02_FixedEffects

QuantFit Estimator Standard Operating Procedure

SOP: Fixed Effects (Within Estimator)

Panel data estimator that controls for time-invariant unobserved heterogeneity

=> Use FE when entity-specific time-invariant factors might be correlated with the regressors.

1. Purpose

FE eliminates time-invariant entity heterogeneity by within-transformation: subtracting the entity mean from each variable. Coefficients are identified from variation over time within each entity, not across entities.

2. When to use this estimator

Panel data with N entities x T periods, where T >= 3 ideally.

Suspected correlation between unobserved entity effects and regressors (Hausman test rejects RE).

Interest in within-entity dynamics rather than between-entity differences.

Country, firm, household, or region panels with sticky structural attributes.

3. Required data structure

A column identifying the entity (group ID) - this drives the within-transformation.

Optional time identifier for two-way FE.

Numeric Y and X with sufficient within-entity variation (low ICC => FE may be inefficient).

4. Mathematical formulation

The within-transformation removes entity means alpha_i, leaving the slope beta identified:

Y_it - ?_i = (X_it - X?_i)' beta + (epsilon_it - epsilon?_i)

beta?_FE = (SumSum (X_it - X?_i)(X_it - X?_i)')^-1 SumSum (X_it - X?_i)(Y_it - ?_i)

alpha?_i = ?_i - X?_i' beta?_FE

sigma?2 = epsilon?'epsilon? / (NT - N - k)

5. Pre-estimation diagnostics

Confirm the panel is balanced or near-balanced; flag unequal T per entity.

Stationarity per variable (CIPS / Fisher-ADF) since panel pooling assumes stationarity.

Cross-sectional dependence (Pesaran CD) - if rejected, use Driscoll-Kraay SE

or CCE/CS-ARDL.

Inspect within R2 vs between R2 to gauge how much variation is within-entity.

6. Estimation procedure

Demean Y and each X by entity: Y_it - ?_i, X_it - X?_i.

Apply OLS to the demeaned data.

Recover entity intercepts alpha?_i = ?_i - X?_i' beta?.

Compute residuals epsilon?_it = Y_it - ?_it; standard errors clustered by entity by default.

Run pooled-vs-FE F-test (H0: all alpha_i equal) to confirm FE is needed.

7. Output produced

8. Output interpretation

beta?_j is the change in Y_it per unit change in X_jit holding entity-specific factors constant.

Significant pooled-vs-FE F-test => entity heterogeneity is real; pooled OLS would be biased.

Compare to RE via Hausman: H_p < 0.05 => FE preferred.

Entity effects alpha?_i can be plotted as a heat-map to spot structural outliers.

9. Post-estimation diagnostics

Cluster-robust SE by entity is the default; verify clustering choice matches data structure.

Heteroskedasticity (modified Wald for groupwise) - if rejected, retain clustered SE.

Serial correlation in residuals (Wooldridge AR(1) test) - if rejected, use HAC or AR(1) FE.

Cross-sectional dependence (Pesaran CD on residuals) - if rejected, prefer CCE / CS-ARDL.

10. Common pitfalls

FE wipes out time-invariant regressors - coefficients on gender, ethnicity, etc. are not identified.

If T is small relative to N, FE is consistent but inefficient; consider RE if Hausman accepts it.

Panel cointegration is not addressed by FE - non-stationarity can still produce spurious results.

FE on non-balanced panels with selection on time can introduce attrition bias.

11. Reporting checklist

Number of entities N, average T, total observations.

Within R2, between R2, overall R2.

Pooled-vs-FE F-test and Hausman test (FE vs RE).

Standard error type clearly named (clustered, Driscoll-Kraay, etc.).

Pesaran CD and slope-homogeneity diagnostics.

12. References

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data, 2nd ed. MIT.

Baltagi, B. H. (2021). Econometric Analysis of Panel Data, 6th ed. Springer.

Hsiao, C. (2014). Analysis of Panel Data, 3rd ed. Cambridge.

Field | Meaning

coefficients | Within-estimator slopes beta?_j

entityEffects | Per-entity intercept alpha?_i

rSquared | Within R2 (variation explained net of entity means)

metadata['betweenR2'] | Between-entity R2 for diagnostics

metadata['overallR2'] | Pooled R2 for completeness

metadata['pooledFTest'] | F-stat / p for pooled-vs-FE test

residuals / fitted | Within-transformed residuals