Package: childpen (via r-universe)

October 23, 2026

Title Identification and Estimation of Child Penalties

Version 0.3.0

Description Estimators from Leventer (2025), ``Identification of Child Penalties'' <doi:10.48550/arXiv.2602.07486>: within-gender difference-in-differences, triple differences, the conventional normalized estimator with a bias-bounding exercise, the effect of parenthood on the gender earnings ratio under normalized and proportional triple differences, the accounting of the gender gap, and aggregation across treatment groups, with influence-function standard errors clustered by person.

License MIT + file LICENSE

URL https://github.com/dorleventer/childpen, https://dorleventer.github.io/childpen/

BugReports https://github.com/dorleventer/childpen/issues

Imports data.table, stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0), purrr, scales, dplyr, tidyr, ggplot2

Config/testthat/edition 3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

VignetteBuilder knitr

Depends R (>= 3.5)

Repository https://dorleventer.r-universe.dev

Date/Publication 2026-10-23 19:34:33 UTC

RemoteUrl https://github.com/dorleventer/childpen

RemoteRef HEAD

RemoteSha e06a257ccfc98a36684e0d596fe6af5b11614829

Contents

aggregate_estimands 2
bias_bounds 4
multiple_treatment_group_analysis 5
simulate_data 7
single_treatment_group_analysis 8

Index **11**

`aggregate_estimands` *Aggregate estimates across treatment groups*

Description

Averages the treatment-group-specific estimates from `multiple_treatment_group_analysis` across treatment groups, by event time (Leventer 2025, Section 5). The default weights are the treatment-group shares in the sample. Standard errors account for persons who serve as controls for several treatment groups.

Usage

```
aggregate_estimands(
  results,
  weights = "sample",
  methods = c("DID_Female", "DID_Male", "TD", "Conventional", "NTD", "TD_E_LOGS",
    "TD_PR"),
  include_pre = FALSE
)
```

Arguments

`results`	The data.frame returned by `multiple_treatment_group_analysis`, unmodified (its attributes carry the influence functions and the treatment-group sizes).
`weights`	How to weight treatment groups: • `"sample"` (default): treatment-group shares. Within gender, $w_{g,d} = n_{g,d} / \sum_k n_{g,k}$, for `DID_Female` and `DID_Male`; across genders, $w_d = n_d / \sum_k n_k$, for all other methods. • `"uniform"` or `NULL`: equal weight per treatment group. • A named numeric vector (names = treatment groups): fixed weights, renormalized to sum to one over the groups present at each event time. Groups missing from the vector get weight zero, with a warning.
`methods`	Methods to aggregate (default: all). `DID_Female`, `DID_Male`: `weighted_avg` and `ratio_of_avgs` of `theta`; `Conventional`: `weighted_avg` of `gap_theta`; `NTD` and `TD_E_LOGS`: `weighted_avg` of `rho_effect` and of the accounting rows `rho_obs`, `rho_cf`, `share_parenthood`, `share_other`; `TD`: `weighted_avg` of `gap_ATE`; `TD_PR`: `weighted_avg` of `gap_theta_pre`.

include_pre — Logical (default FALSE). Also aggregate pre-treatment event times. A pre-treatment event time has one cell per control group (control_offset = $d' - d$), and each offset is aggregated separately, so each row is the validation test of one treatment-control distance.

Details

agg_type = "weighted_avg" is $\sum_d w_d B_d$ over the treatment-group estimates B_d. For DID_Female / DID_Male this is $\theta_{Agg,1}(g, e)$; for Conventional it is the conventional aggregate gap $\Delta\theta_{Agg}(e)$, which under NTD is biased for the gap in normalized effects (Theorem 1); for NTD it is $\rho_{Agg}(e)$, the aggregate effect of parenthood on the gender earnings ratio. agg_type = "ratio_of_avgs" (DID rows only) is $\theta_{Agg,2}(g, e)$, the ratio of the weighted average ATE to the weighted average APO; it puts more weight on higher-earning groups, which is why the paper prefers $\theta_{Agg,1}$ (Section 5.1).

Standard errors. The influence function of an aggregate (aggregate estimator section of the appendix on estimation and inference) sums the treatment-group influence functions with the weights; with weights = "sample" it adds the term for the estimated shares, $\psi_{A(e)} = P_e^{-1} \sum_d [(B_d - A(e))\psi_{p_d} + p_d \psi_{B_d}]$, $\psi_{p_d} = 1\{D = d\} - p_d$, where p_d is the share of persons in group d ($1\{G = g, D = d\} - p_{g,d}$ with the joint share of gender g and group d for the within-gender DID rows) and P_e the summed share of the aggregated groups. For ratio_of_avgs the chain rule is applied to the ratio of the two weighted averages, each with its share term. Standard errors are clustered by person. The share term is put on the person-age scale with the average number of ages per person, exact for a balanced panel.

