Package 'ato'

May 7, 2026

Title Download and Tidy Australian Taxation Office Data

Version 0.1.0

Description Fetch Australian Taxation Office ('ATO') Taxation Statistics and related datasets via the 'data.gov.au' Comprehensive Knowledge Archive Network ('CKAN') API <https://data.gov.au/data/api/3/>. Provides tidy access to individual, company, superannuation, goods and services tax ('GST'), fringe benefits tax ('FBT'), Voluntary Tax Transparency Code ('VTTC'), Pay As You Go ('PAYG') withholding, charity, excise, and Corporate Tax Transparency data, plus Division 293, Petroleum Resource Rent Tax, Medicare Levy Surcharge, fuel tax credits, compliance, and Working Holiday Maker aggregates. Includes reproducibility helpers (snapshot pinning, SHA-256 cache integrity, session manifest, optional 'Zenodo' deposit), classification crosswalks ('ANZSIC' 2006 to 2020, 'ANZSCO' 2013 to 2021), panel harmonisation, reconciliation against Final Budget Outcome totals, and real-terms and per-capita helpers backed by bundled Australian Bureau of Statistics ('ABS') Consumer Price Index and Estimated Resident Population series. Bridges to the 'taxstats' 2 per cent microdata sample via column-schema mapping. Data is published by the Australian Taxation Office under Creative Commons Attribution 2.5 Australia or 3.0 Australia licences (dataset-dependent).

Depends R (>= 4.1.0)

License MIT + file LICENSE

Encoding UTF-8

Language en-US

RoxygenNote 7.3.3

Imports cli (>= 3.6.0), httr2 (>= 1.0.0), jsonlite, readxl (>= 1.4.0), tools, utils

Suggests digest, knitr, openssl, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

URL https://charlescoverdale.github.io/ato/,
https://github.com/charlescoverdale/ato

BugReports https://github.com/charlescoverdale/ato/issues

NeedsCompilation no

Author Charles Coverdale [aut, cre]

Maintainer Charles Coverdale <charlesfcoverdale@gmail.com>

Repository CRAN

Date/Publication 2026-04-28 18:50:19 UTC

Contents

`ato_cache_info` *Inspect the local ato cache*

Description

Inspect the local ato cache

Usage

```
ato_cache_info()
```

Value

A list with `dir`, `n_files`, `size_bytes`, `size_human`, and `files`.

See Also

Other configuration: `ato_clear_cache()`, `ato_meta()`

Examples

```
op <- options(ato.cache_dir = tempdir())
ato_cache_info()
options(op)
```

`ato_catalog` *ATO dataset catalogue*

Description

Returns a summary of all datasets published by the Australian Taxation Office on data.gov.au. Each row is a CKAN "package" with an id (slug), title, licence, modification date, and resource count.

Usage

```
ato_catalog(q = NULL)
```

Arguments

`q`	Optional free-text filter (CKAN Solr query). `NULL` returns the full ATO catalogue.

Value

An `ato_tbl` with one row per dataset.

Source

'data.gov.au' CKAN endpoint https://data.gov.au/data/organization/australiantaxationoffice.

See Also

Other discovery: `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  cat <- ato_catalog()
  head(cat[, c("id", "title", "licence")])
})
options(op)
```

`ato_charities` *Charity and deductible gift recipient data*

Description

Returns the ATO's data on income tax-exempt entities and Deductible Gift Recipients (DGRs): entity counts, income, expenditure, and gift deductions by charity subtype and state. Covers public benevolent institutions, health promotion charities, environmental organisations, and other DGR categories.

Usage

```
ato_charities(year = "latest")
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2021-22"`) or `"latest"`.

Details

Used by Treasury (charity tax expenditure estimates), researchers studying the non-profit sector, and civil society policy analysts.

Value

An `ato_tbl`. Monetary values in nominal AUD.

Source

Australian Taxation Office charity statistics on data.gov.au. Licensed CC BY 2.5 AU.

See Also

Other discovery: `ato_catalog()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  ch <- ato_charities(year = "2021-22")
  head(ch)
})
options(op)
```

`ato_cite` *Cite an ato_tbl (or URL) in BibTeX and plain-text form*

Description

Returns a citation suitable for footnotes, papers, and Treasury-grade briefs. Uses the provenance attributes attached to every `ato_tbl`: source URL, licence, retrieval date, title, snapshot pin, and SHA-256 digest.

Usage

```
ato_cite(x, style = c("text", "bibtex", "apa"), doi = NULL)
```

Arguments

`x`	Either an `ato_tbl` (as returned by any `ato_*` data function) or a character URL pointing to an ATO data.gov.au resource.
`style`	One of `"text"` (default, plain-text footnote), `"bibtex"`, or `"apa"`.
`doi`	Optional DOI (e.g. from `ato_deposit_zenodo()`) to include in BibTeX output as a `doi` field and APA suffix.

Details

BibTeX output includes the SHA-256 digest (first 12 hex chars) and snapshot pin (when set via `ato_snapshot()`) in the `note` field, which is what research reviewers need to verify the provenance of a downstream result.

Value

A character string. For `style = "bibtex"`, a complete `@misc{}` entry.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
x <- data.frame(a = 1)
x <- structure(x,
  ato_source = "https://data.gov.au/data/dataset/example.xlsx",
  ato_licence = "CC BY 2.5 AU",
  ato_retrieved = as.POSIXct("2026-04-23 00:00:00", tz = "UTC"),
  ato_title = "ATO individuals 2022-23",
  ato_sha256 = "abc123def456",
  ato_snapshot_date = "2026-04-23",
  class = c("ato_tbl", "data.frame"))
```

```
ato_cite(x)
ato_cite(x, style = "bibtex")
# DOI style: supply any minted DOI (Zenodo, DataCite, etc.).
# The placeholder below is illustrative only.
ato_cite(x, style = "apa", doi = "10.5281/zenodo.XXXXXXXX")
```

`ato_clear_cache` *Clear the ato cache*

Description

Deletes all locally cached files. The next call to any data function will re-download.

Usage

```
ato_clear_cache()
```

Value

Invisibly returns `NULL`.

See Also

Other configuration: `ato_cache_info()`, `ato_meta()`

Examples

```
op <- options(ato.cache_dir = tempdir())
ato_clear_cache()
options(op)
```

`ato_companies` *Company Taxation Statistics*

Description

Returns the annual Company Taxation Statistics tables. The Company release ships tables covering entity type, turnover band, industry, taxable status, source of income, and expense deductions. Pick the table that matches your question:

Usage

```
ato_companies(
  year = "latest",
  table = c("industry", "snapshot", "key_items_by_size", "entity_type",
    "industry_by_size", "sub_industry", "taxable_status", "source", "expenses"),
  industry = NULL
)
```

Arguments

`year`	`"YYYY-YY"`, `"latest"`, or a vector of years for a multi-year panel. Multi-year requests add a `year` column.
`table`	One of `"snapshot"`, `"key_items_by_size"`, `"entity_type"`, `"industry"` (default), `"industry_by_size"`, `"sub_industry"`, `"taxable_status"`, `"source"`, or `"expenses"`.
`industry`	Optional substring filter on industry name (applied only when the fetched table has an industry column).

Details

- **snapshot** (T1): aggregate counts, total income, net tax across all companies (~1m entities)
- **key_items_by_size** (T2): net tax by company size band
- **entity_type** (T3): split by public/private/co-operative
- **industry** (T4, default): key items by 2-digit ANZSIC subdivision
- **industry_by_size** (T5): industry x turnover band
- **sub_industry** (T6): 4-digit ANZSIC class detail
- **taxable_status** (T7): items by taxable status
- **source** (T8): source of income
- **expenses** (T9): expense and deduction categories

Classification break. Releases from 2022-23 onwards use ANZSIC 2020; earlier releases use ANZSIC 2006. A warning is emitted when the requested year(s) are at or after this boundary, or when a multi-year request spans it.

Value

An `ato_tbl`. Monetary values in nominal AUD of the reporting year.

Source

Australian Taxation Office Taxation Statistics Company Tables. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: Company tables explanatory notes*. Methodology notes on lodgement cut-off, entity-type definitions, and turnover-band thresholds. Accessible from https://www.ato.gov.au/about-ato/research-and-statistics/in-detail/taxation-statistics/.

