

UNU-WIDER GOVERNMENT REVENUE DATASET (GRD):USER GUIDE & FAQs

This short guide introduces the UNU-WIDER Government Revenue Dataset (GRD) and provides guidance on some frequently asked questions. The following files are available:

<i>File Details</i>		
"Core" GRD files	Excel	UNUWIDERGRD_[year].xlsx
		UNUWIDERGRD_[year]_Central.xlsx
		UNUWIDERGRD_[year]_General.xlsx
	Stata	UNU-WIDER GRD [year].dta
		UNU-WIDER GRD Central Government [year].dta
		UNU-WIDER GRD General Government [year].dta
"Full" GRD files	Excel	UNUWIDERGRD_2021_Full.xlsx
		UNUWIDERGRD_[year]_Central_Full.xlsx
		UNUWIDERGRD_[year]_General_Full.xlsx

Those **listed in red** constitute the ‘core’ or ‘merged’ GRD files which most users will consult. The majority of users will access the GRD through these files. These files contain the best available data between General and Central.

The remaining file types are suffixed by ‘Central’ or ‘General’ and contain only Central and General Government data respectively. The Central and General Government Excel files suffixed by ‘Full’ contain all the original formulae used to convert data from LCU to % of GDP, whilst the file *UNUWIDERGRD_2021_Full.xlsx* allows users to alter the choice between Central and General files.

In addition to the present note, the following documentation is available for the GRD:

<i>File</i>	<i>Description</i>	<i>Available from:</i>
Technical Note 1	Describes the source selection criteria employed in the GRD	McNabb, K., A. Oppel, and D. Chachu (2021a). 'Government Revenue Dataset (2021): Source Selection'. WIDER Technical Note 2021/10. Helsinki: UNU-WIDER.
Technical Note 2	Describes each of the variables in the GRD in detail.	Oppel, A., K. McNabb, and D. Chachu (2021). 'Government Revenue Dataset (2021): Variable Description'. WIDER Technical Note 2021/11. Helsinki: UNU-WIDER
Working Paper (2014)	ICTD working paper describing the construction of the initial version of the GRD	Prichard, W., Cobham, A., & Goodall, A., (2014) 'The ICTD Government Revenue Dataset', ICTD Working Paper 19
Working Paper (2017)	UNU-WIDER Working Paper describing 2017 updates.	McNabb, K., (2017) 'Toward Closer Cohesion of International Tax Statistics', WIDER Working Paper 184/2017

UNU-WIDER welcomes all feedback from users of the GRD. Queries should be directed to: GRD@wider.unu.edu

Frequently Asked Questions

What are the various options for accessing the GRD?

The GRD can be accessed in a number of ways. In 2019, UNU-WIDER launched the [GRD Explorer](#). This presents the GRD in a user-friendly, intuitive online tool that allows users to visualize or download custom selections of the data. The *Explorer* tool is the quickest and easiest way to access the data and allows users to download the data in both % of GDP and Local Currency (LCU). For users wishing to access to the entire dataset, this can be downloaded in Excel or Stata format from within the *Explorer* tool, or from the pages listed above.

Should analysis employ data from the merged, general government or central government dataset?

There is no universal answer to this question, as it will depend on the nature of the research question. That said, for studying total revenue collection most users are likely to use the main, or merged, dataset, which employs general government data where available, while inserting central government data if, and only if, there is evidence that subnational revenue collection is limited. By contrast, relying exclusively on central government data will dramatically underestimate total revenue collection in fiscally decentralized states, while general government data is only available for a limited group of countries. While relying on the merged dataset implies a modest underestimation of revenue collection in countries for which general government data is not available, this will generally be consistent over time, and preferable to the major risks introduced by relying exclusively on central or general government data.

Should analysis include or exclusive social contributions from total revenue and total tax revenue?

Conceptually, the answer is ambiguous and highly dependent on the research question: social contributions are not the same as other taxes, as they are contributions toward a specific area of public spending, but they nonetheless represent an important area of government involvement. Pragmatically, however, most users are advised to rely on revenue and tax figures exclusive of social contributions, owing to problems of completeness and comparability for social contributions figures. Specifically, many countries do not consistently report social contributions, leading to more extensive missing data, whilst what is reported in the ‘social contributions’ category does not appear to be entirely consistent across countries, particularly for developing countries and when relying on IMF Article IV reports, thus raising concerns that including social contributions in the analysis could produce misleading results.

If I am interested in resource revenues, should I use the “total resource revenue” variable, or an alternative strategy to measure resource dependence?

We would ideally be able to produce a figure for “total resource revenue” for every country, which would capture all resource revenues, whether those revenues are recorded as tax or non-tax revenue in official government accounts. However, in practice, this is simply not possible: Relatively few countries report a figure for “total resource revenue”, with some additional countries identifying the resource components of total tax revenue or of total non-tax revenue, but not both. Current international initiatives to strengthen the collection of resource revenue data are correspondingly extremely valuable.

In the absence of such data, users may nonetheless wish to include a proxy for resource wealth in econometric analyses. In these cases the best option available to most users is likely to be to calculate “total non tax revenue” as the difference between “total revenue” and “total non-resource tax revenue”. This variable will thus include all types of non-tax revenue, including resource revenue, but acts as a useful proxy for resource revenue because resource revenue explains the vast majority of the variation in “total non tax revenue” across countries. See discussion in Prichard et al. (2014).

Why do the subcomponents of tax revenue not sum to the respective totals?