Not every treatment group has an estimate at every event time; each event time uses the groups present, and n_groups reports how many.

Value

A data.frame with one row per event time, estimand, method and aggregation type: event_time, control_offset ($d' - d$; after treatment it is event_time + 1), estimand, method, agg_type, est, se, n_groups (treatment groups with positive weight), ci_l, ci_h (est -/+ 1.96 se), t, p.

See Also

multiple_treatment_group_analysis, bias_bounds

Examples

```
sim <- simulate_data(n_individuals = 2000)
res <- multiple_treatment_group_analysis(sim, treatment_groups = 24:26,
  periods_post = 2, periods_pre = NULL, verbose = FALSE)
agg <- aggregate_estimands(res)
agg[agg$method %in% c("NTD", "Conventional"), ]
```

`bias_bounds` *Bias-corrected conventional estimates for assumed fathers' effects*

Description

The bias-bounding exercise of Leventer (2025, Section 3.3). Under NTD the conventional estimator (the gender gap in normalized DID) is biased for the gender gap in normalized effects $\theta(f) - \theta(m)$ when parallel trends fails (Theorem 1). If fathers' normalized effect $\theta(m)$ were known, the gap would be identified (Proposition 2). This function computes the bias-corrected gap for each assumed value of $\theta(m)$, for every post-treatment cell of `results` (Proposition 2 applies after treatment, $d' > a \geq d$); comparing it with the conventional estimate bounds the bias.

Usage

```
bias_bounds(results, theta_m = c(-0.1, 0.1))
```

Arguments

`results`	The data.frame returned by `multiple_treatment_group_analysis`, unmodified.
`theta_m`	Numeric vector of assumed fathers' normalized effects (default `c(-0.1, 0.1)`, the paper's range).

Details

For a post-treatment cell (d, d', a), with $\hat{\Delta}$ the `NTD rho_effect` estimate (the effect of parenthood on the gender earnings ratio) and $\hat{\delta}_{APO}(g)$ the DID-imputed counterfactual mean of gender g (rows `APO`, `DID_Female` / `DID_Male`), the bias-corrected gap is (Proposition 2, in the form of the appendix on estimation and inference)

$$\hat{\Delta}\,\frac{\hat{\delta}_{APO}(m)}{\hat{\delta}_{APO}(f)}\,(1+\theta(m)).$$

It is linear in $\theta(m)$, so the values at the ends of an interval bound it over the whole interval. Standard errors come from the delta method on the stored person-level influence functions, clustered by person.

Value

A data.frame with one row per cell and value of `theta_m`: `d`, `dp`, `a`, `event_time`, `theta_m`, `est` (the bias-corrected gap), `se`, `ci_l`, `ci_h`.

See Also

`multiple_treatment_group_analysis`

Examples

```
sim <- simulate_data(n_individuals = 2000)
res <- multiple_treatment_group_analysis(sim, treatment_groups = 24:26,
  periods_post = 2, periods_pre = NULL, verbose = FALSE)
bb <- bias_bounds(res)
# Conventional estimate and the bias-corrected band, d = 25
subset(res, method == "Conventional" & d == 25, c(event_time, est))
subset(bb, d == 25, c(event_time, theta_m, est, se))
```

`multiple_treatment_group_analysis`

Child-penalty estimates for many treatment groups and event times

---

Description

Runs `single_treatment_group_analysis` over a grid of 2x2 comparisons. For each treatment group d (age at first birth) and event time $e = 0, \ldots, H$, the target age $a = d + e$ is compared with the closest not-yet-treated group $d' = a + 1$ (Leventer 2025, Section 1.3). Pre-treatment event times are estimated against every control group used after treatment, because the identifying assumptions must hold pair by pair (Section 2.3). Results are stacked, and the person-level influence functions that `aggregate_estimands` and `bias_bounds` need are attached.

Usage

```
multiple_treatment_group_analysis(
  data,
  treatment_groups,
  periods_post,
  periods_pre = 4,
  max_control_group = Inf,
  min_pre_age = 0,
  pre = 1,
  Y_name = "Y",
  age_name = "age",
  D_name = "D",
  id_name = "id",
  female_name = "female",
  verbose = TRUE
)
```

Arguments

`data`	A data.frame or data.table in long format, one row per person and age; see `single_treatment_group_analysis` for the requirements. It is not modified.
`treatment_groups`	Integer vector of ages at first birth d to estimate (the paper uses 24:34). Groups $d + 1, \ldots, d + H + 1$ must be in `data` to serve as controls.

