# Cheat sheet tables for `fixest`

June 16, 2026

This document contains the following tables:



Table 1: Cheat sheet for including fixed-effects in a **`fixest`** estimation.

| Formula | Comment | Equivalent to |
|---|---|---|
| `y ~ x` | regular estimation | `y ~ x` |
| `y ~ x \| fe` | estimation with fixed-effect `fe` | `y ~ x + as.factor(fe)` |
| `y ~ 1 \| fe` | estimation with only the fixed-effect `fe` | `y ~ as.factor(fe)` |
| `y ~ 1 \| fe1 ^ fe2` | the fixed-effects `fe1` and `fe2` are combined | `y ~ as.factor(fe1):as.factor(fe2)` |
| `y ~ x1 \| fe[x2]` | estimation with fixed-effect `fe` and variable with varying slope `x2` | `y ~ x1 + as.factor(fe) * x2` |
| `y ~ x1 \| fe1[x2, x3] + fe2` | estimation with fixed-effects `fe1` and `fe2`, and variables with varying slope `x2` and `x3` | `y ~ x1 + as.factor(fe1)*(x2 + x3) + as.factor(fe2)` |

Table 2: Cheat sheet for performing 2SLS/IV estimations in `feols`.

| id | Formula | Comment |
|---|---|---|
| (1) | `y ~ x \| endo ~ inst` | one exogenous variable `x`, one endogenous variable `endo`, instrumented with one variable `inst` |
| (2) | `y ~ x \| endo ~ inst1 + inst2` | same as (1), with two instruments `inst1` and `inst2` |
| (3) | `y ~ x \| endo1 + endo2 ~ inst1 + inst2` | same as (2), with two endogenous variables `endo1` and `endo2` |
| (4) | `y ~ 1 \| endo ~ inst` | same as (1), but without any exogenous regressor |
| (5) | `y ~ x \| fe \| endo ~ inst` | same as (1), with (exogenous) fixed-effects `fe` |
| (6) | `y ~ 1 \| fe \| endo ~ inst` | same as (4), the exogenous regressors consist of only the fixed-effects `fe` |

Table 3: Cheat sheet for multiple estimation in **fixest**.

| Formula | Estimates | Comment |
|---|---|---|
| `c(y1, y2) ~x` | • `y1 ~x`<br>• `y2 ~x` | multiple left hand sides |
| `y ~sw(x1, x2)` | • `y ~x1`<br>• `y ~x2` | **S**tep**W**ise |
| `y ~x \| sw0(fe)` | • `y ~x`<br>• `y ~x \| fe` | stepwise, starting with the empty element |
| `y ~x1 + csw0(x2) \| csw(fe1, fe2)` | • `y ~x1 \| fe1`<br>• `y ~x1 \| fe1 + fe2`<br>• `y ~x1 + x2 \| fe1`<br>• `y ~x1 + x2 \| fe1 + fe2` | **C**umulative **S**tep**W**ise |
| `y ~mvsw(x1, x2, x3)` | • `y ~1`<br>• `y ~x1`<br>• `y ~x2`<br>• `y ~x3`<br>• `y ~x1 + x2`<br>• `y ~x1 + x3`<br>• `y ~x2 + x3`<br>• `y ~x1 + x2 + x3` | **M**ulti **V**erse **S**tep**W**ise: includes all the combinations of variables |

Table 4: Cheat sheet for variable interpolation in `fixest` estimations.

