existing data and/or draw on country-level sources in order to achieve improvements. However, these initiatives have ultimately failed to sufficiently address existing challenges, and highlight the importance of more comprehensive strategies.

Reliance on regional datasets has been the most straightforward response. Regional datasets from the OECD, the *Comisión Económica para América Latina y el Caribe* (CEPAL), and the African Economic Outlook (AEO) all offer more complete coverage within their specific regions. However, they are also characterized by significant limitations. Most obviously, regional datasets are limiting for research questions more usefully addressed using a global sample, and they limit comparability across studies. A number of less obvious issues also arise. First, regional datasets often report on total tax revenues, but fail to record non-tax revenue and, by extension, total government revenue, thus providing an incomplete picture of government revenues.<sup>5</sup> Second, the regional datasets generally suffer from the same limitations as the IMF GFS in relation to the classification of natural resource revenues. Third, the AEO dataset in particular lacks clarity about underlying sources of both tax and GDP data, while a significant part of the data matches other international sources only very imperfectly. Finally, small differences in method across sources reduce direct comparability. For example, revenue is categorized somewhat differently between the CEPAL tax dataset and the OECD Revenues Statistics in Latin America dataset for some countries—despite covering the same countries, and being based on the same underlying data.

The alternative to regional datasets has been comparatively *ad hoc* efforts by researchers themselves to construct new datasets by merging data from multiple sources and/or relying on country-level sources. These datasets have offered notable benefits, but have failed to fully address key limitations of available data. While individual datasets achieve significant gains in coverage (e.g., IMF, 2011) and in the consistent treatment of natural resource revenues (e.g., Haber & Menaldo, 2011; Keen & Mansour, 2009), none achieve these two goals simultaneously.<sup>6</sup> Meanwhile, these *ad hoc* datasets have generated significant new concerns. The proliferation of alternative datasets has further reduced comparability across studies, while making replication and verification difficult or impossible. This is particularly true where researchers have merged data from multiple sources, but without careful documentation of those changes.<sup>7</sup> This has been exacerbated by the fact that many of the datasets have not been made publicly available, thus preventing easy replication and dramatically increasing the risk of unintentional errors in data construction, given the absence of external scrutiny.<sup>8</sup>

The best illustration of the potential and pitfalls of privately held research datasets has been a dataset constructed by the IMF Fiscal Affairs Department (IMF FAD). The core aim of the IMF FAD dataset was to fill gaps in the IMF GFS dataset by drawing on additional international sources, most notably the OECD, the World Economic Outlook, and IMF Article IV reports.<sup>9</sup> The approach was superficially sound: Where the preferred source was not available, missing data points would be filled by drawing on the “next best” alternative source. The result was a dataset that achieved dramatically improved data coverage, and which became the basis for both research papers (e.g., Benedek, Crivelli, Gupta, & Muthoor, 2014) and official IMF FAD publications (e.g., IMF, 2011).

However, analysis undertaken during the construction of the ICTD GRD revealed three major issues. First, and by design, the FAD dataset did not seek to arrive at a common treatment of natural resource revenues, and thus suffers from the same inconsistencies and analytical limitations as the IMF GFS and other sources. Second, the FAD dataset failed to address problems with underlying GDP series, and thus

occasionally witnessed large breaks in the data series resulting from GDP rebasing.<sup>10</sup>

Third, and most worryingly, the merging of data from multiple sources inadvertently introduced significant inconsistencies and jumps in the data series.<sup>11</sup> The basic assumption underlying the IMF FAD dataset’s construction was that different sources were always mutually compatible; that is, where there was a missing data point in one dataset it could be replaced from an alternative source. However, this basic assumption does not hold, and has plagued other researcher datasets as well. In many cases different sources report different data in overlapping years owing to differences in, among others, methodology and classification, the level of government being considered, the treatment of natural resources, the treatment of social contributions, and underlying GDP series. By failing to systematically account for these inconsistencies, the merging process introduced jumps in the data large enough to call any related research findings into question. Meanwhile, the fact that the dataset was not publicly available almost certainly contributed to these weaknesses remaining overlooked—even as the dataset provided the basis for high profile and influential research findings.

To their credit, the IMF FAD were eventually willing to share the data with us, and to engage with us during construction of the ICTD GRD. More importantly, as this article was being finalized the FAD made an updated version of their internal research dataset public for the first time, newly named the World Revenue Longitudinal Dataset (WoRLD). The decision was long in coming, but marked an important step forward in seeking to make research data more transparent, widely available, and subject to public scrutiny.<sup>12</sup> The WoRLD dataset also included important improvements over earlier iterations of the FAD dataset, including reducing the number of breaks in the data series and employing a consistent GDP series, similar to the ICTD GRD. However, despite important advances, the new dataset still falls significantly short of the ICTD GRD as of the time of writing. Most notably, it *does not* seek to address the treatment of resource revenues, there remain significant breaks in the data series when incompatible sources are merged, and the total data coverage remains about 15% less complete than the ICTD GRD. Simply put, while somewhat improved, the IMF WoRLD continues to suffer significant limitations, while the ICTD GRD remains by a significant measure the best source of data for most types of research.<sup>13</sup>

The purpose of highlighting the limitations of the earlier IMF FAD dataset, and the more recent WoRLD dataset, is not to “expose” earlier failures. The goal, instead, is to illustrate broader issues related to data quality, and the corresponding integrity of research findings. Despite the best intentions, *ad hoc* efforts to construct new datasets—of which there have been many examples—pose enormous risks. These risks are dramatically amplified when, as has been common, these datasets are not made publicly available for external scrutiny. In practice, existing efforts have failed to sufficiently improve the data foundation for tax and development research, and in some respects may have in fact made the situation worse through the proliferation of alternative datasets. The recent public release of the IMF WoRLD dataset is very encouraging as a signal about future engagement and transparency. But its limitations reinforce the argument here about the importance of transparency and external scrutiny, and reaffirm the overall limitations of available data for tax research. It is these limitations that have motivated the creation of the ICTD GRD, and its emphasis on public availability and external scrutiny.

#### 4. THE INTERNATIONAL CENTRE FOR TAX AND DEVELOPMENT GOVERNMENT REVENUE DATASET

The construction of the ICTD GRD was initiated in late 2010 to address the most glaring weaknesses of existing international data sources, and thus offer a more reliable foundation for cross-country research. Critically, the goal from the outset was to rapidly make the data publicly available, in order to encourage research based on a common data foundation, amenable to comparison and replication, and in order to invite external use and scrutiny as a means to protect against errors.

The ICTD GRD has been constructed by systematically merging, cleaning, and reshaping data from multiple international and country-level sources, in order to maximize coverage while removing errors associated with existing sources. While this approach mirrors some existing efforts—most notably the work of the IMF FAD described above—the ICTD GRD is significantly more ambitious, transparent, and precisely implemented. Construction of the dataset involved five key stages, which are described briefly here, and in greater detail in Prichard, Cobham, and Goodall (2014).

First was the creation of a standard revenue classification into which data from different sources could be incorporated. Five elements are worth highlighting. First, the classification is extensive: It includes all of the major tax categories from the IMF GFS. Second, it includes both tax revenue and *total government revenue*, thus offering a complete view of the fiscal resources of governments. Third, it includes aggregate categories for *direct and indirect taxes*, as well as their subcomponents. These aggregate categories allow slightly more complete time series, and overcome potential inconsistency within and across countries in allocating revenue across the subcomponents of direct and indirect taxes. Fourth, and most critically, the classification systematically distinguishes between *non-resource* tax revenue and revenue from non-renewable natural resources. And, fifth, the classification reports data inclusive and exclusive of social contributions, in order to accommodate differences in reporting across sources.