`periods_post` Whole number $H \geq 0$: post-treatment event times $e = 0, \ldots, H$ (the paper uses 5).

`periods_pre` Whole number $K \geq 0$ (default 4), or `NULL`: number of pre-treatment event times for validation tests, $e = -(pre + K), \ldots, -(pre + 1)$ (with the defaults $-5, \ldots, -2$; $e = -pre$ is the baseline). `0` or `NULL` skips them.

`max_control_group` Largest age at first birth allowed as a control group (the paper's D_{max}; default `Inf`, no bound). Cells whose control group is above it are not estimated.

`min_pre_age` Smallest target age for pre-treatment cells (default 0, no bound). Post-treatment cells are not affected.

`pre` Whole number ≥ 1 (default 1): the baseline age is $d - pre$.

`Y_name`, `age_name`, `D_name`, `id_name`, `female_name` Names of the columns holding earnings, age, age at first birth, the person id and the female indicator.

`verbose` Logical (default `TRUE`). Report progress, one line per treatment group.

Value

A data.frame with one block of rows per estimated cell (`d`, `dp`, `a`), sorted by `d`, `event_time` and `dp`; the columns are those of `single_treatment_group_analysis`. Attributes used by `aggregate_estimands` and `bias_bounds`: `"if_data"` (person-level influence functions per cell), `"n_obs"` (rows of the prepared data), `"n_ids"` (persons), `"id_info"` (`id`, `female`, `D`). Attribute `"skipped"` lists cells that could not be estimated, with the reason (e.g. an empty gender x group x age cell). Pass the result to the aggregation functions unmodified.

See Also

`single_treatment_group_analysis`, `aggregate_estimands`, `bias_bounds`

Examples

```
sim <- simulate_data(n_individuals = 2000)
res <- multiple_treatment_group_analysis(sim, treatment_groups = 24:26,
  periods_post = 1, periods_pre = 2, verbose = FALSE)
# The new estimand, by treatment group and event time
subset(res, method == "NTD" & estimand == "rho_effect" & event_time >= 0,
       c(d, event_time, dp, est, se))
# Pre-treatment validation tests of NTD, one row per control group
subset(res, method == "Conventional" & event_time < 0,
       c(d, event_time, control_offset, est, se))
```

`simulate_data` *Simulate panel data for child-penalty estimation*

Description

Generates a balanced panel of annual earnings designed to show the patterns the paper discusses: selection on the timing of first birth that makes parallel trends fail in the paper's direction, effects of parenthood that differ by age at first birth and fade with time since birth, and a gender earnings ratio that falls with age. The true values are attached, so estimates can be compared with them.

Usage

```
simulate_data(n_individuals = 10000, treatment_groups = 24:28, seed = 42)
```

Arguments

`n_individuals`	Whole number ≥ 2. Number of individuals (default 10 000), half of them women (rounded).
`treatment_groups`	Whole numbers ≥ 21: ages at first birth (default `24:28`). Ages run from 20 to `max(treatment_groups) - 1`, so the largest group is never observed treated and serves only as a control.
`seed`	Whole number or `NULL`. RNG seed (default 42). The caller's RNG state is saved and restored on exit, so calling this function does not alter the global random stream; `set.seed()` before the call has no effect unless `seed = NULL`, which draws from the current RNG state.

Details

Earnings of person i with age at first birth D_i at age a are

$$\log Y_{i,a} = h_{D_i}(a) + \kappa(a)F_i + \tau_{F_i}(D_i, a - D_i)\,1\{a \geq D_i\} + \alpha_i + \varepsilon_{i,a},$$

with $F_i = 1$ for women and:

- **Career entry and growth** (selection on timing): $h_D(a) = 9.6 + 0.02(a - 20) + \beta(D)x - 0.0015x^2$, with experience $x = \max(a - s(D), 0)$, entry age $s(D) = 20 + 0.7(D - 24)$ and slope $\beta(D) = 0.09 + 0.008(D - 28)$. Later first births enter the labor market later (more schooling) and then grow faster. Before treatment, distant control groups are not yet working and grow more slowly, so pre-treatment DID estimates grow in magnitude with the distance to the control group and are negative for distant control groups (for the closest ones they are small and can be positive); after treatment, the later control groups grow faster, so parallel trends fails with a negative violation for earlier treatment groups.
- **Gender gap by age**: $\kappa(a) = -0.04 - 0.015\max(a - 25, 0)$, so the female/male ratio absent children is $e^{\kappa(a)}$: 0.96 up to age 25, then falling.
- **Effects of parenthood**: mothers $\tau_f(D, e) = -0.45 + 0.02(D - 24) + 0.04\min(e, 5)$ (smaller for later first births, fading with event time e); fathers $\tau_m = -0.02$. In normalized terms $\theta(g) = e^{\tau_g} - 1$.