Total tax revenue is generally equal to the sum of the sub-components of tax revenue. However, the total tax figure exceeds the sum of the sub-components in cases where some revenue was not allocated to any specific category in the underlying source, which may occur for a variety of reasons. Thus, for example, Total Taxes on Income, Profits and Capital Gains is sometimes larger than the sum of CIT + PTT, as some sources allocate some revenue to “other income taxes”, which cannot be allocated to either personal or corporate income taxes.

Very occasionally, the subcomponents of revenues are slightly greater than or smaller than the reported totals in the underlying source. The GRD, as a rule, does not modify data from underlying sources, thus retaining these minor discrepancies in order to ensure transparency. If users wish to correct such data for their own analyses, they are free to do so, though changes of this kind to the data should be carefully recorded for any analysis.

For most EU countries, the data in the GRD comes from the OECD Revenue Statistics. The total tax figure in such cases is inclusive of ‘Customs duties collected for the EU’, reflecting the ‘Total Tax Revenue’ figure in the OECD. However, this item is not allocated to any subcomponent, meaning that the sum of subcomponents will often be marginally lower than the total tax figure.

Why does revenue data in the GRD not match that in other sources?

The GRD expresses all data in terms of a common underlying GDP figure, taken from the World Economic Outlook (WEO).¹ This helps to ensure comparability of data across different sources, but might be higher, or lower, than tax ratios expressed in other sources, based on different GDP data.

The underlying GDP figures taken from the most recent version of the WEO will reflect the rebasing of GDP or recent adoption of more modern systems of national accounting. In some cases, this leads large changes in GDP, which might cause users to find lower tax ratios than an older source that does not express tax figures in terms of up-to-date GDP figures.

Tax statistics in other sources may be inflated by revenues earned from natural resources. The GRD, where possible, presents both total revenue and total tax figures both inclusive and exclusive of resource revenues.

Is it possible to estimate total subnational revenue by subtracting Central Government revenue from General Government revenue?

Generally speaking, the difference between General and Central Government revenue should be equal to total subnational revenue. However, any such calculations should be undertaken with extreme caution, and checked carefully, as definitional differences, or other unusual accounting practices, between central and general government sources could lead to over or under-estimating subnational revenue.

Data that I am interested in using has been flagged as potentially problematic.

The authors of the GRD have striven to alert users to cases where data might be potentially unreliable or misleading, through a series of ‘flags’. These are defined in each data file, and often notes relating to the need for the individual flags are included. Users are advised to carefully consider the inclusion of any data that is flagged as *Caution 1 – Caution 4* from their econometric or cross - country analysis, depending on the variables required. In cases where results do not hold following the exclusion of such data, users are recommended to communicate this to their audience and/or exclude the data all together.

More generally, because of well-known imperfections with the collection of government revenue data – and with GDP data used in some countries to calculate tax revenues – best practice for econometric analysis should always be to pursue extensive robustness testing, including sensitivity to the inclusion of data from individual, or small groups, of countries. Caution should also always be employed when comparing tax ratios across countries, as high, or rapidly rising, tax ratios in individual countries may reflect the underestimation and irregular rebasing of GDP.

What are some of the limitations to the GRD?

There are widely recognised concerns over the quality of government revenue data in existing sources. The GRD strives to present the ‘best’ available data consistently and to flag any noticeably problematic or questionable data. This already represents an improvement over any other source. However, the broader accuracy of the numbers contained in the GRD is only as good as the underlying data obtained from each source.

¹ In cases where no data is available in the WEO, GDP figures are taken from the underlying source of the tax data, e.g. Article IV Staff Reports.

In particular, data on natural resource revenues have not historically been collected consistently, or according to strictly enforced definitions. As such, the disaggregation between non-resource taxes and resource taxes is likely to be imperfect at the margin – though still better than the alternative of not making any distinction. There is potential for more precise data in the future, as more systematic data collection efforts are put in place – something toward which both the IMF and OECD are working. In a similar vein, no data on resource revenues is generally available for countries with small levels of resource revenue (generally less than 1% of GDP) and, as such, all revenue is treated as non-resource tax revenue in these cases.

There are also potential inconsistencies across countries, or over time, in the disaggregation of tax revenue between taxes on international trade and taxes on goods and services. This reflects potential inconsistencies in the classification of taxes on goods and services collected at the border, by customs officials. While every effort has been made to ensure consistency, users should be aware of this potential issue in individual countries when working with these disaggregated categories.

There may be cases where the data in the GRD appears slightly less complete (in terms of disaggregation) than that in, for example, the GFS. This would be as a result of using IMF Article IV data that, whilst not as detailed, allows the resource component of tax to be isolated (see Technical Note 1 for more information).

References:

McNabb, K. (2017). 'Toward closer cohesion of international tax statistics: The ICTD/UNU-WIDER GRD 2017', WIDER Working Paper 184/2017. UNU-WIDER: Helsinki. <https://doi.org/10.35188/UNU-WIDER/2017/410-0>.

[Prichard, W., Cobham, A., & Goodall, A., \(2014\) 'The ICTD Government Revenue Dataset', ICTD Working Paper 19](#)

McNabb, K., A. Opper, and D. Chachu (2021a). 'Government Revenue Dataset (2021): Source Selection'. WIDER Technical Note 2021/10. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/WTN/2021-10>

Opper, A., K. McNabb, and D. Chachu (2021). 'Government Revenue Dataset (2021): Variable Description'. WIDER Technical Note 2021/11. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/WTN/2021-11>