The second stage was to compile data from all major international sources into this standard format, as a precursor to merging that data into a “first choice” composite dataset. This included data from the IMF GFS, the OECD Tax Statistics, the OECD Revenue Statistics in Latin America, CEPAL, the African Economic Outlook Africa Tax Dataset, the World Bank World Development Indicators, and the Michigan Ross World Tax Database. Alongside these datasets, we also manually collected data from every publicly available IMF Article IV report. Critically, the Article IV reports often contain more

detailed data than do international datasets—particularly with respect to natural resource revenues—and also often contain data for country years that are not covered by the IMF GFS and other available sources, as illustrated in Figure 1.<sup>14</sup>

The third and most challenging step lay in *merging data from these multiple sources* in order to arrive at a more complete and analytically accurate “first choice” dataset. In principle, this is a straightforward task: Employ data from a “baseline source” where appropriate, and then fill gaps in that series where necessary from alternative sources. However, in practice significant discrepancies across sources means that such mechanistic merging generates major breaks in the data series, as with the original IMF FAD dataset, described above. As such, the process is necessarily manual: Inspection of all available data for a given country, selection of the best available “baseline” source for the country, and filling gaps using other sources if, and only if, the sources are clearly compatible. In general, the “best” source in this context is the one which best combines temporal coverage, detailed data disaggregation, and consistency with other sources.

The only partial exception is in resource-rich countries, where there is a need to rely only on sources that offer clear disaggregation between the resource and non-resource components of tax revenue. In practice, this significantly reduces the range of potential sources for resource-rich countries, and the ICTD GRD dataset draws overwhelmingly on IMF Article IV reports, which offer the most consistent disaggregation.<sup>15</sup> Meanwhile, the non-resource tax revenue variable is left missing where the resource component of tax revenue cannot be isolated, as a missing data point is preferable to one that is clearly misleading.

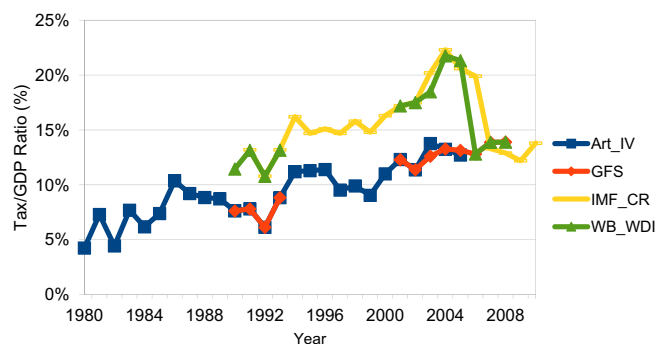


Figure 2. Ghana: Total tax as% of source-specific GDP. Source: Data compiled by the authors.

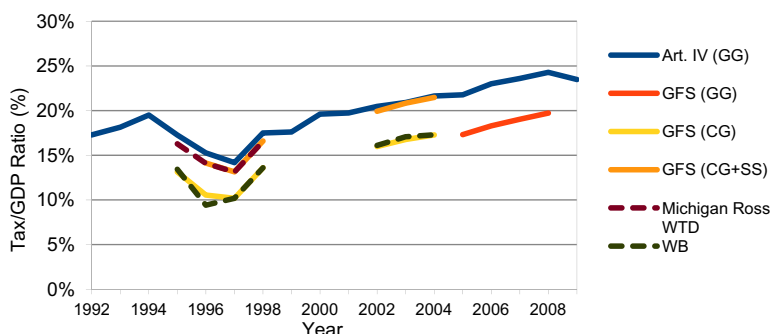


Figure 1. Additional data coverage using Article IV data: Albania 1992–2009. Source: Data compiled by the authors. GG denotes General Government data, CG denotes Central Government data and SS denotes social contributions.



The result is the successful closing of “gaps” in individual data series, as illustrated in Figure 2. However, Figure 2 also reveals sharp differences across sources in the levels of the tax to GDP ratio. This is driven by difference in the underlying GDP series. As discussed earlier, employing source-specific GDP series when merging data from multiple sources can lead to major inconsistencies, the worst of which result from a failure to update historical GDP series following GDP rebasing exercises.

These discrepancies motivated the fourth step in the process: *the construction of a standard GDP series* that would smooth these differences. Rather than relying on source-specific GDP series, all of the tax figures were compiled in local currency units and merged with a common GDP series, which relied on a single, and most up-to-date, GDP base year for each country. This common GDP series was drawn primarily from the WEO database, with missing data occasionally filled from other sources where we could verify that the GDP series were compatible.<sup>16</sup> The result is illustrated in comparing Figure 2, above, with Figure 3, below: the resolution of large apparent discrepancies in tax ratios across countries.

Finally, most existing studies have drawn exclusively on central government data, which are comparatively widely available. However, central government data vastly understate actual government revenue in comparatively decentralized countries. As such, the ICTD GRD also gathers general government (the sum of national, state and local) data wherever available. In turn, the final dataset employs general government data for more decentralized countries, while continuing to employ central government data for comparatively centralized countries (which generally do not report subnational data). While this approach moderately underestimates government revenue in centralized states, this is far preferable

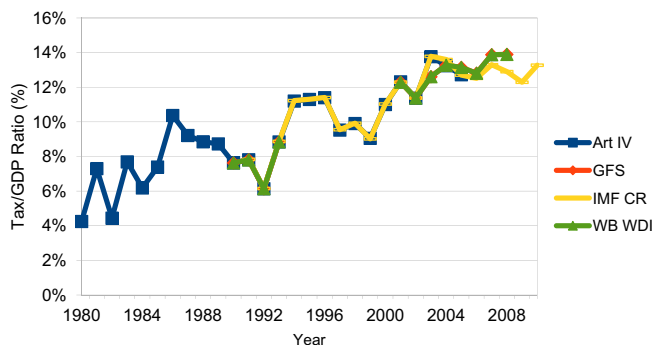


Figure 3. Ghana: Total tax revenue as% common GDP series. Source: Data compiled by the authors.

to the large errors implied by relying solely on central government data.<sup>17</sup>

The end result is a dataset that is significantly more complete and analytically accurate than previous options. The new dataset covers almost 70% more developing country observations than the IMF GFS (Table 1), and addresses the major sources of inaccuracy in earlier datasets: inconsistent treatment of natural resources, inconsistent GDP figures, problematic merging of data, and lack of attention to subnational revenue. And, as critically, the dataset, including details of its construction and revisions, is publicly available.

In turn, these improvements in the data yield clearer insights into global trends in tax collection than has previously been the case—and, again, this is particularly true for developing countries. While these trends are covered in greater detail elsewhere (Prichard, Cobham *et al.*, 2014), a few key messages warrant particular attention.

First, the ICTD GRD reveals consistent increases over time in non-resource tax collection in the developing world, across all regions. While this broad message has been highlighted in earlier IMF reports (e.g., IMF, 2011), the higher quality data available from the ICTD GRD reveals a clearer and more pronounced trend (Figure 4). Meanwhile, consistent with most commentary on the topic, the data reveal that the largest increases in collection have come from expanded taxation of goods and services, while trade taxes have declined significantly. That said, the data reveal significant gains in direct taxes as well, offering some optimism about the capacity of developing country tax administrations to tackle these politically and technically complex taxes (Figure 5).

Second, despite these gains, total tax collection in low-income countries still very frequently falls below even the minimum levels advocated by international organizations. It has long been common, though somewhat arbitrary, to suggest that countries should raise a minimum of 15% of GDP in domestic tax revenue, and some alternative data sources have implied that most developing countries have reached this level. However, by employing a more consistent and up to date GDP series, and focusing only on *non-resource* tax revenue, we find that the average level of non-resource tax collection in low-income countries is only 13% of GDP, with fully 65 countries collecting less than 15% of GDP in non-resource taxation.<sup>18</sup> The ICTD GRD data thus highlight not only recent gains, but the very large challenges that remain.

Third, the ICTD GRD distinguishes between non-resource taxation and all other types of revenue, thus making it possible to compare trends in non-resource tax collection to trends in the collection of non-tax revenue (which here is taken to include *all* resource revenues). This comparison highlights one of the key arguments in favor of robust domestic taxation:

Table 1. ICTD GRD data coverage, relative to other source, 1990–2010 Source: Authors’ calculation. Each cell captures the number of available country-year observations captured by each dataset.