- **Noise**: $\alpha_i \sim N(0, 0.20^2)$ (permanent), $\varepsilon_{i,a} \sim N(0, 0.15^2)$ (transitory), independent of everything else.
- **Age at first birth**: drawn from `treatment_groups` with probabilities proportional to a normal density with mean 29 and standard deviation 4, independently of gender.

Under this design DID-PT and TD-PT fail; TD-E-LOGS holds exactly (the gender terms cancel in log means); NTD-PT and TD-PR hold only approximately, because the gender ratio changes with age between the baseline and the target age. The NTD estimator therefore differs slightly from the true effect on the gender earnings ratio (by up to about 0.01 in this design); the attached true values show by how much.

Value

A balanced long panel (data.frame), `n_individuals` x (`max(treatment_groups)` - 20) rows sorted by id and age, with columns `id`, `female` (1 = female), `age`, `D` (age at first birth) and `Y` (earnings, positive). Attribute `"truth"`: the population values for every treatment group `d` and age `a >= d` (`event_time = a - d`): `APO_f`, `APO_m` (mean earnings absent children), `ATE_f`, `ATE_m`, `theta_f`, `theta_m`, `gap_theta`, `rho_cf` ($\rho(\infty)$), `rho_obs` ($\rho(d)$), `rho_effect` ($\rho(d)-\rho(\infty)$), `share_parenthood`. Attribute `"params"`: the parameter values.

See Also

`multiple_treatment_group_analysis`

Examples

```
sim <- simulate_data(n_individuals = 2000)
head(sim)
head(attr(sim, "truth"))
```

`single_treatment_group_analysis`

Child-penalty estimates for one 2x2 comparison

Description

For one treatment group d (age at first birth), control group d' and target age a, estimates the descriptive estimands of Leventer (2025): within-gender DID, the gender gaps under TD, the conventional estimator, the effect of parenthood on the gender earnings ratio (NTD and proportional TD), and the accounting of the gender gap. Returns 19 rows with standard errors clustered by person.

Usage

```
single_treatment_group_analysis(
  data,
  d,
  dp,
  a,
  pre = 1,
  Y_name = "Y",
  age_name = "age",
  D_name = "D",
  id_name = "id",
  female_name = "female"
)
```

Arguments

`data`	A data.frame (or data.table) in long format: one row per person and age, with the person id, age, age at first birth, a female indicator and earnings. The panel need not be balanced. Each person's female indicator and age at first birth must be constant; a missing age at first birth means the person never has a child. Rows with missing earnings are dropped (with a message). Earnings may be zero, but the ratio rows need positive cell means. `data` is not modified.
`d`	Whole number. Treatment group: age at first birth.
`dp`	Whole number. Control group d': an age at first birth that is not yet treated at the target age, `dp > a` (and `dp > d`). The paper uses the closest not-yet-treated group, `dp = a + 1` (Section 1.3), as `multiple_treatment_group_analysis` does.
`a`	Whole number. Target age; event time is $e = a - d$. With `a >= d` the rows are post-birth estimates; with `a < d - pre` they are pre-treatment validation (placebo) estimates (Sections 2.3, 4.2).
`pre`	Whole number ≥ 1 (default 1). The baseline age is $d - pre$. The default, $d - 1$, is the paper's choice; larger values move the baseline earlier in case earnings at $d - 1$ reflect anticipation (Section 3.1).
`Y_name`, `age_name`, `D_name`, `id_name`, `female_name`	Names of the columns holding earnings, age, age at first birth, the person id (clustering unit) and the female indicator (`1 / TRUE =` female; also accepted: labels such as "female"/"male", "f"/"m", "woman"/"man"; any other label is an error).