Australian Bureau of Statistics (2020). *Australian and New Zealand Standard Industrial Classification (ANZSIC)*, 2006 revision with 2020 update. Catalogue 1292.0.

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  s <- ato_companies(year = "2022-23", table = "snapshot")
  head(s)
  m <- ato_companies(year = "2022-23", industry = "mining")
  head(m)
  # Multi-year industry panel
  panel <- ato_companies(year = c("2021-22", "2022-23"))
})
options(op)
```

---

ato_compliance *ATO compliance program outcomes*

Description

Returns the ATO's annual compliance program outcomes: audit yield (tax raised from audits), settled disputes, collectable debt, and compliance cost recovery. These appear in the ATO annual report and related data.gov.au releases.

Usage

```
ato_compliance(year = "latest", metric = c("overview", "debt", "audit"))
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.
`metric`	One of `"overview"` (default), `"debt"` (collectable vs insolvency vs disputed), or `"audit"` (liabilities raised by program area).

Value

An `ato_tbl`.

Source

Australian Taxation Office annual report data. Licensed CC BY 3.0 AU.

See Also

Other specialist: ato_division293(), ato_fuel_tax_credits(), ato_international(), ato_medicare_levy(), ato_prrt(), ato_rba(), ato_state_tax(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_compliance(year = "2022-23", metric = "debt"))
options(op)
```

---

ato_crosswalk *Load a bundled ATO crosswalk table*

Description

Returns one of the bundled classification crosswalks. Used internally by ato_harmonise() and available for user-level panel work.

Usage

```
ato_crosswalk(name = c("anzsic", "anzsco", "postcode", "cpi", "erp", "budget"))
```

Arguments

name	One of "anzsic", "anzsco", "postcode", "cpi", "erp", "budget".

Details

Bundled crosswalks (at division/major-group level):

- "anzsic": ANZSIC 2006 to 2020 (19 divisions, complete)
- "anzsco": ANZSCO 2013 to 2021 (8 major groups, complete)
- "postcode": postcode first-digit to state anchors
- "cpi": ABS CPI annual, base 2011-12 = 1.0
- "erp": ABS Estimated Resident Population, June 30 annual
- "budget": Final Budget Outcome reference totals

For 4-digit ANZSIC, 6-digit ANZSCO, or postcode-to-SA2/LGA/CED crosswalks, fetch the full tables from ABS. The bundled division/major-group level covers cross-year ATO Taxation Statistics joins at the industry headings used in all ATO tables.

Value

A data frame.

References

Australian Bureau of Statistics (2006). *Australian and New Zealand Standard Industrial Classification (ANZSIC)*. Catalogue 1292.0.

Australian Bureau of Statistics (2020). *ANZSIC 2006 Update*, cat. 1292.0, divisional structure. Used by ATO Taxation Statistics from 2022-23.

Australian Bureau of Statistics (2013). *Australian and New Zealand Standard Classification of Occupations (ANZSCO)*. Catalogue 1220.0.

Australian Bureau of Statistics (2022). *ANZSCO Revised Edition*, cat. 1220.0. Used by ATO Taxation Statistics from 2022-23 onward.

See Also

Other harmonisation: `ato_deflate()`, `ato_harmonise()`, `ato_per_capita()`, `ato_reconcile()`, `ato_schema_map()`, `ato_to_taxstats()`

Examples

```
ato_crosswalk("anzsic")
ato_crosswalk("cpi")
```

---

ato_deflate *Deflate nominal AUD to real AUD*

Description

Converts a numeric vector of nominal AUD figures indexed by financial year to real AUD of a chosen base year using the bundled ABS CPI series (annual, All Groups Australia, 2011-12 = 1.0). For the user's `inflateR` workflow in non-Australian contexts, bundle a matching CPI series and call this with a custom `cpi =` argument.

Usage

```
ato_deflate(x, year, base = "2022-23", cpi = NULL)
```

Arguments

`x`	Numeric vector of nominal AUD values.
`year`	Character vector of financial years for each entry in `x`, in `"YYYY-YY"` form. Must be the same length as `x`.
`base`	Base financial year for real terms (default `"2022-23"`).
`cpi`	Optional override: a data frame with columns `financial_year` and `cpi_all_groups_australia`. Default uses the bundled ABS series.

Details

Uses proportional (Laspeyres-style) adjustment: $real = nominal \times (CPI_{base}/CPI_{source})$. The bundled CPI is the ABS annual All Groups Australia index published in cat. 6401.0, rebased so that 2011-12 = 1.000. This is the standard rebasing used in most Australian time-series work and is consistent with ABS System of National Accounts methodology (cat. 5204.0).

The formula is exact for a chain-linked index after 1949 (when the ABS CPI was introduced) and approximate for earlier values that rely on Commonwealth Statistician retail-price series. Use a custom `cpi =` argument if you need a different deflator (e.g. GDP deflator, wage price index, or industry-specific PPI).

Value

Numeric vector of real AUD values in base-year prices.

References

Australian Bureau of Statistics (2024). *Consumer Price Index, Australia: Concepts, Sources and Methods*. Catalogue 6461.0.

Australian Bureau of Statistics (2024). *Consumer Price Index, Australia*. Catalogue 6401.0.

Diewert, W.E. (1998). "Index Number Issues in the Consumer Price Index." *Journal of Economic Perspectives*, 12(1), 47-58. doi:10.1257/jep.12.1.47

See Also

Other harmonisation: `ato_crosswalk()`, `ato_harmonise()`, `ato_per_capita()`, `ato_reconcile()`, `ato_schema_map()`, `ato_to_taxstats()`

Examples

```
ato_deflate(c(100, 100, 100),
            year = c("2012-13", "2017-18", "2022-23"),
            base = "2022-23")
```

`ato_deposit_zenodo` *Prepare a Zenodo deposit payload for the session manifest*

Description

Builds the JSON metadata payload Zenodo expects for a data deposit, using the current `ato_manifest()` and the snapshot pin set via `ato_snapshot()`. The function does NOT upload by default; it returns the payload and saved manifest path so you can inspect before calling with `upload = TRUE`.

Usage

```
ato_deposit_zenodo(
  title = NULL,
  description = NULL,
  creators = list(list(name = "Anonymous")),
  keywords = c("ATO", "taxation", "Australia", "reproducibility"),
  upload = FALSE,
  sandbox = FALSE,
  token = Sys.getenv("ZENODO_TOKEN")
)
```

Arguments

`title`	Deposit title. Defaults to "ATO data snapshot YYYY-MM-DD" using the current snapshot pin.
`description`	Free-text description. Defaults to a short auto-generated note listing the datasets fetched.
`creators`	List of creator records. Each should be a list with `name`, optional `affiliation`, `orcid`. Defaults to a single anonymous entry; override for published work.
`keywords`	Character vector of keywords. Defaults to `c("ATO", "taxation", "Australia", "reproducibility")`.
`upload`	Logical; if `TRUE`, POSTs the deposit to Zenodo and uploads the manifest CSV. Default `FALSE` (dry run).
`sandbox`	Logical; if `TRUE`, uses Zenodo Sandbox (sandbox.zenodo.org) for testing. Default `FALSE`.
`token`	Zenodo personal access token. Defaults to `Sys.getenv("ZENODO_TOKEN")`.

Details

To upload, supply a Zenodo personal access token via the `ZENODO_TOKEN` environment variable (or the `token` argument). Tokens can be generated at https://zenodo.org/account/settings/applications/.

Value

A list with `payload` (the JSON metadata), `manifest_path` (where the CSV manifest was staged), and if `upload = TRUE`, `deposit_id`, `doi_prereserve`, and `url`.

See Also

Other reproducibility: `ato_manifest()`, `ato_manifest_clear()`, `ato_manifest_write()`, `ato_sha256()`, `ato_snapshot()`

Examples

```
ato_snapshot("2026-04-24")
ato_deposit_zenodo(
  title = "ATO data snapshot for working paper v1",
  creators = list(list(name = "Coverdale, Charles")),
  upload = FALSE
)
```

`ato_division293` *Division 293 tax assessments (high-income super contributions)*

Description

Returns Division 293 tax data: number of assessments, average Division 293 liability, and distribution by income band. Division 293 applies an extra 15% tax on concessional super contributions for individuals with combined income plus low-tax super contributions above AUD 250,000. Central to retirement-income reform analysis (e.g. Grattan's "Better Super" proposals).

Usage

```
ato_division293(year = "latest")
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.