		ICTD GRD	IMF FAD	IMF WoRLD	IMF GFS	WDI
All Developing countries	Total revenues	2,317	1,913	2,134	1,391	1,060
	Total tax	2,348	1,976	1,876	1,396	1,060
	Tax on income, profit, and capital gains	1,900	1,909	1,608	1,395	1,060
	Tax on goods and services	1,952	1,856	1,435	1,395	1,060
All countries	Total revenues	3,465	3,022	3,130	2,070	1,581
	Total tax	3,513	3,071	2,783	2,091	1,581
	Tax on income, profit, and capital gains	2,956	3,040	2,530	2,089	1,581
	Tax on goods and services	3,035	3,025	2,262	2,089	1,581

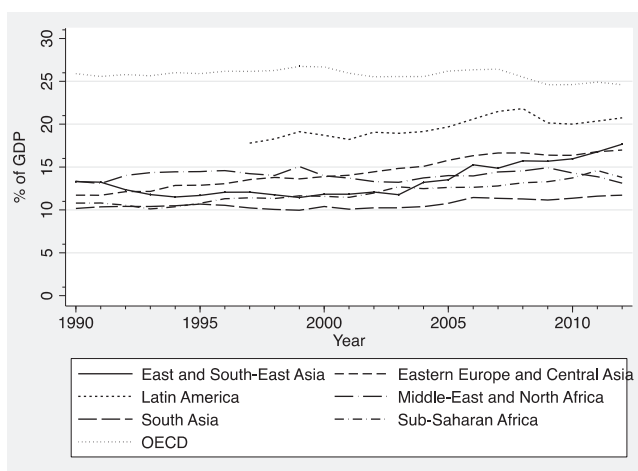


Figure 4. Total Tax Collection as % of GDP by Region. Source: ICTD GRD (2015). Trend lines include all countries with data available for at least 15 years during the time period. Date for Eastern Europe and Central Asia begin in 1997 owing to major data problems during and after the transition from Soviet control.

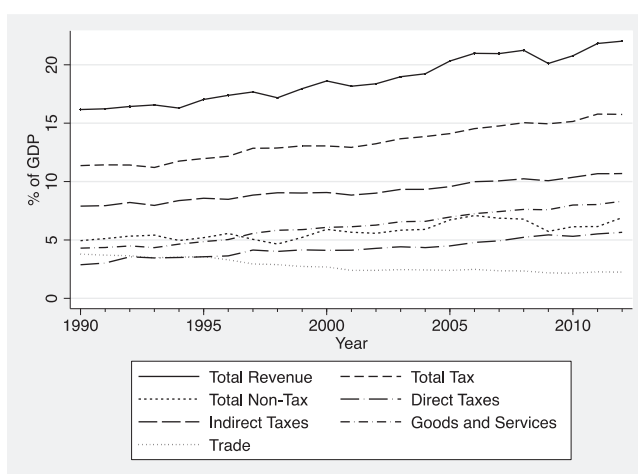


Figure 5. Composition of Government Revenue for Developing Countries. Source: ICTD GRD (2015). Trend lines include all countries with data available for at least 15 years during the time period.

greater predictability. Following the global financial crisis of 2008–09 non-resource tax collection remained relatively stable across developing countries, and quickly resumed its upward trend. By contrast, non-tax revenues—of which the most important and volatile component is resource revenues—declined significantly in 2009, with only a very limited recovery in subsequent years, thus creating significant fiscal stress in countries that lacked robust non-resource taxation (Figure 5).

## 5. IMPROVED DATA, NEW RESULTS

The ICTD GRD’s improvements in data coverage and accuracy thus offer a more precise descriptive picture of tax collection around the world. However, a larger question remains unanswered: does employing better data also produce different and more robust econometric research findings? With this question in mind, the remainder of this paper reports the

results of the first four studies conducted employing the new ICTD GRD. Each study began with a simple question: Do existing research findings appear to be robust to employing the new ICTD GRD? Each of the studies correspondingly revisits well-known research findings, and employs relatively standard methods from the existing literature, in order to explore the empirical benefits of employing improved data.<sup>19</sup>

The results make clear the value of the new data. In the first two cases, employing the ICTD GRD casts doubt on the robustness of earlier findings (relating to aid and taxation, and to the impact of elections on tax collection). In the third case, the new data help to resolve a long-running empirical controversy about the existence of a “political resource curse”, with the new results significantly more consistent than earlier studies. Finally, in the fourth case the new data make it possible to explore new dimensions of the relationship between tax structure and growth, and to uncover important new findings.

## 6. AID AND DOMESTIC TAX EFFORT

Over the past decade, research looking at the impact of foreign aid on domestic tax collection has generated sharply conflicting findings. It has focused on a straightforward hypothesis: Countries that receive higher levels of foreign assistance will collect less domestic tax revenue, owing to weaker incentives to pursue politically costly local tax collection. This highly intuitive possibility found its strongest early support in work by Gupta, Clemens, Pivovarsky, and Tiongson (2004). These IMF authors published influential econometric results demonstrating that foreign grants have a negative impact on domestic tax collection, while loans—which eventually need to be repaid—have a positive impact on tax collection. The finding was not only of academic interest, but had major policy implications. It offered support to policy voices calling for greater reliance on loans rather than grants in delivering foreign assistance, and for the expansion of donor conditionality targeting increased tax collection in order to offset the negative incentive effects of grants.<sup>20</sup>

However, skepticism about these research findings was immediate, prompting a range of subsequent studies. At least one such study replicated the negative impact of grants specifically (Thornton, 2014), while others offered support for a generally negative impact of aid on tax collection (Bräutigam & Knack, 2004; Remmer, 2004). However, other studies presented evidence that there was likely no consistent effect of aid on domestic tax collection (Clist & Morrissey, 2011), or that any relationship was contingent on the quality of institutions in recipient countries (Brun, Chambas, & Guerinneau, 2008). These initial results sparked a subsequent round of studies, and continued disagreement. Members of the original team from the IMF—along with new collaborators—employed the newly created IMF FAD dataset, described above, to revisit the results. They again reported a negative impact of grants on tax collection, and a positive impact of loans—though both smaller in magnitude than in their original study (Benedek et al., 2014). These results again elicited an immediate response. Carter (2013) argued that the results were not robust to more appropriate econometric methods, while Clist (2014) found that the results could not be fully replicated using the listed data sources, which he attributed in part to significant errors—described earlier—in the IMF FAD dataset.<sup>21</sup>

The findings reported in Clist (2014) in particular suggest that data problems may be critical to explaining divergent

findings across studies.<sup>22</sup> They also hold a more subtle lesson: Continued empirical disagreement has been facilitated by each set of researchers employing their own datasets, while the [Benedek et al. \(2014\)](#) dataset was not made publicly available. With this in mind, [Morrissey, Prichard, and Torrance \(2014\)](#) and [Morrissey and Torrance \(2015\)](#) re-examine the relationship between aid and taxation employing the ICTD GRD. The former employs four-year averaged tax data, in order to smooth year-to-year shocks, while the latter employs annual data. Both studies are based on a standard model specification:

$$\begin{aligned} \ln(\text{Tax Revenue}_{it}) = & \beta_0 + \beta_1 \text{gdppc}_{it} + \beta_2 \text{agr}_{it} + \beta_3 \text{ind}_{it} + \beta_4 \text{imports}_{it} \\ & + \beta_5 \text{exports}_{it} + \beta_6 \text{Aid}_{it} + \beta_7 \text{Aid}_{it}^2 + \varepsilon_{it} \end{aligned} \quad (1)$$

In the model, tax revenue is a function of a set of standard determinants of tax collection, and of the level of aid. Both studies include a squared aid term to account for potential non-linearity, while the aid term is also broken into its component parts, grants, and loans. The model is tested employing both contemporaneous levels of aid, and a one-period lag of the aid variable. The contemporaneous version is favored by [Benedek et al. \(2014\)](#), but the lagged version is preferred by most other studies, based on a belief that, because tax collection changes slowly, high levels of aid are more likely to affect levels of tax collection one or more years in the future.