| Preceding code | Formula | Interpolates as |
|---|---|---|
| *all examples with the* `airquality` *data set, with variables:* `Ozone`, `Solar.R`, `Wind`, `Temp`, `Month`, `Day` | | |
| INTERPOLATION WITH `.[]` | | |
| `x = c("Temp", "Month")` | `Ozone ~ .[x]` | `Ozone ~ Temp + Month` |
| `x = ~ Temp + Month` | `Ozone ~ .[x]` | `Ozone ~ Temp + Month` |
| `x = character(0)` | `Ozone ~ .[x]` | `Ozone ~ 1` |
| `x = ~ Temp \| Month` | `Ozone ~ .[x]` | `Ozone ~ Temp \| Month` |
| `y = c("Ozone", "Wind")` | `.[y] ~ Temp + Month` | `c(Ozone, Wind) ~ Temp + Month` |
| *note:* passing several variables in the left-hand-side nests them in `c()` to trigger multiple estimations | | |
| `x = c("Temp", "Month")` | `Ozone ~ sw(.[, x])` | `Ozone ~ sw(Temp, Month)` |
| `n = 1:3` | `y ~ x.[n]` | `y ~ x1 + x2 + x3` |
| MACRO VARIABLES STARTING WITH `..` | | |
| `setFixest_fml(`<br>`  ..x = ~ Temp + Month`<br>`)` | `Ozone ~ ..x` | `Ozone ~ Temp + Month` |
| `setFixest_fml(`<br>`  ..y = ~ c(Ozone, Wind)`<br>`)` | `..y ~ Temp + Month` | `c(Ozone, Wind) ~ Temp + Month` |
| INTERPOLATIONS TIED TO A DATA SET | | |
| *using* `..("reg.-expr.")` *or* `regex("reg.-expr.")` | | |
| | `Ozone ~ ..("Te\|Mo")` | `Ozone ~ Temp + Month` |
| | `Ozone ~ regex("Te\|Mo")` | `Ozone ~ Temp + Month` |
| *using name completion with* `name..` | | |
| | `Ozone ~ T.. + M..` | `Ozone ~ Temp + Month` |

Table 5: Description of `fixest` built-in VCOVs, and how to write the `vcov` argument included in many functions of this package.

| Keyword | Requirement | Provide requirement with `vcov = function` or `formula` |
|---|---|---|
| *Spherical errors* | | |
| `"iid"` | *No requirement* | |
| *Heteroskedastic errors* | | |
| `"hetero"`, `"hc1"`, `"hc2"` `"hc3"` | *No requirement* | |
| *Clustered, within group correlation* | | |
| `"cluster"` | req. 1: group identifier<br>*default (if available, in that order):*<br>• the unit identifier from the panel<br>• the first fixed-effect | `vcov_cluster("id")`<br>`cluster ~ id`<br>`~ id` (i.e., no formula lhs) |
| *Clustered, within groups correlations* | | |
| `"twoway"` | req. 1: two group identifiers<br>*default (if available, in that order):*<br>• the unit and time identifiers from the panel<br>• the first two fixed-effects | `vcov_cluster(c("id1", "id2"))`<br>`cluster ~ id1 + id2`<br>`~ id1 + id2` (i.e., no formula lhs) |
| *Newey West, serial correlation* | | |
| `"NW"` | req. 1: unit and time identifiers<br>*default (if available):*<br>• the panel unit and time identifiers<br>req. 2: lag<br>*default:*<br>• default lag provided by `sandwich::bwNeweyWest` | `vcov_NW(unit = "id", time = "year")`<br>`NW ~ id + year`<br><br>`vcov_NW(lag = 2, "id", "year")`<br>`NW(lag = 2) ~ id + year` |
| *Driscoll Kraay, serial correlation* | | |
| `"DK"` | req. 1: time identifier<br>*default (if available):*<br>• the time identifier from the panel<br>req. 2: lag<br>*default:*<br>• default lag equal to $N_T^{0.25}$ | `vcov_DK(time = "year")`<br>`DK ~ year`<br><br>`vcov_DK("year", lag = 4)`<br>`DK(lag = 4) ~ year` |
| *Conley, spatial corelation* | | |
| `"conley"` | req. 1: latitude and longitude<br>*default (if available):*<br>• lat. and long. variables from the data set<br>req. 2: distance cutoff<br>*default:*<br>• the default cutoff is based on an internal algorithm using the median distance across a sample of units | `vcov_conley("lat", "lng")`<br>`conley ~ lat + lng`<br><br>`vcov_conley("lat", "lng", "100mi")`<br>`conley(cutoff = "100mi") ~ lat + lng` |