Details

Published as part of the Individuals Taxation Statistics (Table 3b in recent releases).

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics Individuals. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Income Tax Assessment Act 1997*, Division 293. Extra 15 per cent tax on concessional super contributions for high-income earners.

Daley, J., Coates, B. and Wood, D. (2018). *Money in retirement: more than enough*. Grattan Institute. Uses Division 293 distributional data in reform analysis.

See Also

Other specialist: ato_compliance(), ato_fuel_tax_credits(), ato_international(), ato_medicare_levy(), ato_prrt(), ato_rba(), ato_state_tax(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_division293(year = "2022-23"))
options(op)
```

---

ato_download *Download a resource from an ATO dataset*

Description

Low-level helper for arbitrary CKAN resources. Resolves the package by id (slug) and picks the first resource matching `pattern`, or the first resource if `pattern` is `NULL`.

Usage

```
ato_download(
  id,
  pattern = NULL,
  parse = c("auto", "csv", "xlsx", "none"),
  sheet = 1
)
```

Arguments

`id`	CKAN package id (e.g. `"taxation-statistics-2022-23"` or `"corporate-transparency"`).
`pattern`	Optional regex applied to the resource filename and name.
`parse`	One of `"auto"` (default), `"csv"`, `"xlsx"`, or `"none"` (returns the cached file path).
`sheet`	For XLSX resources: sheet index or name.

Value

Either a file path (`parse = "none"`) or an `ato_tbl`.

See Also

Other discovery: ato_catalog(), ato_charities(), ato_cite(), ato_excise(), ato_fbt(), ato_help(), ato_irpd(), ato_payg(), ato_rdti(), ato_sme_benchmarks(), ato_tax_gaps(), ato_top_taxpayers(), ato_vttc()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  cat <- ato_download("corporate-transparency",
                      pattern = "2023",
                      parse = "csv")
})
options(op)
```

`ato_excise` *Excise and fuel tax credit rates and clearances*

Description

Returns ATO excise data, covering four sub-releases:

- **beer** : beer clearances summary (volumes by product class)
- **spirits** : spirits and other excisable beverages clearances
- **excise_rates** : historical excise rate schedule (all excise categories, quarterly indexed rates)
- **ftc_rates** : historical Fuel Tax Credit rates

Usage

```
ato_excise(table = c("excise_rates", "ftc_rates", "beer", "spirits"))
```

Arguments

`table`	One of `"beer"`, `"spirits"`, `"excise_rates"` (default), or `"ftc_rates"`.

Value

An `ato_tbl`. Rates are in AUD per litre (or per kg for tobacco); volumes are in megalitres or similar.

Source

Australian Taxation Office excise data. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Excise Act 1901*; *Excise Tariff Act 1921*; *Fuel Tax Act 2006*.

Australian Taxation Office (annual). *Excise data: methodology and indexation notes*. Excise rates are indexed to the Consumer Price Index twice a year (February and August) for most commodities.

Productivity Commission (2016). *Migrant Intake into Australia* (for tobacco excise distributional analysis); *Harmful Drinking* inquiry (for alcohol excise distributional analysis).

See Also

Other discovery: ato_catalog(), ato_charities(), ato_cite(), ato_download(), ato_fbt(), ato_help(), ato_irpd(), ato_payg(), ato_rdti(), ato_sme_benchmarks(), ato_tax_gaps(), ato_top_taxpayers(), ato_vttc()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  rates <- ato_excise("excise_rates")
  head(rates)
})
options(op)
```

---

ato_fbt *Fringe Benefits Tax statistics*

Description

Returns the ATO's annual Fringe Benefits Tax (FBT) Taxation Statistics: employer counts, gross taxable value, FBT payable, and employee benefit counts by benefit type and industry. Used by Treasury, PBO, and researchers evaluating the FBT concession system (electric vehicles, remote area exemptions, novated leases).

Usage

```
ato_fbt(year = "latest")
```

Arguments

year	Income year in "YYYY-YY" form (e.g. "2022-23") or "latest".

Value

An ato_tbl. Monetary values in nominal AUD.

Source

Australian Taxation Office FBT Taxation Statistics on data.gov.au. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Fringe Benefits Tax Assessment Act 1986*. Substantive FBT law; ATO rulings (TR series) elaborate taxable-value methodology.

Australian Taxation Office (annual). *FBT explanatory notes*. Definitions of reportable benefits, gross-up factors (Type 1 and Type 2), and otherwise-deductible rule.

Treasury (2022). *Electric Car Discount Bill*. Explanatory memorandum for the EV FBT exemption introduced 1 July 2022.

See Also

Other discovery: ato_catalog(), ato_charities(), ato_cite(), ato_download(), ato_excise(), ato_help(), ato_irpd(), ato_payg(), ato_rdti(), ato_sme_benchmarks(), ato_tax_gaps(), ato_top_taxpayers(), ato_vttc()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  fbt <- ato_fbt(year = "2022-23")
  head(fbt)
})
options(op)
```

---

ato_fuel_tax_credits *Fuel Tax Credits by industry and claim period*

Description

Returns the Fuel Tax Credits scheme data: entitlement rates by fuel type, claim totals by industry. FTC is a major implicit fossil-fuel subsidy and is a key lens for decarbonisation policy cost-benefit analysis.

Usage

```
ato_fuel_tax_credits(year = "latest", by = c("industry", "fuel", "period"))
```

Arguments

year	"YYYY-YY" or "latest".
by	One of "industry" (default, by ANZSIC division), "fuel" (by fuel type), or "period" (quarterly rates).

Details

The ATO publishes FTC data as part of the Excise Data release and in standalone FTC tables.

Value

An ato_tbl.

Source

Australian Taxation Office Excise and Fuel Tax Credit data. Licensed CC BY 3.0 AU.

References

Commonwealth of Australia. *Fuel Tax Act 2006*; *Fuel Tax (Consequential and Transitional Provisions) Act 2006*.

Denniss, R. and Grudnoff, M. (2021). *Fossil fuel subsidies in Australia*. The Australia Institute. FTC-as- subsidy framing used in decarbonisation policy analysis.

Intergovernmental Panel on Climate Change (2022). *Climate Change 2022: Mitigation of Climate Change*. Chapter 13 covers fossil-fuel subsidy reform.

See Also

Other specialist: `ato_compliance()`, `ato_division293()`, `ato_international()`, `ato_medicare_levy()`, `ato_prrt()`, `ato_rba()`, `ato_state_tax()`, `ato_tax_expenditures()`, `ato_whm()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try(head(ato_fuel_tax_credits(year = "latest", by = "industry")))
options(op)
```

---

ato_gst *GST and activity statement ratios*

Description

Returns the Taxation Statistics GST tables (T1-T5) or the Activity Statement Ratios (A1-A5) for the requested year.

Usage

```
ato_gst(year = "latest", table = c("overview", "state", "industry", "ratios"))
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.
`table`	One of `"overview"` (default, GST T1), `"state"` (GST by state), `"industry"` (GST by ANZSIC), or `"ratios"` (Activity Statement Ratios).

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: GST and Activity Statement Ratios explanatory notes*.

Commonwealth of Australia. *A New Tax System (Goods and Services Tax) Act 1999*. Enabling legislation for the 10 per cent value-added tax introduced 1 July 2000.

Productivity Commission (2018). *Horizontal Fiscal Equalisation*. Background reference on the GST distribution formula across states.