Having established the model, the authors test the relationship using a range of econometric methods from the existing literature.<sup>23</sup> Reflecting a focus on assessing the robustness of earlier findings, they focus on comparatively straightforward models: fixed effects, a lagged dependent variable (LDV) model with panel-corrected standard errors (PCSEs) and Feasible Generalized Least Squares. The most recent study by [Benedek et al. \(2014\)](#) also argues for employing the System-GMM estimator in order to account for possible endogeneity. This approach has been criticized elsewhere ([Carter, 2013](#)), but it is widely used, and is thus also tested.

While constraints of space do not allow for a full review of the results, the basic message from both studies is straightforward: There is no consistently significant relationship between aid and tax collection. [Morrissey et al. \(2014\)](#) find that in the specifications where they are significant, net aid, grants, and loans are more often positively associated with government tax revenues, although the significance is often weak and not robust. This is true when using either contemporaneous or lagged aid variables, and is true when the sample is restricted to countries in sub-Saharan Africa.

The study also demonstrates more directly the value of the improved data. Existing studies have asked whether access to aid reduces tax effort, but have not included a right-hand-side variable to account for the impact of access to non-tax revenue—including natural resource revenues—on tax collection. This is a surprising omission, given widespread evidence that natural resource wealth reduces tax collection ([Crivelli & Gupta, 2014](#)). This implies a significant risk of omitted variable bias in *all* earlier results, which appears to have been driven entirely by the inadequacy of available data. In any event, the ICTD GRD allows for the inclusion of a control for non-tax revenue and, as expected, it is negative and highly significant in all of the econometric models. Because the results are in any case generally insignificant across specifications, adding this new variable does not alter the broad pattern. However, the absence of this critical variable from earlier studies again points to problems of robustness arising from inadequate data.

Ultimately, the results indicate that earlier IMF results are highly sensitive to choices about data and econometric speci-

fication. In turn, access to the ICTD GRD allows a common foundation for interrogating this sensitivity, and suggests that earlier findings of a significant relationship are likely misleading. This is quite intuitive: Foreign aid may sometimes create disincentives to domestic tax collection, but this effect is likely to be muted by several factors. “Foreign aid” is a hugely diverse category. Much aid does not go directly to recipient governments, and is anyway an imperfect substitute for own-revenue in the eyes of recipient governments. Meanwhile, technical assistance, donor conditionality, and related pressure for the expansion of services may all generate a countervailing positive impact of aid on tax collection. Through access to improved data, this more nuanced relationship is brought to the fore.

## 7. ELECTORAL CYCLES AND TAXATION

While less controversial than debates over aid and taxation, there has been similar disagreement within the existing literature about the impact of elections on pre-election tax collection. Models of “political budget cycles” predict that tax collection will decline—and public spending will increase—prior to elections, as opportunistic politicians seek to secure short-term political support ([Drazen, 2001](#)). This possibility has spawned a large set of cross-country studies.

With respect to public spending, there has been significant agreement that elections do, in fact, spur increased pre-election spending, with a somewhat larger effect in the generally weaker institutional environments of developing countries ([de Haan & Klomp, 2013](#)).<sup>24</sup> By contrast, there has been little agreement about taxation. Several studies have found that elections increase pre-election spending but have no impact on pre-election revenue collection (e.g., [Brender & Drazen, 2005](#)). This is presumed to reflect the comparative difficulty and cost of short-term reductions in taxation.<sup>25</sup> By contrast, a recent study by [Ehrhart \(2013\)](#) reports evidence that elections have had a negative impact on pre-election tax collection, amounting to about 0.2% of GDP, and concentrated among more broad-based indirect taxes.

However, there are again grounds for concerns about data quality, reflected in two illustrative examples. The most widely cited evidence that elections *do not* reduce pre-election tax collection comes from [Brender and Drazen \(2005\)](#), who draw on panel data from the IMF covering 68 countries from 1960 to 2001. To their significant credit, they noted the imperfections of IMF data, attempted to “clean” the data, and, most notably, made their dataset publicly available along with careful documentation. However, their data nonetheless suffer from many of the issues noted here: limited coverage of low-income countries, imperfect treatment of natural resources, unexplained jumps in the data series from one year to the next, and limited comparability to other studies. The more recent study from [Ehrhart \(2013\)](#) faces similar issues. Recognizing the limitations of existing data, she constructs a “cleaned” dataset of 56 developing countries from a combination of the IMF GFS and IMF Article IV reports. However, there remains significant missing data, while, unlike [Brender and Drazen \(2005\)](#), she did not post the data publicly, nor provide detailed documentation.

Collectively, these limitations motivated [Prichard \(2014, in press\)](#) to revisit these earlier findings using the ICTD GRD, in order offer more robust evidence based on more complete and transparent data. Following the literature, the study employs a simple model in which tax collection is a function of a battery of standard control variables,  $X$ , fixed effects,  $\alpha$ ,



time dummies,  $\lambda$ , and a dummy variable for the presence of an election, *Election*:

$$Tax_{i,t} = \varphi Tax_{i,t-1} + \beta_1 Election_{i,t} + X'_{i,t-1} \delta + \alpha_i + \lambda_t + e_{i,t} \quad (2)$$

The model is, again following the literature, tested using both a standard fixed-effects model and the System-GMM estimator. The results are unambiguous, finding no consistent evidence of a statistically significant impact of elections on pre-election tax collection. This remains true using both estimators, over different time periods, employing different subsets of countries and when disaggregating total tax collection between direct and indirect taxes. Simply put, the new results suggest that the statistically significant results reported by Ehrhart (2013) are highly sensitive to the completeness and quality of the data.

However, the value of the improved ICTD GRD data lies not only in re-evaluating existing results, but making it possible to push existing analysis further. To this end, Prichard (2014, in press) subsequently asks whether specific *types of elections* have statistically significant effects on pre-election tax collection, even if elections *in general* do not. The results indicate that *competitive elections* (in which the winning party receives less than 60% of legislative seats) are unique in having a statistically significant negative impact on pre-election tax collection.<sup>26</sup> This finding is, again, robust to a wide range of alternative choices about data and sample. This finding is not only relevant to understanding taxation, but also has implications for democratic theory, by implying that more competitive elections have a stronger impact on the incentives of political leaders. The bottom line, again, is that access to the new data makes it possible to reassess earlier results and to open up new lines of research inquiry—while making all of those new results open to replication and scrutiny.

## 8. TAXATION, NON-TAX REVENUE, AND DEMOCRACY

The relationship between tax revenue, non-tax revenue (primarily from non-renewable natural resource revenues), and democracy is another research question that has elicited high-profile disagreement across an expansive range of cross-country econometric studies. And, again, recent research employing the ICTD GRD suggests that this disagreement has, at least in part, been driven by the weaknesses of previously available government revenue data.

Research in this area has centered on the existence (or not) of the so-called “political resource curse”. That is, on the possibility that governments with access to extensive revenues from non-renewable natural resources are less likely to be democratic than those that rely primarily on raising tax revenue from their own citizens. Governments with access to extensive “unearned” revenue from captive natural resources have weaker incentives to be accountable: They have less need to secure revenue from, and thus to bargain with, their citizens, while they have extensive revenue to invest in repression or in purchasing political support through public services, transfers, and patronage. By contrast, a government reliant on local tax revenue has incentives to be more accountable in order to encourage “quasi-voluntary” tax compliance, while it is likely to face greater demands for reciprocity from citizens interested in ensuring that their tax payments are used effectively.

