

# Impact of ERDF on Regional Growth and Convergence in EU NUTS2 Regions (2014–2020): Panel Evidence and Policy-Rule Identification

EUcohesion Project

February 8, 2026

## Abstract

This report studies the causal effect of ERDF funding intensity on regional GDP per-capita growth and convergence across EU NUTS2 regions. We use the project's harmonized panel and policy-rule artifacts already generated in V3.1, combining TWFE benchmarks with policy-rule-based designs (RD and IV). The main outcome is real GDP per-capita growth; PPS growth is reported as robustness. Diagnostics show that the 75% eligibility cutoff does not generate a strong funding discontinuity (first-stage F-statistics below 2 across bandwidths), so fuzzy RD is not treated as a causal ERDF LATE design. We therefore use diagnostics-driven IV selection among candidate instruments and select a cross-sectional cumulative-exposure specification instrumented by eligibility below 90%, which delivers the strongest first stage ( $F=15.65$ ). The headline IV estimates are negative and marginally imprecise (p-values around 0.07–0.09), while TWFE benchmarks remain small and specification-sensitive. Convergence evidence remains strong: both sigma and beta metrics support convergence over the sample. Results should be interpreted with caution because identification still depends on exclusion assumptions and partial eligibility coverage.

## 1 Introduction

**Problematic.** This report targets the causal effect of ERDF intensity on GDP per-capita growth and convergence, with policy period focus on 2014–2020 and post-period outcomes.

The empirical challenge is that regional funding intensity is endogenous to initial conditions and policy targeting. V3.1 therefore combines descriptive panel evidence with explicit design diagnostics for policy-rule identification, rather than presenting one estimator without testing its viability.

## 2 Data

### 2.1 Sources and identifiers

Table 1 lists the exact datasets used in the repository build and analysis pipeline, with IDs and URLs. Primary source portals are DG REGIO and Eurostat [[European Commission, DG REGIO, 2026, Eurostat, 2026a,b](#)], with ERDF finance records from the Cohesion Data Platform [[European Commission, Cohesion Data Platform, 2026a,b](#)].

Table 1: Data sources used in V3.1

Domain	Dataset ID	URL
Cohesion funding	99js-gm52	<a href="#">link</a>
Cohesion funding	tc55-7ysv	<a href="#">link</a>
Eurostat GDP (PPS)	nama_10r_2gdp	<a href="#">link</a>
Eurostat regional accounts	nama_10r_2gvagr	<a href="#">link</a>
Eurostat population	demo_r_d2jan	<a href="#">link</a>
Eurostat unemployment	tgs00010	<a href="#">link</a>
Eurostat employment	tgs00007	<a href="#">link</a>
Eurostat tertiary education	tgs00109	<a href="#">link</a>
Eurostat R&D	tgs00042	<a href="#">link</a>
DG REGIO hub	Data for research	<a href="#">link</a>

## 2.2 Construction summary

The processed panel is built at the NUTS2-year level with harmonized keys (`nuts2_id`, `year`), ERDF per-capita treatment and lags, GDP outcomes, controls, and eligibility metadata. NUTS harmonization is imperfect across historical revisions, so eligibility/running-variable coverage is partial relative to the full panel.

Real GDP per capita is reconstructed from regional volume indices (`nama_10r_2gvagr`, B1GQ, I15) anchored to each region's nominal reference level (preferably 2015):

$$GDPpc_{it}^{real} = GDPpc_{i,ref}^{nom} \times \frac{I_{it}^{vol}}{I_{i,ref}^{vol}}.$$

Panel coverage metrics from the generated overview are reported in Table 2.

Table 2: Panel overview metrics

metric	value
Min year	1986
Max year	2025
Regions	398
Rows	15920

## 3 Empirical strategy

### 3.1 TWFE benchmark

The benchmark panel models are:

$$(A) \quad y_{it} = \beta ERDFpc_{i,t-1} + \gamma' X_{it} + \alpha_i + \tau_t + \varepsilon_{it}, \quad (1)$$

$$(B) \quad y_{it} = \beta ERDFpc_{i,t-1} + \gamma' X_{it} + \alpha_i + \delta_{c(i),t} + \varepsilon_{it}, \quad (2)$$

$$(C) \quad y_{it} = \sum_{k=1}^3 \beta_k ERDFpc_{i,t-k} + \gamma' X_{it} + \alpha_i + \tau_t + \varepsilon_{it}, \quad (3)$$

where  $y_{it}$  is headline real growth (with PPS robustness),  $\alpha_i$  are region FE, and errors are clustered by region.