See Also

Other gst: `ato_industry()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  g <- ato_gst(year = "2022-23", table = "industry")
  head(g)
})
options(op)
```

`ato_harmonise` *Harmonise column names in a multi-year ATO panel*

Description

ATO renames columns across annual releases; a stacked panel from `ato_individuals_postcode(year = c("2020-21", "2021-22"))` may have inconsistent names like `total_income` vs `total_income_or_loss`. `ato_harmonise()` renames columns to the first variant in `ATO_COL_VARIANTS` so panels are join-ready.

Usage

```
ato_harmonise(df)
```

Arguments

`df`	A data frame (typically an `ato_tbl` with `year` column from a multi-year call).

Details

Unknown columns are left alone. Columns that collide after renaming (because two variants map to the same canonical name) emit a warning; the first column wins.

Value

A data frame with harmonised names. `ato_tbl` class and provenance attributes are preserved.

See Also

Other harmonisation: ato_crosswalk(), ato_deflate(), ato_per_capita(), ato_reconcile(), ato_schema_map(), ato_to_taxstats()

Examples

```
df <- data.frame(postcode = "2000",
                 total_income_or_loss = 100,
                 state_territory = "NSW")
ato_harmonise(df)
```

---

ato_help *Study and Training Support Loan data (HELP, AASL, VSL)*

Description

Returns aggregate statistics on Australia's three main education-loan schemes:

- **HELP** (Higher Education Loan Program, ~3m borrowers, AUD 80bn+ outstanding debt)
- **AASL** (Australian Apprenticeship Support Loans, previously Trade Support Loans)
- **VSL** (VET Student Loans, vocational education loans)

Usage

```
ato_help(scheme = c("help", "aasl", "vsl"))
```

Arguments

scheme	One of "help" (default), "aasl", or "vsl".

Details

Headline covers: new loans by income range, outstanding debt by age and gender, repayment rates, median debt on entry. Used by Treasury (PBO costings of HELP indexation changes) and education policy researchers.

Value

An ato_tbl. All dollar values in nominal AUD.

Source

Australian Taxation Office Study and Training Support Loans statistics. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Higher Education Support Act 2003*; *VET Student Loans Act 2016*.

Australian Department of Education (annual). *Higher Education Statistics: HELP statistics collection*.

Norton, A. and Cherastidtham, I. (2018). *Mapping Australian higher education*. Grattan Institute. Methodology reference for HELP repayment projections.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  help <- ato_help(scheme = "help")
  head(help)
})
options(op)
```

---

ato_individuals *Individual Taxation Statistics snapshot*

Description

Returns the Individuals Table 1 snapshot: aggregate counts, total income, taxable income, tax payable, and deductions across all individual returns (roughly 14 million per year). The snapshot is the headline table; for finer cuts use the dedicated functions:

- `ato_individuals_postcode()` for geographic breakdowns,
- `ato_individuals_occupation()` for occupation × sex × income-range detail, or
- `ato_download()` with a custom `pattern` for specific Tables 2 to 27 (age, sex, state, industry, source of income, deductions, offsets, CGT, non-residents).

Usage

```
ato_individuals(year = "latest")
```

Arguments

`year`	Year in `"YYYY-YY"` form (e.g. `"2022-23"`) or `"latest"`. `"latest"` resolves to the most recently published release (currently 2022-23).

Details

Monetary values are nominal AUD of the reporting year. Use `inflateR::inflate()` or the ABS CPI series if you need real-term comparisons.

Value

An `ato_tbl` with one row per aggregate line-item and columns for count and amount in nominal AUD.

Source

Australian Taxation Office Taxation Statistics https://www.ato.gov.au/about-ato/research-and-statistics/. Licensed CC BY 2.5 AU.

See Also

Other individuals: `ato_individuals_age()`, `ato_individuals_occupation()`, `ato_individuals_postcode()`, `ato_individuals_sex()`, `ato_individuals_state()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  ind <- ato_individuals(year = "2022-23")
  head(ind)
})
options(op)
```

`ato_individuals_age` *Individual tax data by age range*

Description

Returns Taxation Statistics Individuals Table 2 (approximately): counts, total income, taxable income, and tax payable by age range and (usually) sex. Age ranges are 5-year bands for most of working life plus wider bands at the tails.

Usage

```
ato_individuals_age(year = "latest", sex = c("all", "male", "female"))
```

Arguments

`year`	`"YYYY-YY"`, `"latest"`, or a vector of years.
`sex`	One of `"all"` (default), `"male"`, or `"female"`.

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics Individuals. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: Individuals explanatory notes*. Age-range breakdowns use the taxpayer's reported date of birth at lodgement; sex is self-reported on the return.

See Also

Other individuals: `ato_individuals()`, `ato_individuals_occupation()`, `ato_individuals_postcode()`, `ato_individuals_sex()`, `ato_individuals_state()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_individuals_age(year = "2022-23", sex = "female"))
options(op)
```

---

ato_individuals_occupation

Individual tax data by occupation

Description

Returns the Individuals Table 14 (occupation by sex by taxable income range). Around 1,000 occupations classified by ANZSCO with aggregate counts, total income, taxable income, and tax payable. The ATO migrated from ANZSCO 2013 to ANZSCO 2021 across the 2022-23 release; cross-year joins on occupation name or code must account for the recode.

Usage

```
ato_individuals_occupation(
  year = "latest",
  occupation = NULL,
  sex = c("all", "male", "female", "m", "f")
)
```

Arguments

year	"YYYY-YY", "latest", or a vector of years for a multi-year panel (e.g. c("2020-21", "2021-22", "2022-23")).
occupation	Optional substring filter (case-insensitive) applied to the occupation description column.
sex	One of "all" (default), "male", or "female". Rows with sex recorded as "Not stated" are dropped when filtering to male or female. Short forms "m"/"f" are accepted.

Details

Classification break. Releases from 2022-23 onwards use ANZSCO 2021; earlier releases use ANZSCO 2013. A warning is emitted when the requested year(s) are at or after this boundary, or when a multi-year request spans it.

Value

An ato_tbl with one row per occupation-sex-income combination. Multi-year queries add a year column. Monetary values in nominal AUD of the reporting year.

Source

Australian Taxation Office Taxation Statistics. Licensed CC BY 2.5 AU.

See Also

Other individuals: ato_individuals(), ato_individuals_age(), ato_individuals_postcode(), ato_individuals_sex(), ato_individuals_state()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  occ <- ato_individuals_occupation(year = "2022-23",
                                    occupation = "economist")
  head(occ)
  # Multi-year panel
  panel <- ato_individuals_occupation(year = c("2021-22", "2022-23"),
                                      occupation = "nurse")
})
options(op)
```

`ato_individuals_postcode` *Individual tax data by postcode*

Description

Returns the Individuals Table 6 (or standalone postcode dataset): taxable income, tax payable, and return counts by 4-digit postcode. Headline dataset for income-distribution journalism.

Usage

```
ato_individuals_postcode(year = "latest", state = NULL, postcode = NULL)
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`. Pass a vector of years (e.g. `c("2020-21", "2021-22", "2022-23")` or `2018:2022`) to stack multiple years with a `year` column added to the output. Useful for time-series analysis.
`state`	Optional character vector of state codes (e.g. `"NSW"`, `c("VIC", "QLD")`).
`postcode`	Optional character vector of 4-digit postcodes.

Details

Privacy suppression. The ATO suppresses postcodes with fewer than 50 returns; those cells are returned as `NA` after parsing (the package maps `"np"`, `"*"`, and similar tokens to `NA` so numeric columns stay numeric). Small or remote postcodes will be silently missing from the output.

Monetary values are nominal AUD of the reporting year. Use `inflateR::inflate()` for real-term series.

Value

An `ato_tbl` with one row per postcode (or per postcode per year for multi-year queries), including state, return count, total income, taxable income, and tax payable. Schema drifts year to year (SA3/SA4 columns present from 2017 onwards).

Source

Australian Taxation Office Taxation Statistics postcode release. Licensed CC BY 2.5 AU.

References

Atkinson, A.B. and Leigh, A. (2007). "The Distribution of Top Incomes in Australia." *Economic Record*, 83(262), 247-261. doi:10.1111/j.14754932.2007.00412.x

Burkhauser, R.V., Hahn, M.H. and Wilkins, R. (2015). "Measuring top incomes using tax record data: a cautionary tale from Australia." *Journal of Economic Inequality*, 13(2), 181-205. doi:10.1007/s108880149281z

See Also

Other individuals: ato_individuals(), ato_individuals_age(), ato_individuals_occupation(), ato_individuals_sex(), ato_individuals_state()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  # Single year
  p <- ato_individuals_postcode(year = "2022-23", state = "NSW")
  head(p)
  # Multi-year stack with year column
  panel <- ato_individuals_postcode(year = c("2020-21", "2021-22"),
                                    state = "NSW")
})
options(op)
```

---

ato_individuals_sex *Individual tax data by sex*

Description

Returns counts and aggregates split by sex. Thin wrapper around the ATO "Selected items by sex" table.

Usage

```
ato_individuals_sex(year = "latest")
```

Arguments

year	"YYYY-YY", "latest", or a vector of years.

Value

An ato_tbl.

Source

Australian Taxation Office Taxation Statistics Individuals. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: Individuals explanatory notes*. Age-range breakdowns use the taxpayer's reported date of birth at lodgement; sex is self-reported on the return.

See Also

Other individuals: ato_individuals(), ato_individuals_age(), ato_individuals_occupation(), ato_individuals_postcode(), ato_individuals_state()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_individuals_sex(year = "2022-23"))
options(op)
```

---

ato_individuals_state *Individual tax data by state or territory*

Description

Returns counts and aggregates by state. Thin wrapper around the ATO "Selected items by state/territory" table.

