

ONLINE APPENDIX

ADB Economics Working Paper No. 665

DEVELOPING ASIA'S FISCAL LANDSCAPE AND CHALLENGES

Eugenia Go, Sam Hill, Maria Hanna Jaber, Yothin Jinjarak, Donghyun Park, and Anton Ragos

June 2022

APPENDIX

A. Data Sources

Our comparative analysis presented in the main section required a range of tax revenue data that provided good temporal and coverage of economies, particularly of developing Asian economies, as well as details on individual tax revenue components, ideally presented on a general government basis. Against this objective, existing revenue data sources across economies all suffered from various shortcomings, requiring us to combine data from different sources to maximize data coverage.

The fiscal revenue data used in our analyses mainly comes from two widely used sources: the OECD Revenue Statistics Database and the IMF-GFS Revenue Database. Each of these databases have desirable features as well as limitations. The OECD database has superior temporal coverage of 113 economies from 1995 to 2019, presents revenue on a general government basis, and has a full and consistent decomposition of revenues by tax type.

In contrast, the IMF-GFS covers significantly more economies: 181. The IMF-GFS has data for 41 developing Asian economies of ADB, compared with only 20 for OECD data. The GFS database is updated on a weekly basis and, at the time of writing in early 2022, included some data for 2020, valuable for an initial assessment of the effects of COVID-19 on fiscal revenues. However, its temporal coverage is patchier, and the presentation of tax subcomponents are less harmonized across economies. For example, Singapore's value-added tax is missing in the IMF-GFS and subsumed under taxes on goods and services. By comparison, the two components are reported separately for Thailand. Finally, data in the IMF-GFS is a mix of general government (36%) and central government (64%) reporting across economies.

By combining data sources, we create a dataset that maximizes the desirable features of the IMF-GFS and OECD revenue statistics. Table A1 provides a summary of economy, temporal, and general government coverage comparisons across the three datasets. Our core dataset includes data for 192 economies from 1995 to 2020, including the main tax aggregates of personal and corporate income, value-added tax and other goods and services, and international trade. The gains are most evident in terms of coverage of economies. Data for 66% of economies is presented on a general government basis. The lower temporal and general government coverage are simply artefacts of the dataset combination rather than poorer coverage.

Table A1: Comparing Economy, Temporal, and General Government Reporting Coverage in Different Datasets

Variables	Total (%)			ADB Member Economies (%)			Latin America (%)			Sub Saharan Africa (%)			High Income OECD (%)		
	GFS	OECD	ADB Core	GFS Core	OECD	ADB Core	GFS	OECD	ADB Core	GFS Core	OECD	ADB Core	GFS Core	OECD	ADB Core
Economy	181	113	192	41	20	45	29	28	31	43	27	45	33	33	33
Taxes	80	83	81	75	71	76	89	95	97	63	75	65	99.5	100	100
Taxes: income, profits, and capital gains	75	83	78	80	71	68	83	96	95	63	75	65	99.5	100	100
Payable by individuals	64	80	72	56	66	61	66	87	90	54	73	62	93	100	100
Payable by corporations and other enterprises	67	80	74	56	66	60	70	90	90	54	72	61	93	100	100
Taxes on payroll and workforce	25	82	57	16	68	39	24	96	84	19	75	49	53	100	100
Taxes on property	58	80	68	47	70	54	77	96	97	32	65	48	98	100	100
Taxes on goods and services	74	83	77	64	83	66	85	96	97	62	75	64	99	100	100
VAT	51	83	64	32	82	46	57	94	89	43	74	53	90	100	100
Excise tax	64	82	70	51	82	56	70	92	86	54	71	60	94	100	100
Taxes on international trade and transaction	67	82	77	65	70	66	88	92	94	62	71	63	60	100	100
Share of general government observations in OECD core	36	100	65	44	100	57	83	100	81	5	100	54	100	100	100

ADB = Asian Development Bank, GFS = Government Finance Statistics, OECD = Organisation for Economic Co-operation and Development.

Note: (i) Green-shaded columns represent superior country coverage.

(ii) Green-highlighted numbers represent superior temporal coverage measured as the share of non-missing entries to total potential entries from 1995 to 2019.

Sources: Authors' calculations; Organisation for Economic Co-operation and Development. Global Revenue Statistics Database. <https://www.oecd.org/tax/tax-policy/>; International Monetary Fund. Government Finance Statistics Database. <https://data.imf.org/> (both accessed 31 January 2022).

It must be noted that the practice of combining databases is common for fiscal data analyses across economies. For example, the IMF World Revenue Longitudinal Data (WoRLD) was compiled using data from the OECD, the IMF-GFS, World Economic Outlook, and IMF staff estimates. The UN-WIDER Government Revenue Dataset similarly draws from multiple sources.

For our purposes, the process of supplementing databases with each other was subject to careful comparisons across the databases, ensuring that (i) data remains comparable within an economy over time; (ii) data across economies are as comparable as possible; and (iii) data remains internally consistent, i.e., the subcomponents, in principle, add up to the aggregate components. The process is described in the next section.

B. Data Compilation Notes

The OECD revenue database served as the base dataset. Data from 75 economies not covered in the OECD were imported from the IMF-GFS. Data for four economies (Brunei Darussalam; Taipei,China; Turkmenistan; and Tuvalu), which are neither in OECD nor in IMF-GFS were sourced from the ADB Key Indicators Database (KIDB). In total, our core database covers 192 economies. Table A2 presents the matching of variables across the three databases.