The clarity of the hypothesized relationship appears to make this question particularly amenable to cross-country econometric testing, and such studies have dominated existing

research. However, existing studies have yielded highly varied results. Early results almost universally uncovered the expected negative relationship between natural resource wealth and democracy (e.g., Jensen & Wantchekon, 2004; Ross, 2001, 2004). However, subsequent studies argued that the finding was not robust, and that there was, in fact, no significant relationship—particularly when focusing on within-country variation over time (Dunning, 2008; Haber & Menaldo, 2011; Herb, 2003). These dissenting findings have, in turn, been met by further studies reporting a negative impact of resource wealth on democracy (Andersen & Ross, 2014; Aslaksen, 2010; Wiens, Poast, & Clark, 2014).

Owing to the lack of alternatives, these studies have often relied on problematic data of two different types. Some studies have relied on highly incomplete and imperfect government revenue data from the IMF GFS, the WB WDI, and regional sources, the consequences of which have already been discussed at length. Meanwhile other studies, recognizing the limitations of available government revenue data, have opted instead to rely on measures of total *resource income*—that is, the total value of resource extraction in a country, which is more widely available, but only an imperfect proxy for the theoretical propositions of interest. Most theories of the political resource curse center on government *revenue*: a government that has greater access to *revenue* from natural resources is less likely to be accountable to its citizens, owing to a reduced need to collect *tax revenue* from citizens, and access to *revenue* to invest in repression and securing support. While *resource income* is a plausible proxy for government *revenue* from nature resources, the connection is highly imperfect, thus subtly distorting the research question.<sup>27</sup>

Against this background, the ICTD GRD offers a major step forward, as it allows researchers to implement more direct tests of the relationship between the composition of government revenue and democracy. Prichard, Salardi, and Segal (2014) correspondingly re-test the most widely cited econometric models from the existing literature using the GRD. As with the bulk of the literature, they begin with a standard dynamic specification:

$$Democracy_{i,t} = Democracy_{i,t-1} + \beta_1 Tax\_rel_{i,t-1} + X'_{i,t-1} \gamma + \alpha_i + \lambda_t + e_{i,t} \quad (3)$$

In this model, the level of democracy is a function of existing institutions (the lagged level of democracy), a revenue variable, a battery of control variables—including income, economic growth, the existence of conflict, and regional democratic diffusion—fixed effects,  $\alpha$ , and time dummies,  $\lambda$ .

They employ three alternative government revenue variables to reflect slightly different theoretical claims. First, they test a measure of *tax reliance*, which is the share of non-resource tax revenue in total government revenue. This captures the idea that a government that is relatively more dependent on tax revenue is more likely to be accountable to its citizens. Second, they split tax reliance into its component parts, non-resource *tax revenue* and total *non-tax revenue*, of which natural resource revenue is the dominant component across countries. Both are, following theory, measured as shares of GDP, and this second set of tests attempts to distinguish the *governance-enhancing* benefits of domestic taxation from the *negative impacts* on governance of high levels of non-tax revenue, primarily from natural resources.

Reflecting the diversity of the existing literature, each of these relationships is tested using five alternative econometric methods: System-GMM, the Common-Correlated Errors Mean Group estimator developed by Pesaran (2006), the



Error Correction Mechanism regressions favored by Haber and Menaldo (2011), and random and fixed effects logit models, which follow Wiens *et al.* (2014). Each method has its limitations, thus motivating the decision to test them side-by-side in order to assess the robustness of the results.

While there is, again, no space to review the results in detail, they offer relatively consistent support for the existence of a political resource curse. Both the *tax reliance* variable and the variable for *total non-tax revenue* are statistically significant with the expected sign across almost all of the estimations. Meanwhile, the magnitude of the estimates is relatively large, suggesting that the composition of government revenue is a powerful determinant of governance outcomes.<sup>28</sup> Finally, and critically, the results prove generally robust to changes in time period, country coverage, control variables, and the level of government under consideration. The greater consistency of the results using the ICTD as compared to earlier studies thus offers analytical clarity, and a clear indication of the value of improved data.

Meanwhile, a series of more specific tests further supports the belief that employing higher quality data can contribute to more reliable findings. When the time series is extended back into the 1980s, and thus to include less reliable and complete data, the strength and significance of the results gradually declines. When the results are tested using central government data exclusively, instead of the more robust final version of the ICTD GRD, again the results are generally weaker, though still significant in most cases. Finally, and most tellingly, most of the results are more statistically significant and consistent when employing the ICTD GRD measures of *government revenue* than when running the same tests using the *resource income* variables from Haber and Menaldo (2011). In each case, more complete data that better reflects the theoretical propositions of interests yields stronger results, consistent with claims about the value of the ICTD GRD.

## 9. TAXATION AND GROWTH

The final subject of the initial round of research has been the relationship between taxation and growth. Existing literature has, again, focused on a relatively straightforward prediction: All else equal, growth will be slower in countries that are more reliant on taxes that are expected to be comparatively economically distorting: corporate income taxes, personal income taxes, and trade taxes.<sup>29</sup> The broad proposition that marginal increases in these taxes is bad for growth has found relatively consistent support in a series of existing cross-country econometric studies. More strikingly, it has come to be widely viewed as policy orthodoxy, reflected in international support, led by the IMF and the World Bank, for reductions in income and corporate tax rates in developing countries, and for replacing trade taxes with value-added taxes (IMF, 2011).<sup>30</sup> However, the empirical foundation for these views has been surprisingly limited: cross-country research has been largely limited to more developed economies, with data limitations having restricted efforts to explore the relationship in detail in the developing world.

A survey of the literature by McNabb and LeMay-Boucher (2014) concludes that only one cross-country study testing the tax-growth relationship using tax collection data has included a large range of developing countries.<sup>31</sup> In that study, Acosta-Ormaechea and Yoo (2012) employ an *ad hoc* dataset created specifically for their research, which draws on data

from the IMF GFS, OECD, and CEPAL, and which covers 25 middle-income and 23 low-income countries. However, while their decision to construct an *ad hoc* dataset results in improved data coverage, it comes with all of the costs discussed previously: lack of comparability and replicability, and the risk of undetected errors or bias in the sample. Both problems are exacerbated by the fact that the dataset is not publicly available.<sup>32</sup>

Against this background, McNabb (2014) re-examine the relationship between tax collection and economic growth using the ICTD GRD. Following Acosta-Ormaechea and Yoo (2012), and others, they estimate the impact of revenue-neutral changes in tax structures on growth. They employ a Common Correlated Effects Mean Group (CMG) estimator (Pesaran, 2006), which allows for parameter heterogeneity and controls for cross-sectional dependencies.

They begin by focusing on the question that has dominated the relevant literature: Does a revenue-neutral shift away from consumption taxes in favor of income taxes have a negative effect on growth? Consistent with Acosta-Ormaechea and Yoo (2012), they find a modest but statistically significant effect of such a shift: a one percentage point revenue neutral shift from trade taxes to personal and corporate income taxes reduces growth by between 0.1% and 0.2% depending on the specification, with increases in personal income taxes having the largest effect. That said, the results are somewhat inconclusive when the focus is on low-income countries alone, with the strongest results driven by middle and high-income countries.

The overall results are thus broadly consistent with Acosta-Ormaechea and Yoo (2012). However, a key advantage of the ICTD GRD is not only greater completeness and analytical accuracy, but also the ability to explore a broader range of questions. With this in mind, McNabb and LeMay-Boucher explore a second question of particular interest to developing countries: Have there been significant growth benefits associated with the shift from taxes on international trade to taxes on domestic goods and services and, most notably, value-added taxes?

There has been a major shift from trade to consumption taxes across the developing world over the past two decades, spurred by advice from the IMF and the World Bank. Yet while this shift has been pursued largely in order to accelerate economic growth, McNabb and LeMay-Boucher report no previous efforts to estimate empirically the impact of this shift using cross-country methods—in part, it seems, because of the absence of adequate data. Ultimately, McNabb and LeMay-Boucher find no evidence that the shift from trade taxes to consumption taxes has enhanced economic growth. This is striking: Influential international advice has been premised in significant part on the growth benefits of such a shift, but the authors find no evidence that such growth benefits have materialized. Taken alongside results from Baunsgaard and Keen (2005), highlighting the difficulty many developing countries have faced in replacing foregone trade tax revenue, these results highlight the challenges that have accompanied trade tax reform in many developing countries.