Usage

```
ato_individuals_state(year = "latest")
```

Arguments

year	"YYYY-YY", "latest", or a vector of years.

Value

An ato_tbl.

Source

Australian Taxation Office Taxation Statistics Individuals. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: Individuals explanatory notes*. Age-range breakdowns use the taxpayer's reported date of birth at lodgement; sex is self-reported on the return.

See Also

Other individuals: ato_individuals(), ato_individuals_age(), ato_individuals_occupation(), ato_individuals_postcode(), ato_individuals_sex()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_individuals_state(year = "2022-23"))
options(op)
```

`ato_industry` *Industry aggregates across entity types*

Description

Derived helper that returns an ANZSIC industry breakdown based on either individual, company, or all entities for the year.

Usage

```
ato_industry(
  year = "latest",
  entity = c("company", "individual", "all"),
  anzsic = NULL
)
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.
`entity`	One of `"individual"`, `"company"` (default), or `"all"`.
`anzsic`	Optional substring filter on industry name.

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics. Licensed CC BY 2.5 AU.

See Also

Other gst: `ato_gst()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  i <- ato_industry(year = "2022-23", entity = "company",
                    anzsic = "manufacturing")
  head(i)
})
options(op)
```

`ato_international` *OECD Revenue Statistics comparison*

Description

Fetches OECD Revenue Statistics for cross-country tax-to-GDP benchmarking. Returns tax revenue as percent of GDP by tax category. Use to contextualise Australian ATO aggregates in cross-country policy arguments (e.g. OECD average corporate tax-to-GDP, international ranks for personal income tax).

Usage

```
ato_international(country = "AUS", year = "latest")
```

Arguments

`country`	Country ISO code or name (default `"AUS"`).
`year`	Four-digit year or `"latest"`.

Details

Thin wrapper pointing users to `readoecd::` for full OECD API access; returns a minimal tax-to-GDP slice here for convenience.

Value

An `ato_tbl` with columns `country`, `year`, `tax`, `pct_gdp`.

Source

OECD Revenue Statistics https://www.oecd.org/tax/tax-policy/revenue-statistics.htm.

See Also

Other specialist: `ato_compliance()`, `ato_division293()`, `ato_fuel_tax_credits()`, `ato_medicare_levy()`, `ato_prrt()`, `ato_rba()`, `ato_state_tax()`, `ato_tax_expenditures()`, `ato_whm()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_international(country = "AUS"))
options(op)
```

`ato_irpd` *International Related Party Dealings (IRPD)*

Description

Returns the ATO's International Related Party Dealings data, which captures intra-group cross-border payments and receivables reported by Australian corporate taxpayers. Core dataset for BEPS and transfer-pricing research, transfer pricing risk assessment, and multinational tax analysis.

Usage

```
ato_irpd(year = "latest", table = 1L)
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2023-24"`) or `"latest"`.
`table`	Integer 1, 2, or 3. Default 1.

Details

The IRPD data is published as a separate CKAN package per income year (2019-20 through 2023-24). Each annual package contains three tables:

- **Table 1** : IRPD totals from 2015-16 to the current year
- **Table 2** : IRPDs by jurisdiction
- **Table 3** : Index of chart data

Value

An `ato_tbl`. Monetary values in nominal AUD.

Source

Australian Taxation Office International Related Party Dealings release. Licensed CC BY 2.5 AU.

References

Organisation for Economic Co-operation and Development (2015). *Transfer Pricing Documentation and Country-by-Country Reporting, Action 13: 2015 Final Report*. OECD/G20 Base Erosion and Profit Shifting Project, Paris. doi:10.1787/9789264241480en

Commonwealth of Australia. *Income Tax Assessment Act 1997*, Subdivision 815-B (Transfer Pricing); *Multinational Anti-Avoidance Law* (MAAL) and *Diverted Profits Tax*.

Australian Taxation Office (annual). *International Dealings Schedule (IDS) instructions*. Reporting framework underlying the IRPD dataset.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  by_jurisdiction <- ato_irpd(year = "2023-24", table = 2)
  head(by_jurisdiction)
})
options(op)
```

---

ato_manifest *Return the session manifest of fetched ATO datasets*

Description

Every call to a data function (`ato_individuals()`, `ato_companies()`, etc.) appends one row to the session manifest, recording URL, dataset title, CKAN resource and package IDs where resolvable, SHA-256 of the cached file, size, retrieval timestamp, and the snapshot pin set via `ato_snapshot()`. Duplicate URLs within a session are deduplicated (last fetch wins).

Usage

```
ato_manifest(format = c("df", "yaml", "json"))
```

Arguments

`format`	One of `"df"` (default, tidy data frame), `"yaml"`, or `"json"`.

Details

Attach the output to your paper's appendix, deposit it to Zenodo with `ato_deposit_zenodo()` to mint a DOI, or export with `ato_manifest_write()` for CI artefacts.

Value

A data frame, YAML string, or JSON string depending on `format`.

See Also

Other reproducibility: `ato_deposit_zenodo()`, `ato_manifest_clear()`, `ato_manifest_write()`, `ato_sha256()`, `ato_snapshot()`

Examples

```
op <- options(ato.cache_dir = tempdir())
ato_manifest_clear()
ato_snapshot("2026-04-24")
try(ato_individuals(year = "2022-23"))
ato_manifest()
options(op)
```

---

ato_manifest_clear *Clear the session manifest*

Description

Clear the session manifest

Usage

```
ato_manifest_clear()
```

Value

Invisibly `NULL`. Useful at the top of a script when running repeatedly.

See Also

Other reproducibility: `ato_deposit_zenodo()`, `ato_manifest()`, `ato_manifest_write()`, `ato_sha256()`, `ato_snapshot()`

Examples

```
ato_manifest_clear()
```

ato_manifest_write *Write the session manifest to a file*

Description

Writes the manifest to a file in the requested format. Call at the end of an analysis script; commit the manifest alongside the paper for full reproducibility.

Usage

```
ato_manifest_write(path, format = c("auto", "csv", "yaml", "json"))
```

Arguments

path	Output file path. Extension determines format if format = "auto": .csv to CSV, .yaml/.yml to YAML, .json to JSON.
format	One of "auto" (infer from extension), "csv", "yaml", or "json".

Value

Invisibly, the absolute path to the written file.

See Also

Other reproducibility: ato_deposit_zenodo(), ato_manifest(), ato_manifest_clear(), ato_sha256(), ato_snapshot()

Examples

```
p <- tempfile(fileext = ".csv")
ato_manifest_clear()
ato_manifest_write(p)
```

ato_medicare_levy *Medicare Levy and Medicare Levy Surcharge*

Description

Returns aggregate Medicare Levy and MLS data from Taxation Statistics Individuals. The 2% Medicare Levy is on most taxable income; MLS is an additional 1.0 to 1.5% on high-income earners without adequate private hospital cover. Used in private health insurance reform analysis.

Usage

```
ato_medicare_levy(year = "latest", component = c("levy", "surcharge"))
```

Arguments

year	"YYYY-YY" or "latest".
component	One of "levy" (default, standard Medicare Levy) or "surcharge" (MLS).

Value

An ato_tbl.

Source

Australian Taxation Office Taxation Statistics Individuals. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Medicare Levy Act 1986*; *A New Tax System (Medicare Levy Surcharge – Fringe Benefits) Act 1999*.

Productivity Commission (2015). *Efficiency in Health*. Analysis of Medicare Levy and MLS distributional effects.