Table A2: Variable Mapping across Databases

IMF-GFS	OECD	ADB Key Indicators
Total revenue		Current revenue/Total revenue
Tax revenue	Total taxes (less social security contributions)	Taxes
Income tax	Taxes on income, profits, and capital gains	Taxes: income, profits, and capital gains
Individual taxes	Of individuals	Payable by individuals
Corporate taxes	Of corporations	Payable by corporations
Payroll taxes	Taxes on payroll and workforce	Taxes on payroll and workforce
Property taxes	Taxes on property	Taxes on property
Taxes and goods and services	Taxes on goods and services - (Customs and import duties + Taxes on exports + Other taxes on international trade and transactions)	Taxes on goods and services
VAT	Value added taxes	
Excise tax	Excises	
International trade taxes	Customs and import duties + Taxes on exports + Other taxes on international trade and transactions	Taxes on international trade and transaction
Revenue from Social contributions		Social contributions
Grants		Grants
Other revenue		Other revenue

ADB = Asian Development Bank, GFS = Government Finance Statistics, IMF = International Monetary Fund, OECD = Organisation for Economic Co-operation and Development, VAT = value-added tax.

Source: Authors.

The ultimate data source for each economy was decided using the following criteria:

- (i) The database with the most complete and unbroken series from 1995 to 2019.

For example, the data for the PRC in the OECD database only begins from 2009 to 2019, whereas the data in the IMF-GFS is from 1995 to 2019. In this case, we opted to keep the data from the latter source. The final source for each economy is documented in the source column of the database.

(ii) Preference for general over central government reporting:

Whereas the OECD presents data for general government, the IMF-GFS presents both central and general government data for some economies. There is a preference for general government, which ideally represents the totality of an economy's revenues across government levels as described in equation (1).

$$\text{general government} = \text{central} + \text{state} + \text{local} + \text{social security} \quad (1)$$

When possible, general government series were derived by adding up subcomponents of equation (1). For example, the general government reporting of the PRC did not become available until 2005, whereas data on central budgetary government and local government are available from 1995. General government data was derived for years before 2005 by adding together entries for central budgetary and local governments, while state government in equation (1) and extrabudgetary and social security entries in equation (2) are assumed to be zero. The resulting sums for years 2005 onward were confirmed to be equal to the data reported as general government. Similarly, the general government data for India was derived by summing up budgetary central and state government entries.

$$\text{central government} = \text{budgetary} + \text{extrabudgetary} + \text{social security} \quad (2)$$

Notwithstanding the derivations, less than a third of the economies in the IMF-GFS dataset have good temporal and subcomponent coverage on general government entries. The choice between general or central government was determined based on the series that provided the maximum temporal coverage. Our core dataset is, therefore, a mix of general government (65%) and central government reporting (35%).

In general, we refrain from mixing central and government entries within an economy. However, there are a few cases where an economy shifted from central to general government reporting without any overlapping years when both are reported: Albania (2004), Armenia (2004), the Kyrgyz Republic (2014), and Turkmenistan (2011). We joined the two series after verifying that doing so (a) did not introduce large deviations from overall time trends, and (b) resulting revenue trends are comparable to the total revenue reported in the World Economic Outlook. These cases are indicated by the variable *break*, which is equal to one for the joining year for each of these economies. Changes in fiscal and calendar year reporting are likewise marked by the *break* variable. Annotations on the nature of the break can be found in *breaknotes*.

There are cases where general government is the same as central government (i.e., Hong Kong, China, and Singapore), and this is most common for smaller economies such as Fiji, the Lao People's Democratic Republic, Maldives, Papua New Guinea, and Vanuatu. In these cases, we consider the series as general government.

(iii) Within economy mixing of sources and internal consistency:

Both the OECD and GFS databases exhibit data gaps for some economies. An example is Kenya where OECD only reports data from 2001–2018 for general government, whereas GFS has data from 1995 to 2019 with central government reporting. In such cases,

we supplement the OECD data with the GFS data if it can be verified from overlapping observations that the two sources are close in magnitude. Specifically, we supplement the OECD data (missing observations) with the GFS data (non-missing observations) provided that $\left| \frac{OECD_t - GFS_t}{OECD_t} \right| \leq .05$ for $t \pm 2$, where t represents the year where data is missing from OECD dataset. This discrepancy check is carried out variable-by-variable to preserve the internal consistency of our database.

For developing Asia, ADB KIDB provides a valuable resource for extending temporal coverage. For example, in the GFS, data for Armenia (not covered in OECD), begins in 2003, whereas in ADB KIDB, it starts in 2000. We augmented Armenia's data using ADB KIDB provided the 5% discrepancy rule of thumb noted above is satisfied. The Asian Development Outlook Update database also proved a useful source albeit only for the aggregate tax revenue variable.

One area where mixing of data sources is common surrounds the years 2019 and 2020 because data coverage for OECD ends in 2018 and 2019 for most economies, whereas ADB KIDB has 2020 data for 19 developing Asian economies, and the IMF-GFS has 2020 data entries for 64 economies, nine of which are from developing Asia (with coverage expected to grow gradually over its weekly updates. The data in this background reflect GFS data as of 22 January 2022). Data augmentations for 2019 and 2020 generally follow the 5% discrepancy guideline.