None of the results reported by McNabb and LeMay-Boucher (2014) directly challenge or overturn earlier research findings. However, their access to high-quality, complete, and publicly available data has allowed them to go beyond the questions asked by earlier research, and to cast some doubt on the empirical basis for a key set of IMF policy recommendations. As importantly, their results are now easily open to scrutiny and replication, in order to advance policy research in a way not previously possible.

## 10. LESSONS FOR DEVELOPMENT RESEARCH

Recent years have seen increased public discussion of a so-called “statistical tragedy” in developing countries, with international focus turning to the relative absence and low quality of development data from many developing countries. It is increasingly understood that the weaknesses of available data can result in the neglect of issues that are not well captured by existing data, difficulty in evaluating the effectiveness of development interventions, and the potential for misleading research and policy advice. This has, in turn, carried clear practical implications. For researchers, a need to exercise much greater care in working with cross-country data. For policy makers, the need for more concerted investments in strengthening existing data, gathering new data, and strengthening the capacity of statistical departments, particularly, but not exclusively, in the developing world (e.g., [Jerven & Johnston, 2015](#)).

Attention to data questions has arisen out of the work of scholars highlighting the limitations of existing data (e.g., [Jerven, 2013a](#)), but has also been closely linked to discussion around the Sustainable Development Goals (SDGs). Because the SDGs aim to cover a much wider range of topics than their predecessors, the Millennium Development Goals, a spotlight has been placed on the need for new data against which to measure progress against key goals. In broad terms, this enhanced focus on the need for more and better data is extremely welcome, and is consistent with the motivation behind the ICTD GRD. However, it also carries a risk that the desire for *more* data may in practice outweigh the push for *better* data—and for ensuring that the data that we have are deployed with greater caution.

This distinction between *more* data and *better* data is central to the work described in this paper. One of the problems facing earlier international tax data was of quantity: cross-country data coverage within the IMF GFS was highly incomplete, and those gaps undermined the credibility of cross-country analysis. Researchers responded by taking steps to fill these gaps by merging data from multiple sources. However, as they did so the desire for *more* data sometimes trumped necessary attention to the *quality* of that data—and to how it was being used—with important consequences for research quality and reliability. These *ad hoc* datasets too often included errors, while basic limitations of the data were often addressed only briefly, if at all. These concerns were reinforced by a striking lack of transparency. These data problems have, in turn, shaped potentially misleading but influential research results.

For this reason, the focus of the ICTD GRD has been as much on quality as quantity. Like earlier efforts, it has sought to close gaps in individual international data series by merging data from multiple sources. However, much greater effort has gone into ensuring that the merging of that data has been done with sufficient care to avoid errors. This has, as discussed already, included painstaking manual cleaning of the data, careful attention to the classification of resource revenues, and efforts to minimize imperfections in underlying GDP series. The result is, as demonstrated here, a much-improved dataset, which has allowed researchers to generate more reliable findings. As important, it has created a common platform for like-minded researchers to similarly revisit existing results—including those reported here—using improved data.

While this paper has stressed the benefits of the new dataset, it remains equally important to stress that, despite these efforts, the dataset remains imperfect in various ways. Fundamentally, the ICTD GRD—or any alternative—is only as good as the underlying data from international sources. Sometimes data are missing, while we have growing evidence that these numbers are often imperfect approximations, compiled by overstretched local and international staff ([Jerven, 2013a](#)).<sup>33</sup> Our efforts to deal with natural resource revenues have similarly been limited by underlying sources. A few larger resource producers simply do not distinguish resource and non-resource revenues, and are thus excluded. Smaller resource producers often do not report resource revenues separately owing to their relatively small size—a small, but not entirely negligible, distortion in the resultant data. Meanwhile, there is not yet an explicit and universally accepted definition of “resource revenues”—though the IMF is working toward developing one—with some risk that the categorization of resource revenues varies at the margin across countries.

There is no evidence that any of these issues are seriously distorting the data in the ICTD GRD, which is the best available option for research. However, these continued limitations are a reminder of the imperfect underlying reality, and of substantial scope for investing in supporting improved data collection both nationally and internationally. They also highlight a broader point: No data are perfect, and when we speak of data quality we should be speaking not only about the data itself, but about *transparency* around the data being used, and its limitations. The most jarring issue surrounding much recent tax and development research is not the weakness of the data used, but the fact that the data were not, in many cases, made publicly available. The absence of such transparency creates major challenges for scrutinizing research findings, and dramatically increases the likelihood that errors will persist, unidentified, within researcher datasets—even when these are the basis for influential policy results.

There is a tendency, due to the proliferation of data and ease of access to powerful statistical packages, to consider the present as something of a golden age for data-driven research. An alternative perspective is, however, to also urge caution. Imperfect data, handled with insufficient care, and lacking the potential for replication and scrutiny, create a risk that the quantity of data-driven research may increase, but without adequate attention to quality and reliability. Transparency about the limitations of even the best data is thus essential in order to ensure that researchers employ adequate data, in a way consistent with its limitations—and cognizant of its imprecisions. Against this background, the best guarantee of the integrity of the ICTD GRD is that it was posted publicly online almost as soon as it was created, with a clear guide to its own limitations. This was accompanied by efforts (this paper included) to publicize it, to invite public scrutiny, and to encourage broad and consistent use of a relatively standard data source for tax and development research. Updating of the dataset to cover recent years, and address feedback from users, is already underway. The hope is that these continued efforts will contribute to the development of more robust research findings moving forward.