See Also

Other specialist: ato_compliance(), ato_division293(), ato_fuel_tax_credits(), ato_international(), ato_prrt(), ato_rba(), ato_state_tax(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_medicare_levy(year = "2022-23", component = "surcharge"))
options(op)
```

---

ato_meta *Fetch CKAN metadata for an ATO dataset*

Description

Returns structured metadata for any ATO dataset on data.gov.au: title, notes, licence, last-modified timestamp, resource count, and all resource URLs. Useful for detecting silent updates before clearing the cache, or for auditing what version of data you have.

Usage

```
ato_meta(x)
```

Arguments

x	Either an ato_tbl (as returned by any ato_* data function) or a character CKAN package ID / slug (e.g. "taxation-statistics-2022-23", "corporate-transparency").

Value

A list with elements:

- `id`: CKAN package slug
- `title`: human-readable title
- `notes`: dataset description (truncated to 400 chars)
- `licence`: licence title
- `metadata_modified`: ISO timestamp of last CKAN update
- `n_resources`: number of downloadable files
- `resource_urls`: character vector of all resource URLs

See Also

Other configuration: `ato_cache_info()`, `ato_clear_cache()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  # By package ID
  m <- ato_meta("taxation-statistics-2022-23")
  m$metadata_modified

  # From an ato_tbl
  tbl <- ato_individuals(year = "2022-23")
  ato_meta(tbl)
})
options(op)
```

---

ato_payg *PAYG withholding data*

Description

Returns the ATO's Pay As You Go (PAYG) withholding data: employer counts, total withholding amounts, and employee counts by industry and state. Used by researchers studying labour market taxation, wage growth, and employer compliance.

Usage

```
ato_payg(year = "latest")
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2022-23"`) or `"latest"`.

Value

An `ato_tbl`. Monetary values in nominal AUD.

Source

Australian Taxation Office PAYG withholding data on data.gov.au. Licensed CC BY 2.5 AU.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  payg <- ato_payg(year = "2022-23")
  head(payg)
})
options(op)
```

`ato_per_capita` *Express an aggregate per capita using ABS ERP*

Description

Express an aggregate per capita using ABS ERP

Usage

```
ato_per_capita(x, year, erp = NULL)
```

Arguments

`x`	Numeric vector of aggregate values (same length as `year`).
`year`	Character vector of financial years.
`erp`	Optional override: data frame with columns `financial_year` and `erp_june_australia_thousands`.

Details

Divides the input by Estimated Resident Population at 30 June of the financial year's end (a stock measure). For flow-style measures where a mid-year-average population is preferable, substitute a custom `erp =` argument. ERP is ABS's preferred population-denominator concept for per-capita economic statistics (see cat. 3101.0 methodology).

Value

Numeric vector of per-capita values (same units as `x` per person).

References

Australian Bureau of Statistics (2024). *National, State and Territory Population*. Catalogue 3101.0.

See Also

Other harmonisation: `ato_crosswalk()`, `ato_deflate()`, `ato_harmonise()`, `ato_reconcile()`, `ato_schema_map()`, `ato_to_taxstats()`

Examples

```
# Income tax per person, 2022-23 FBO headline
ato_per_capita(316.4e9, "2022-23")
```

`ato_prrt` *Petroleum Resource Rent Tax (PRRT) annual data*

Description

Returns PRRT revenue and assessments. PRRT is a 40% tax on the profits of offshore petroleum projects; revenues are volatile and project-specific. Key dataset for resource-tax reform analysis.

Usage

```
ato_prrt(year = "latest")
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics Company. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Petroleum Resource Rent Tax Assessment Act 1987*. Enabling legislation for the 40 per cent rent tax on offshore petroleum projects.

Callaghan, M. (2017). *Review of the Petroleum Resource Rent Tax*. Treasury-commissioned review; reference for PRRT-reform analysis.

See Also

Other specialist: ato_compliance(), ato_division293(), ato_fuel_tax_credits(), ato_international(), ato_medicare_levy(), ato_rba(), ato_state_tax(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_prrt(year = "2022-23"))
options(op)
```

---

ato_rba *RBA Commonwealth receipts (H1 statistical table)*

Description

Pointer to the RBA's H1 series on Commonwealth receipts for long-run time series. RBA compiles since 1959-60, filling gaps in ATO Taxation Statistics which start 1994-95.

Usage

```
ato_rba(series = c("receipts", "income_tax"))
```

Arguments

series	One of "receipts" (default, all Commonwealth receipts by category) or "income_tax" (income tax only).

Details

The RBA publishes H1 as an XLSX with stable URL. This function fetches it and returns a tidy tibble.

Value

An ato_tbl.

Source

Reserve Bank of Australia Statistical Tables H1 https://www.rba.gov.au/statistics/tables/.

See Also

Other specialist: ato_compliance(), ato_division293(), ato_fuel_tax_credits(), ato_international(), ato_medicare_levy(), ato_prrt(), ato_state_tax(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_rba(series = "receipts"))
options(op)
```

`ato_rdti` *R&D Tax Incentive claimants*

Description

Returns the annual "Report of data about Research and Development Tax Incentive entities": claimants, claimed expenditure, refundable and non-refundable tax offsets by industry and company size. Treasury and DISR use this series to evaluate the R&D Tax Incentive programme, which is the largest single element of Australia's business innovation policy (AUD 2 billion+ per year).

Usage

```
ato_rdti(year = "latest")
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2022-23"`) or `"latest"`. Current releases cover 2021-22 and 2022-23.

Value

An `ato_tbl` with one row per entity (or aggregated cell, depending on the release schema). Monetary values in nominal AUD.

Source

Australian Taxation Office Research and Development Tax Incentive report. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Income Tax Assessment Act 1997*, Division 355 (Research and Development Tax Incentive).

Department of Industry, Science and Resources and Australian Taxation Office (annual). *R&DTI Transparency Report*. Jointly administered programme methodology.

Ferris, B., Finkel, A. and Fraser, J. (2016). *Review of the R&D Tax Incentive*. Australian Government review (the "Three Fs review") informing subsequent programme design.

Organisation for Economic Co-operation and Development (annual). *R&D Tax Incentives Database*. International comparator data for R&D tax expenditures.

See Also

Other discovery: ato_catalog(), ato_charities(), ato_cite(), ato_download(), ato_excise(), ato_fbt(), ato_help(), ato_irpd(), ato_payg(), ato_sme_benchmarks(), ato_tax_gaps(), ato_top_taxpayers(), ato_vttc()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  rdti <- ato_rdti(year = "2022-23")
  head(rdti)
})
options(op)
```

---

ato_reconcile *Reconcile an aggregate against Commonwealth budget totals*

Description

Compares a scalar (or data frame total) against the published Final Budget Outcome figure for the same year and revenue line. Useful as a sanity check on an ATO Taxation Statistics sum before reporting it in a paper or brief.

Usage

```
ato_reconcile(value, year, measure, sum_column = NULL)
```

Arguments

value	Numeric; the figure to check, in AUD (not AUD billions). An ato_tbl can also be passed: pass sum_column to pick which numeric column to sum.
year	Financial year, e.g. "2022-23".
measure	One of the measure codes in ato_crosswalk("budget"), for example "individuals_income_tax_net" "company_tax_net", "gst_net", "fuel_excise_net".
sum_column	Column name to sum when value is a data frame. Default NULL (errors if multiple numeric columns exist).

Details

Discrepancies between ATO Taxation Statistics aggregates and the Final Budget Outcome (FBO) are expected and meaningful:

- Taxation Statistics are based on assessments made by a cut-off date (usually October of the following calendar year) and may exclude late-lodging returns.
- FBO figures are cash-basis Commonwealth receipts; Taxation Statistics are accrual-basis tax assessed.

- GST, excise, and fuel credits have timing and refund effects that further distort the cash-vs-assessment gap.

A 1-3 per cent gap is consistent with the accrual-to-cash reconciliation Treasury publishes in the FBO statement of revenues; larger gaps warrant investigation. The bundled reference totals in `inst/extdata/budget_reference_totals.csv` are taken from the relevant FBO release, with the precise table cited in the `source` column of each row.

Value

A one-row data frame: `measure`, `year`, `value_aud`, `reference_aud`, `diff_aud`, `pct_diff`, `source`. Emits a warning if `abs(pct_diff) > 0.05`.

References

Commonwealth of Australia (various years). *Final Budget Outcome*. The Treasury, Canberra. https://budget.gov.au/content/fbo/index.htm

Australian Bureau of Statistics (various years). *Taxation Revenue, Australia*. Catalogue 5506.0.

Australian Taxation Office (annual). *Australian tax gaps – overview*, methodology notes on accrual-vs-cash reconciliation.

See Also

Other harmonisation: `ato_crosswalk()`, `ato_deflate()`, `ato_harmonise()`, `ato_per_capita()`, `ato_schema_map()`, `ato_to_taxstats()`

Examples