Details

Let $\mu(g, k, t)$ be the mean earnings of gender $g \in \{f, m\}$ with age at first birth k at age t, estimated by the cell mean, and $b = d - pre$ the baseline age. The additive (DID) and multiplicative imputations of counterfactual earnings at age a are

$$\delta_{APO}(g) = \mu(g, d, b) + \mu(g, d', a) - \mu(g, d', b), \qquad \delta^{\times}_{APO}(g) = \mu(g, d, b)\,\mu(g, d', a)/\mu(g, d', b).$$

The rows, with the assumption under which each identifies its target (all also need no anticipation):

method	estimand	estimator; target
DID_Female, DID_Male	APO	$\delta_{APO}(g)$; $APO(g,\infty)$ under DID-PT
DID_Female, DID_Male	ATE	$\mu(g,d,a)-\delta_{APO}(g)$; $ATE(g)$ under DID-PT
DID_Female, DID_Male	theta	ATE row / APO row; $\theta(g)$ under DID-PT
TD	gap_ATE	ATE row f minus m; $ATE(f)-ATE(m)$ under TD-PT
Conventional	gap_theta	theta row f minus m; $\theta(f)-\theta(m)$ under DID-PT only (under NTD-PT it is b
NTD	rho_effect	$\mu(f,d,a)/\mu(m,d,a)-\delta_{APO}(f)/\delta_{APO}(m)$; $\rho(d)-\rho(\infty)$ under NTD-PT (
TD_E_LOGS	rho_effect	as NTD with $\delta^{\times}_{APO}$; $\rho(d)-\rho(\infty)$ under TD-E-LOGS (proposition in the onli
TD_PR	gap_theta_pre	$[\mu(g,d,a)-\delta^{\times}_{APO}(g)]/\mu(g,d,b)$, f minus m; $ATE(g)/APO(g,d,\infty,b)$, f n
NTD, TD_E_LOGS	rho_obs	$\mu(f,d,a)/\mu(m,d,a)$; the observed gender earnings ratio $\rho(d)$
NTD, TD_E_LOGS	rho_cf	$\delta_{APO}(f)/\delta_{APO}(m)$ (NTD) or with $\delta^{\times}_{APO}$ (TD_E_LOGS); $\rho(\infty)$ under NTD
NTD, TD_E_LOGS	share_parenthood	$(\rho_{cf}-\rho_{obs})/(1-\rho_{obs})$; parenthood's share of the observed gender gap (Sec
NTD, TD_E_LOGS	share_other	$1-$ share_parenthood; the share of other factors

Gaps are female minus male. The conventional estimator is the 2x2 version of the normalized event-study child penalty; under NTD it equals the target gap times a gender-invariant bias factor (Theorem 1), see `bias_bounds` for the paper's bias-bounding exercise. The accounting rows and their standard errors are an extension of the paper's Section 4.2 accounting exercise.

Standard errors. Each row is a smooth function of the eight cell means; its influence function follows by the chain rule from those of the cell means, $\psi_\mu = S(Y-\mu)/P(S)$ with S the cell indicator (the appendix on estimation and inference). The variance is $n^{-2}\sum_c(\sum_{i\in c}\psi_i)^2\, G/(G-1)$, clustering by person ($n$ person-age rows, G persons). `vignette("estimators")` writes out every estimator and influence function.

Value

A data.frame with 19 rows and columns `d`, `dp`, `a`, `event_time` ($a-d$), `control_offset` ($d'-d$), `estimand`, `method`, `est`, `se`, `ci_l`, `ci_h` (est -/+ 1.96 se), `t`, `p` (two-sided, normal), and `n_female_treat`, `n_female_control`, `n_male_treat`, `n_male_control` (persons with $D=d$ and $D=d'$). Attribute `"if_scores"`: person-level sums of the influence functions, used by `aggregate_estimands` and `bias_bounds`.

See Also

`multiple_treatment_group_analysis` for the grid of comparisons the paper uses, `aggregate_estimands`, `bias_bounds`, `vignette("estimators")`

Examples

```
sim <- simulate_data(n_individuals = 2000)
# Treatment group d = 25 one year after first birth (a = 26),
# closest not-yet-treated control group d' = a + 1 = 27
res <- single_treatment_group_analysis(sim, d = 25, dp = 27, a = 26)
res[res$method %in% c("Conventional", "NTD"), c("method", "estimand", "est", "se")]
# Pre-treatment validation (placebo) test: a < d - 1
val <- single_treatment_group_analysis(sim, d = 25, dp = 27, a = 22)
val[val$method %in% c("Conventional", "NTD") & val$estimand %in% c("gap_theta", "rho_effect"),
    c("method", "estimand", "est", "se")]
```

Index

aggregate_estimands, 2, *5, 6*, *10*

bias_bounds, *3*, 4, *5, 6*, *10*

multiple_treatment_group_analysis, *2–4*, *5*, *8–10*

simulate_data, 7
single_treatment_group_analysis, *5, 6*, 8