## NOTES

1. While the ICTD initiative began in late 2010, two of the creators of the dataset had been involved in similar efforts dating back almost a decade—while similar, but less complete, efforts by other researchers date back still further.
2. These limitations, as well as several comparatively modest issues, are described at greater length in [Prichard, Cobham, and Goodall \(2014\)](#).
3. The incompleteness of the IMF GFS is, in turn, driven by limited reporting from member countries, and the fact that the IMF GFS is only able to publish data where it is deemed to be of very high quality, and where relatively complete national accounts data can be constructed. As explained below, this explains the potential to fill gaps in the GFS by employing IMF Article IV reports, which are moderately less precise, but far more complete.
4. [Jerven \(2013b\)](#), [Cobham \(2014\)](#), [Mukanga \(2014\)](#), [Leo \(2014\)](#).
5. For CEPAL data total revenue figures can often be drawn from alternative national accounts tables, though with risks resulting from methodological inconsistencies. The African Economic Outlook and the OECD Revenue Statistics in Latin America offer no figures for total government revenue.
6. The best known is that of [Keen and Mansour \(2009\)](#), which draws on data from IMF Article IV reports in order to construct a complete dataset for sub-Saharan Africa over the period 1980–2006 which distinguishes natural resource revenues from other types of tax revenue. It has since been updated ([Mansour, 2014](#)). However, despite its many strengths, the dataset only covers one region, and only covers a small subset of the tax and revenue variables offered by the IMF GFS, thus limiting its applications.
7. For example, [Ehrhart \(2013\)](#) tells us that for non-African developing countries “the data are taken from the same sources, namely GFS and Article IV data,” (p. 10) without any further clarification. In a similar vein, [Garcia and von Haldenwang \(2015\)](#) draw on a dataset which merges data from many of the same sources described here, but for which more detailed information is not readily available.
8. Illustratively, the ICTD GRD, despite requiring three years to construct, nonetheless benefitted from new users pointing out minor errors upon its public release in September 2014. Openness to scrutiny, and to revision, seems an important benefit from publishing datasets with an engaged, technical audience.
9. The latter primarily indirectly by way of the research dataset constructed by [Keen and Mansour \(2009\)](#), which draws on data from Article IV reports.
10. This is true, among others, for Ghana, as discussed earlier.
11. A detailed account of every such “break” in the data series resulting from inappropriate merging of data from multiple sources is available on request from the author.
12. There is a suspicion in many quarters that the decision to revise the earlier IMF data, and make it public was, in part, a response to the ICTD GRD, and to the efforts described here to highlight the limitations of the earlier IMF FAD dataset. That said, this has never been stated explicitly, and remains speculation.
13. More detailed discussion of the differences between the ICTD GRD and IMF WORLD, as of the time of writing, can be found in [Prichard and McNabb \(2015\)](#).
14. While the exclusion of some Article IV data from the IMF GFS reflect concerns about data quality, we follow much earlier work—including work at the IMF itself—in concluding that such imprecisions are generally comparatively negligible. That said, all Article IV data have been interrogated carefully, and excluded where they are significantly at odds with other sources.
15. In practice, there are also a large number of countries that secure a very small share of government revenue from non-renewable natural resources—including many countries home to significant, but weakly taxed, mining sectors. However, there are not *any* international sources that systematically distinguish natural resource revenues where resource revenues are small (normally less than 1% of GDP). In these cases resource revenue is treated as equal to 0, in order to make the exercise tractable, as any resultant distortions are comparatively negligible.
16. Specifically, where GDP figures from the two sources were equivalent, plus or minus 2.5%, during overlapping years.
17. The dataset also allows users to customize their preferences, while a version of the data employing central government data exclusively is also provided to all users, in order to verify the robustness of results. The results reported below are, in general (a) robust to using only central government data, but (b) stronger when employing the preferred ICTD dataset, thus offering support for the approach adopted. The approach here is also preferable to that employed by the IMF WORLD dataset, which often merges central and general government data within the same country, resulting in sometimes large jumps in the data when transitioning between sources.
18. While low non-resource tax collection may be a rational policy choice in some resource-rich states, the majority of the countries below 15% of GDP—42 in total in the dataset—are *not* resource rich, and thus simply struggle to collect adequate revenue.
19. The papers are not precise replications of earlier findings. This reflects, among other things, the existence of multiple earlier studies in each area, a desire to consider the most econometrically appropriate approaches, and, critically, the fact that replication data and files for most of the studies were not generally available.
20. Revenue-related conditionality at the IMF did, in fact, increase in subsequent years, although this cannot be attributed to the impact of this study alone.
21. The authors of [Benedek et al. \(2014\)](#) subsequently shared their dataset and it produced results similar to those reported in the published article, but also confirmed the significant discontinuities and errors in the data noted earlier.
22. Illustrative of the problem, one of the most recent studies, [Thornton \(2014\)](#), employs incomplete and analytically problematic data from the IMF GFS, but does not pay any broader attention to data quality beyond mentioning the existence of extensive missing data.
23. They do not employ instrumental variables analysis, despite it having been employed in two recent studies ([Benedek et al., 2014](#); [Thornton, 2014](#)), owing to lack of confidence in the instruments that have been employed, and the questionable value of employing complex IV strategies. Specifically, instrumental variables analysis is intended to correct for potential problems of endogeneity. However, in this case any endogeneity is generally expected to make a negative relationship between aid and taxation *more likely*, as countries with low tax capacity may receive higher

aid as a result. Given that potential endogeneity likely makes a negative relationship *more likely*, there is little justification for employing instrumental variables unless other, more straightforward, methods also suggest a negative relationship—which the papers do not find to be the case. See discussion in Carter (2013, pp. 7–8).

24. There may, of course, also be weakness and concerns with public expenditure data, though that is not the focus here.

25. Pre-election reductions in tax collection can generally only be achieved through difficult to reverse changes in tax rates or by intentionally relaxing tax enforcement, with potentially negative long-term implications for the quality of administration.

26. Tests for the effects of contested elections (at least two competing parties), first elections, free and fair elections, and transitional elections are all consistently insignificant.

27. Recent studies, by Haber and Menaldo (2011) and Wiens *et al.* (2014), acknowledge that a focus on revenues would be theoretically preferable, but explain that data limitations have led them to focus primarily on resource income. Haber and Menaldo (2011) acknowledge this implicitly, as they construct a measure of *fiscal reliance*—the share of government revenue from natural resource wealth—but are able to do so for only 19 countries. Wiens *et al.* (2014) make their preference for measures of revenue, rather than resource income, more explicit, writing that “fiscal reliance best captures our theoretical quantity of interest” (p. 786), while nonetheless employing resource income as the best available proxy for the variable of interest.

28. By contrast, the variable for *total tax revenue* is comparatively inconsistent across estimation methods, suggesting either a weaker relationship or that the relationship is comparatively complex, and difficult to capture using cross-country methods. The latter possibility is advanced at length in Prichard (2015).

29. Note that these studies do not, for the most part, ask whether higher taxes *in general* are good or bad for growth. They are focused on which *types* of taxes are *relatively* better or worse for growth.

30. Value-added taxes are generally felt to be less economically distorting than trade taxes, though this claim has been challenged theoretically by Emran and Stiglitz (2005), who argue that this might not be the case in the imperfect markets and enforcement environments that characterize developing countries.

31. Lee and Gordon (2005) also included a wide range of developing countries in their study of tax and growth, but employ data on statutory tax rates rather than actual tax collection.

32. McNabb and LeMay-Boucher (2014) were unable to gain access to the dataset during the course of their research, though Acosta-Ormaechea later commented favorably on their paper at a workshop in September 2014.