```
ato_reconcile(value = 316.4e9,
              year = "2022-23",
              measure = "individuals_income_tax_net")
```

`ato_schema_map` *Print the ATO -> taxstats schema map*

Description

Convenience accessor for the bundled column-name mapping.

Usage

```
ato_schema_map()
```

Value

A data frame with columns `ato_aggregate` and `taxstats_microdata`.

See Also

Other harmonisation: ato_crosswalk(), ato_deflate(), ato_harmonise(), ato_per_capita(), ato_reconcile(), ato_to_taxstats()

Examples

```
head(ato_schema_map())
```

`ato_sha256` *Compute the SHA-256 digest of a file*

Description

Wraps tools::md5sum() style behaviour for SHA-256 via the digest package when available, or falls back to a pure-R implementation via tools::md5sum() + file length as a weaker check. For integrity work PBO/Grattan-grade, install the digest package (Suggests).

Usage

```
ato_sha256(file)
```

Arguments

file	Path to a local file.

Value

A length-1 character string (hex digest), or NA if the file does not exist.

See Also

Other reproducibility: ato_deposit_zenodo(), ato_manifest(), ato_manifest_clear(), ato_manifest_write(), ato_snapshot()

Examples

```
f <- tempfile()
writeLines("hello", f)
ato_sha256(f)
```

`ato_sme_benchmarks` *Small Business Benchmarks*

Description

Returns the ATO's Small Business Benchmarks: industry-specific performance ranges (cost of sales / turnover, total expenses / turnover, labour / turnover, etc.) derived from small-business income tax returns. Used by the ATO to identify outlier taxpayers, by small-business advisors for comparative analysis, and by tax integrity researchers.

Usage

```
ato_sme_benchmarks(year = "latest")
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2023-24"`) or `"latest"`. Releases available from 2016-17 onwards.

Value

An `ato_tbl` with one row per (industry, turnover band, ratio) combination. Ratios are percentages.

Source

Australian Taxation Office Small Business Benchmarks. Licensed CC BY 2.5 AU.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_tax_gaps()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  bm <- ato_sme_benchmarks(year = "2023-24")
  head(bm)
})
options(op)
```

`ato_snapshot` *Pin or inspect the session snapshot date*

Description

Call once at the top of an analysis script to declare the vintage of ATO data you intend to use. Every subsequent `ato_*` fetch records this date in the `ato_tbl` provenance header, in `ato_manifest()` entries, and in `ato_cite()` output. Combined with SHA-256 integrity (see `ato_sha256()` and `ato_manifest()`), this gives a reproducible audit trail acceptable for PBO or Grattan-style published work.

Usage

```
ato_snapshot(date)
```

Arguments

`date`	ISO `"YYYY-MM-DD"` character, `Date`, or `POSIXct`. Pass `NULL` to clear.

Details

If called with no arguments, returns the current pin (or `NULL` if unset).

Value

Invisibly, the new pinned date (as `Date`), or `NULL`.

See Also

Other reproducibility: `ato_deposit_zenodo()`, `ato_manifest()`, `ato_manifest_clear()`, `ato_manifest_write()`, `ato_sha256()`

Examples

```
ato_snapshot("2026-04-24")
ato_snapshot()
ato_snapshot(NULL)
```

ato_state_tax *State and territory tax revenue (ABS 5506.0)*

Description

Fetches the ABS Taxation Revenue collection (cat. 5506.0), which gives land tax, payroll tax, stamp duty, motor vehicle taxes, and other state taxes by jurisdiction. Needed for complete-tax-system analysis alongside ATO Commonwealth data.

Usage

```
ato_state_tax(year = "latest")
```

Arguments

year	"YYYY-YY" or "latest".

Value

An ato_tbl.

Source

Australian Bureau of Statistics, Taxation Revenue, catalogue 5506.0 https://www.abs.gov.au/statistics/economy/government/taxation-revenue-australia. Licensed CC BY 4.0.

See Also

Other specialist: ato_compliance(), ato_division293(), ato_fuel_tax_credits(), ato_international(), ato_medicare_levy(), ato_prrt(), ato_rba(), ato_tax_expenditures(), ato_whm()

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_state_tax(year = "latest"))
options(op)
```

`ato_super_funds` *Superannuation fund aggregates*

Description

Returns Taxation Statistics Super Funds tables or Self-Managed Superannuation Fund ('SMSF') aggregates, depending on `type`.

Usage

```
ato_super_funds(year = "latest", type = c("apra", "smsf", "all"))
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.
`type`	One of `"apra"` (APRA-regulated funds, default), `"smsf"` (SMSF statistical overview), or `"all"`.

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics Super Funds tables + SMSF statistical overview. Licensed CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Taxation Statistics: Super funds and SMSF explanatory notes*. Distinguishes reporting populations: APRA-regulated large funds, SMSFs, and Pooled Superannuation Trusts.

Australian Prudential Regulation Authority (annual). *Annual Superannuation Bulletin*. Complementary APRA-regulated fund statistics.

Commonwealth of Australia. *Superannuation Industry (Supervision) Act 1993* (SIS Act); *Superannuation Guarantee (Administration) Act 1992* (SGAA).

Productivity Commission (2018). *Superannuation: Assessing Efficiency and Competitiveness*. Inquiry report.

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  s <- ato_super_funds(year = "2022-23", type = "apra")
  head(s)
})
options(op)
```

`ato_tax_expenditures` *Tax Expenditures and Insights Statement (TEIS)*

Description

Returns the Treasury TEIS annual table of concession-by-concession tax expenditure estimates in AUD millions. TEIS is the authoritative cost-of-concessions dataset used in PBO and Grattan tax reform costings.

Usage

```
ato_tax_expenditures(year = "latest")
```

Arguments

`year`	Reference year for the TEIS release, e.g. `"2024"` or `"latest"`. Treasury publishes one TEIS per calendar year.

Details

TEIS is published by Treasury, not ATO; the function attempts a CKAN search on data.gov.au for the TEIS release, and falls back to the Treasury web URL if not indexed.

Key concessions covered: CGT main residence exemption, CGT 50% discount, superannuation earnings tax concession, franking credit refundability, work-related deductions, fuel tax credit scheme, R&D tax incentive, GST food exemption, and many more.

Value

An `ato_tbl` with one row per tax expenditure: label, category, estimated revenue forgone in AUD millions by year.

Source

Treasury Tax Expenditures and Insights Statement https://treasury.gov.au/publication/p2025-721342.

References

Commonwealth of Australia (annual). *Tax Expenditures and Insights Statement*. The Treasury, Canberra. https://treasury.gov.au/publication/p2025-721342

See Also

Other specialist: `ato_compliance()`, `ato_division293()`, `ato_fuel_tax_credits()`, `ato_international()`, `ato_medicare_levy()`, `ato_prrt()`, `ato_rba()`, `ato_state_tax()`, `ato_whm()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try(head(ato_tax_expenditures("latest")))
options(op)
```

`ato_tax_gaps` *Australian tax gaps estimates*

Description

Returns the ATO's annual Tax Gap publication: estimates of the difference between the tax theoretically payable under current law and the tax actually collected, across each tax type and taxpayer population (individuals not in business, small business, large corporate, GST, excise, fuel tax credits, PRRT, superannuation guarantee).

Usage

```
ato_tax_gaps(sheet = 1)
```

Arguments

`sheet`	Optional sheet name or index. The workbook contains separate sheets for each tax-gap population (e.g. "Large corporate", "Small business", "Individuals"). Pass the sheet name to extract a specific population. `NULL` (default) returns sheet 1 (overview).

Details

The Tax Gap series is used by Treasury (every MYEFO), the Parliamentary Budget Office, and academic researchers as the headline measure of revenue integrity.

Value

An `ato_tbl`. Tax-gap estimates are in nominal AUD millions of the reporting year and typically accompanied by a percentage-gap column.

Source

Australian Taxation Office Tax Gaps publication, CC BY 2.5 AU.

References

Australian Taxation Office (annual). *Australian tax gaps – overview*. Methodology notes on bottom-up, top-down, and random-inquiry approaches to the tax-gap estimation.

HMRC (annual). *Measuring tax gaps*. Sister methodology paper applied by HM Revenue and Customs in the UK; the ATO series was partly inspired by this literature.

Organisation for Economic Co-operation and Development (2017). *Shining Light on the Shadow Economy: Opportunities and Threats*. Paris. Synthesises tax-gap measurement practice across OECD member countries.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_top_taxpayers()`, `ato_vttc()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  gaps <- ato_tax_gaps()
  head(gaps)
})
options(op)
```

---

ato_top_taxpayers *Corporate Tax Transparency*

Description

Returns the ATO's annual Corporate Tax Transparency release, mandated by Part 5-25 of the *Taxation Administration Act 1953*. Covers every Australian public company, foreign-owned company, or Australian-owned private company above the AUD 100 million total-income threshold (the private-company threshold was lowered from AUD 200 million to AUD 100 million for the 2022-23 income year onwards, making all three categories uniform). The 2023-24 release was published 1 October 2025 and covered 4,110 entities.

Usage

```
ato_top_taxpayers(
  year = "latest",
  entity_type = c("all", "public", "private", "foreign"),
  sheet = c("income_tax", "prrt")
)
```

Arguments

year	"YYYY-YY" (e.g. "2023-24") or "latest".
entity_type	One of "all" (default), "public", "private", or "foreign". Matches the CTT Entity type column values "Australian public", "Australian private", "Foreign-owned".
sheet	One of "income_tax" (default, the ~4,000-entity income-tax sheet) or "prrt" (petroleum resource rent tax filers, typically 10-20 entities).

Details

The underlying XLSX has three sheets:

- **Information** (cover/metadata, ~7 rows).
- **Income tax details** (the headline dataset, ~4,000 entities: total income, taxable income, tax payable).
- **PRRT details** (petroleum resource rent tax filers, typically 10-20 entities).

Licensed under **CC BY 3.0 Australia** (the Corporate Tax Transparency and Voluntary Tax Transparency Code releases use CC BY 3.0 AU; most other Taxation Statistics use CC BY 2.5 AU).

Value

An ato_tbl with one row per disclosed entity. All monetary values are nominal AUD of the reporting year.

Source

Australian Taxation Office Corporate Tax Transparency release. Licensed CC BY 3.0 AU.

References

Commonwealth of Australia. *Taxation Administration Act 1953*, Part 5-25 (Corporate Tax Transparency).

Australian Taxation Office (annual). *Report of entity tax information*. The statutory Corporate Tax Transparency release.

Commonwealth Treasury (2013). *Improving the transparency of Australia's business tax system: Exposure draft explanatory memorandum*. Rationale for the Part 5-25 regime.

See Also

Other discovery: ato_catalog(), ato_charities(), ato_cite(), ato_download(), ato_excise(), ato_fbt(), ato_help(), ato_irpd(), ato_payg(), ato_rdti(), ato_sme_benchmarks(), ato_tax_gaps(), ato_vttc()

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  top <- ato_top_taxpayers(year = "2023-24")
  head(top)
  # Petroleum resource rent tax sheet
  prrt <- ato_top_taxpayers(year = "2023-24", sheet = "prrt")
  head(prrt)
})
options(op)
```

`ato_to_taxstats` *Remap an ato_tbl to the taxstats microdata column schema*

Description

Takes an `ato_tbl` with aggregate column names (produced by any `ato_*` function) and renames columns to match the `taxstats` (or `taxstats2`) 2% microdata sample schema used by Hugh Parsonage's DRAT package. Enables consistent variable definitions when moving between aggregate views and microdata prototyping.

Usage

```
ato_to_taxstats(df, direction = c("to_taxstats", "from_taxstats"))
```

Arguments

`df`	An `ato_tbl` or data frame.
`direction`	`"to_taxstats"` (default, aggregate -> microdata) or `"from_taxstats"` (microdata -> aggregate).

Details

The bundled schema map (`ato_schema_map()`) mirrors the column names from Parsonage's `taxstats` and `taxstats2` packages, which in turn use the ATO Individual Sample File variable names. Because `taxstats` is DRAT-distributed and not on CRAN, this function imposes the mapping as a static table rather than programmatically introspecting the `taxstats` namespace. Re-check the bundled map against the `taxstats` NAMESPACE when the ATO publishes a revised Sample File schema.

Unknown columns pass through unchanged. Use `ato_harmonise` first if the input panel has drift in source column names.

Value

A data frame with renamed columns. `ato_tbl` class and provenance attributes preserved.

References

Parsonage, H. (2019). *taxstats: 2 per cent Individual Sample File from the Australian Taxation Office*. R package (DRAT). https://github.com/HughParsonage/taxstats

Parsonage, H. (2024). *grattan: Perform Common Quantitative Tasks for Australian Analysts*. R package version 2026.1.1. https://cran.r-project.org/package=grattan

Australian Taxation Office (2024). *Taxation Statistics: Individual Sample File documentation*.

See Also

Other harmonisation: `ato_crosswalk()`, `ato_deflate()`, `ato_harmonise()`, `ato_per_capita()`, `ato_reconcile()`, `ato_schema_map()`

Examples

```
df <- data.frame(postcode = "2000", taxable_income = 80000,
                 medicare_levy = 1600)
ato_to_taxstats(df)
```

---

ato_vttc *Voluntary Tax Transparency Code disclosures*

Description

Returns the ATO's Voluntary Tax Transparency Code (VTTC) disclosures: large private companies that voluntarily publish tax information beyond the Corporate Tax Transparency mandate. Covers total income, taxable income, tax payable, and effective tax rate for each disclosing entity.

Usage

```
ato_vttc(year = "latest")
```

Arguments

`year`	Income year in `"YYYY-YY"` form (e.g. `"2022-23"`) or `"latest"`.

Details

The VTTC complements `ato_top_taxpayers()` (which covers mandatory CTT disclosures for entities above AUD 100m total income). VTTC signatories may be below or above the CTT threshold.

Licensed under **CC BY 3.0 Australia** (same as CTT data).

Value

An `ato_tbl`. Monetary values in nominal AUD.

Source

Australian Taxation Office Voluntary Tax Transparency Code disclosures on data.gov.au. Licensed CC BY 3.0 AU.

See Also

Other discovery: `ato_catalog()`, `ato_charities()`, `ato_cite()`, `ato_download()`, `ato_excise()`, `ato_fbt()`, `ato_help()`, `ato_irpd()`, `ato_payg()`, `ato_rdti()`, `ato_sme_benchmarks()`, `ato_tax_gaps()`, `ato_top_taxpayers()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try({
  vttc <- ato_vttc(year = "2022-23")
  head(vttc)
})
options(op)
```

---

ato_whm *Working Holiday Maker tax data*

Description

Returns aggregate Working Holiday Maker tax data: number of backpackers, total earnings, tax paid. Relevant for migration and labour-market policy analysis.

Usage

```
ato_whm(year = "latest")
```

Arguments

`year`	`"YYYY-YY"` or `"latest"`.

Value

An `ato_tbl`.

Source

Australian Taxation Office Taxation Statistics. Licensed CC BY 2.5 AU.

References

Commonwealth of Australia. *Migration Act 1958*, visa subclasses 417 and 462; *Working Holiday Maker Reform Act 2016*. Establishes the 15 per cent flat tax rate from the first dollar of WHM earnings.

Productivity Commission (2016). *Migrant Intake into Australia*. Includes WHM labour-market analysis.

See Also

Other specialist: `ato_compliance()`, `ato_division293()`, `ato_fuel_tax_credits()`, `ato_international()`, `ato_medicare_levy()`, `ato_prrt()`, `ato_rba()`, `ato_state_tax()`, `ato_tax_expenditures()`

Examples

```
op <- options(ato.cache_dir = tempdir())
try(ato_whm(year = "2022-23"))
options(op)
```

---

print.ato_tbl *Print an ato_tbl*

Description

Prints a provenance header (title, source, licence, retrieval time, dimensions) followed by the data frame.

Usage

```
## S3 method for class 'ato_tbl'
print(x, ...)
```

Arguments

`x`	An `ato_tbl` object.
`...`	Passed to the next print method.

Value

Invisibly returns `x`.

Examples

```
x <- data.frame(postcode = "2000", taxable_income = 82000)
x <- structure(x, ato_title = "Demo", ato_source = "https://data.gov.au",
               ato_licence = "CC BY 2.5 AU", ato_retrieved = Sys.time(),
               class = c("ato_tbl", "data.frame"))
print(x)
```

Index