33. Indeed, this is part of why the IMF GFS often does not publish data from IMF Article IV reports.

## REFERENCES

- Acosta-Ormaechea, S., & Yoo, J. (2012). *Tax composition and growth: A broad cross-country perspective (IMF Working Paper 12/257)*. Washington DC: IMF.
- Andersen, J., & Ross, M. (2014). Big oil change: A closer look at the Haber-Menaldo analysis. *Comparative Political Studies*, 47(7), 993–1021. <http://dx.doi.org/10.1177/0010414013488557>.
- Aslaksen, S. (2010). Oil and democracy: More than a cross-country correlation?. *Journal of Peace Research*, 47(4), 421–431. <http://dx.doi.org/10.1177/0022343310368348>.
- Baungsaard, T., & Keen, M. (2005). *Tax revenue (or?) trade liberalizations. (IMF Working Paper 05/112)*. Washington DC: IMF.
- Benedek, D., Crivelli, E., Gupta, S., & Muthoora, P. (2014). Foreign aid and revenue: Still a crowding out effect?. *FinanzArchiv: Public Finance Analysis*, 70(1), 67–96. <http://dx.doi.org/10.1628/001522114X679156>.
- Bräutigam, D., & Knack, S. (2004). Foreign aid, institutions and governance in sub-Saharan Africa. *Economic Development and Cultural Change*, 52, 255–285. <http://dx.doi.org/10.1086/380592>.
- Brender, A., & Drazen, A. (2005). “Political budget cycles in new versus established democracies. *Journal of Monetary Economics*, 52, 1271–1295. <http://dx.doi.org/10.1016/j.jmoneco.2005.04.004>.
- Brun, J.-F., Chambas, G., & Guerinéau, S. (2008). *Aide et mobilisation fiscale dans les pays en développement. (CERDI, Études et Documents E 2008.12)*. Clermont-Ferrand: CERDI.
- Carter, P. (2013). Does foreign aid displace domestic taxation?. *Journal of Globalization and Development*, 4(1), 1–47. <http://dx.doi.org/10.1515/jgd-2013-0023>.
- Clist, P. (2014). *Foreign aid and domestic taxation: Multiple sources, one conclusion (ICTD Working Paper)*. Brighton: Institute of Development Studies.
- Clist, P., & Morrissey, O. (2011). Aid and tax revenue: Signs of a positive effect since the 1980s. *Journal of International Development*, 23(2), 165–180. <http://dx.doi.org/10.1002/jid.1656>.
- Cobham, A. (2014, April 9). *Nigeria's upward revision of GDP should sound alarm on tax-to-GDP ratio (Center for Global Development Blog 9)*. <<http://www.cgdev.org/blog/nigerias-upward-revision-gdp-should-sound-alarm-tax-gdp-ratio>>.
- Crivelli, E., & Gupta, S. (2014). Resource blessing, resource curse? Domestic revenue effort in resource-rich countries. *European Journal of Political Economy*, 35, 88–101. <http://dx.doi.org/10.1016/j.ejpoleco.2014.04.001>.
- de Haan, J., & Klomp, J. (2013). Conditional political budget cycles: A review of recent evidence. *Public Choice*, 157, 387–410. <http://dx.doi.org/10.1007/s11127-013-0106-6>.
- Drazen, A. (2001). The political business cycles after 25 years. In B. Bernanke, & K. Rogoff (Eds.), *NBER macroeconomics annual 2000* (Vol. 15, pp. 75–138). Cambridge: MIT Press, <http://www.nber.org/chapters/c11055>.
- Dunning, T. (2008). *Crude democracy: Natural resource wealth and political regimes*. Cambridge: Cambridge University Press.
- Emran, S., & Stiglitz, J. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89, 599–623. <http://dx.doi.org/10.1016/j.jpubeco.2004.04.007>.
- Ehrhart, H. (2013). Elections and the structure of taxation in developing countries. *Public Choice*, 156(2), 195–211. <http://dx.doi.org/10.1007/s11127-011-9894-8>.
- García, M. M., & von Haldenwang, C. (2015). Do democracies tax more? Political regime type and taxation. *Journal of International Development*. <http://dx.doi.org/10.1002/jid.3078>.
- Gupta, S., Clemens, B., Privoarsky, A., & Tiongson, E. (2004). Foreign aid and revenue response: Does the composition of aid matter?. In S. Gupta, B. Clemens, & G. Inchauste (Eds.), *Helping countries develop: The role of fiscal policy* (pp. 385–406). Washington DC: IMF.
- Haber, S., & Menaldo, V. (2011). Do natural resources fuel authoritarianism? A reappraisal of the resource curse. *American Political Science Review*, 105(1), 1–26. <http://dx.doi.org/10.1017/S0003055410000584>.
- Herb, M. (2003). Taxation and representation. *Studies in Comparative International Development*, 38(3), 3–31. <http://dx.doi.org/10.1007/BF02686197>.
- Herrera, Y., & Kapur, D. (2007). Improving data quality: Actors, incentives and capabilities. *Political Analysis*, 15, 365–386. <http://dx.doi.org/10.1093/pan/mpm007>.
- IMF (2011). *Revenue mobilization in developing countries*. Washington DC: IMF.



- Jensen, N., & Wantchekon, L. (2004). Resource wealth and political regimes in Africa. *Comparative Political Studies*, 37(7), 816–841. <http://dx.doi.org/10.1177/0010414004266867>.
- Jerven, M. (2013a). *Poor numbers: How we are misled by African development statistics and what to do about it*. Ithaca: Cornell University Press.
- Jerven, M. (2013b). Comparability of GDP estimates in sub-Saharan Africa: The effect of revisions in sources and methods since structural adjustment. *Review of Income and Wealth*, 59, S16–S36. <http://dx.doi.org/10.1111/roiw.12006>.
- Jerven, M., & Johnston, D. (2015). Statistical tragedy in Africa? Evaluating the data base for African economic development. *Journal of Development Studies*, 51(2), 111–115. <http://dx.doi.org/10.1080/00220388.2014.968141>.
- Keen, M., & Mansour, M. (2009). *Revenue mobilization in sub-Saharan Africa: Challenges from globalization (IMF Working Paper 09/157)*. Washington DC: IMF.
- Lee, Y., & Gordon, R. (2005). Tax structure and economic growth. *Journal of Public Economics*, 89(5–6), 1027–1043. <http://dx.doi.org/10.1016/j.jpubeco.2004.07.002>.
- Leo, B. (2014, April 8). *Beyond the headlines: 3 under-reported stories about Nigeria's 90% GDP boost (Center for Global Development Blog 8)*. <<http://www.cgdev.org/blog/beyond-headlines-3-under-reported-stories-about-nigerias-90-gdp-boost>>.
- Mansour, M. (2014). *A tax revenue dataset for sub-Saharan Africa: 1980–2010 (FERDI Working Paper 19)*. Clermont-Ferrand: FERDI.
- McNabb, K., & LeMay-Boucher, P. (2014). *Tax structures, economic growth and development (ICTD Working Paper 22)*. Brighton: Institute of Development Studies.
- Moore, M. (1998). Death without taxes: Democracy, state capacity and aid dependence in the Fourth World. In M. Robinson, & G. White (Eds.), *The democratic development state: Politics and institutional reform* (pp. 67–88). Oxford: Oxford University Press.
- Morrissey, O., Prichard, W., & Torrance, S. (2014). *Aid and taxation: Exploring the relationship using new data (ICTD Working Paper 21)*. Brighton: Institute of Development Studies.
- Morrissey, O., & Torrance, S. (2015). Aid and taxation. In B. M. Arvin, & B. Lew (Eds.), *Handbook on the economics of foreign aid* (pp. 555–576). Cheltenham: Edward Elgar.
- Mukanga, C. (2014). Zambia releases GDP (25% larger). *Zambian Economist*. Retrieved from <<http://www.zambian-economist.com/2014/02/zambia-rebases-gdp-25-larger.html>>.
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967–1012. <http://dx.doi.org/10.1111/j.1468-0262.2006.00692.x>.
- Prichard, W. (in press). Electoral competitiveness, tax bargaining and political incentives in developing countries: Evidence from political budget cycles affecting taxation. *British Journal of Political Science*.
- Prichard, W. (2015). *Taxation, responsiveness and accountability in developing countries: The dynamics of tax bargaining*. Cambridge: Cambridge University Press.
- Prichard, W. (2014). *Electoral competitiveness, political budget cycles and taxation in developing countries (ICTD Working Paper 24)*. Brighton: Institute of Development Studies.
- Prichard, W., Cobham, A., & Goodall, A. (2014). *The ICTD Government Revenue Dataset (ICTD Working Paper 19)*. Brighton: Institute of Development Studies.
- Prichard, W. & McNabb, K. (2015). The ICTD Government Revenue Dataset: Still the best option for researchers. International Center for Tax and Development. Retrieved from <[www.ictd.ac/images/ICTD%20GRD%20Research%20Note.pdf](http://www.ictd.ac/images/ICTD%20GRD%20Research%20Note.pdf)>.
- Prichard, W., Salardi, P., & Segal, P. (2014). *Taxation, non-tax revenue and democracy: New evidence using new cross-country data (ICTD Working Paper 23)*. Brighton: Institute of Development Studies.
- Remmer, K. (2004). Does foreign aid promote the expansion of government?. *American Journal of Political Science*, 48(1), 77–92. <http://dx.doi.org/10.1111/j.0092-5853.2004.00057.x>.
- Ross, M. (2001). Does oil hinder democracy?. *World Politics*, 53, 325–361. <http://dx.doi.org/10.1353/wp.2001.0011>.
- Ross, M. (2004). Does taxation lead to representation?. *British Journal of Political Science*, 34, 229–249. <http://dx.doi.org/10.1017/S0007123404000031>.
- Thornton, J. (2014). Does foreign aid reduce tax revenue? Further evidence. *Applied Economics*, 46(4), 359–373. <http://dx.doi.org/10.1080/00036846.2013.829207>.
- Wiens, D., Poast, P., & Clark, W. (2014). The political resource curse: An empirical re-evaluation. *Political Research Quarterly*, 67(4), 783–794. <http://dx.doi.org/10.1177/1065912914543836>.

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

**ScienceDirect**