### 3.2 RD diagnostics and sharp RD ITT

Using the running variable

$$r_i = \left( \frac{GDPpc_i^{PPS}}{EU \text{ average } GDPpc^{PPS}} \right) \times 100,$$

we test whether funding jumps at the 75 threshold. Local linear RD with triangular kernels is estimated over a bandwidth grid  $\{5, 7.5, 10, 12.5, 15, 20\}$ .

Because first-stage funding jumps are weak in V3.1, fuzzy RD is treated as non-viable for causal ERDF LATE. Sharp RD is retained as reduced-form eligibility ITT evidence.

### 3.3 IV candidate set and selection

IV candidates are diagnosed via first-stage strength (F-stat, partial  $R^2$ ) before second-stage estimation. Candidate instruments include interactions of eligibility indicators with time exposure and cross-sectional threshold instruments for cumulative ERDF exposure.

Headline IV is selected by the documented rule: choose the strongest candidate with first-stage  $F \geq 10$  for the headline outcome window; otherwise revert to conservative TWFE interpretation.

## 4 Results

### 4.1 TWFE benchmark estimates

Table 3 reports Models A and B for real growth; Table 4 reports distributed lags. Figure 1 visualizes dynamic lag estimates.

Table 3: TWFE benchmark (headline outcome: real growth)

Model	Coef	SE	p	N obs	N regions	SE cluster
Model A	0.0057	0.0030	0.0604	1175	223	nuts2
Model B	0.0034	0.0029	0.2444	1175	223	nuts2

Table 4: Distributed lags (Model C, headline outcome)

Term	Coef	SE	p	N obs
ERDF pc lag 1	0.0120	0.0048	0.0134	861
ERDF pc lag 2	0.0079	0.0038	0.0366	861
ERDF pc lag 3	-0.0031	0.0027	0.2470	861

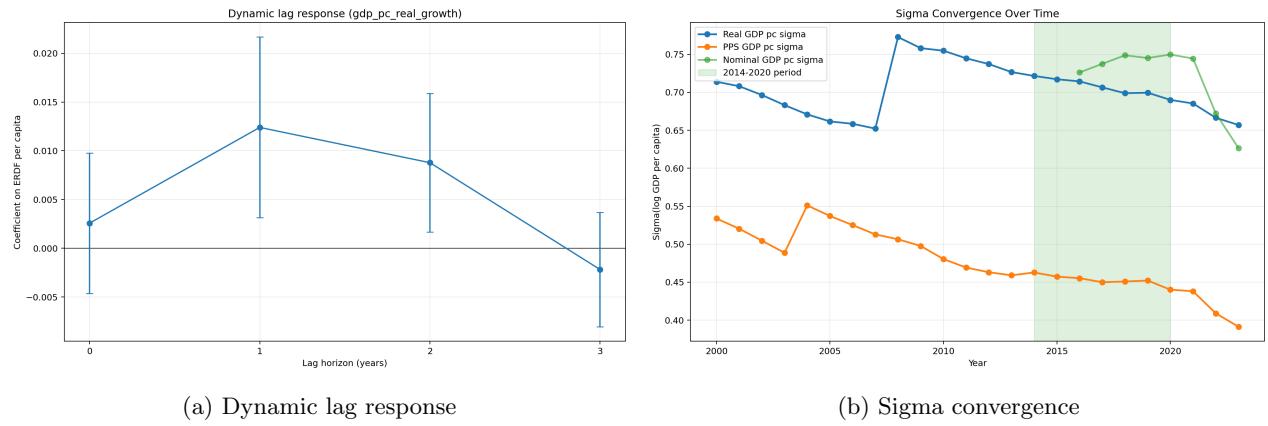


Figure 1: Dynamic treatment profile and convergence trend

### 4.2 RD diagnostics and sharp RD

Table 5 and Figure 2 show weak funding discontinuity at 75 (F-statistics below conventional strength thresholds). Table 6 reports sharp-RD ITT estimates; Table 7 reports placebo pre-trend.

Table 5: RD first-stage funding jump diagnostics at 75 cutoff

BW	Jump	SE	p	F	Fuzzy viable
5.000	605.066	582.900	0.299	1.077	False
7.500	628.144	482.716	0.193	1.693	False
10.000	521.676	405.835	0.199	1.652	False
12.500	406.564	356.718	0.254	1.299	False
15.000	323.631	308.232	0.294	1.102	False
20.000	277.562	251.487	0.270	1.218	False

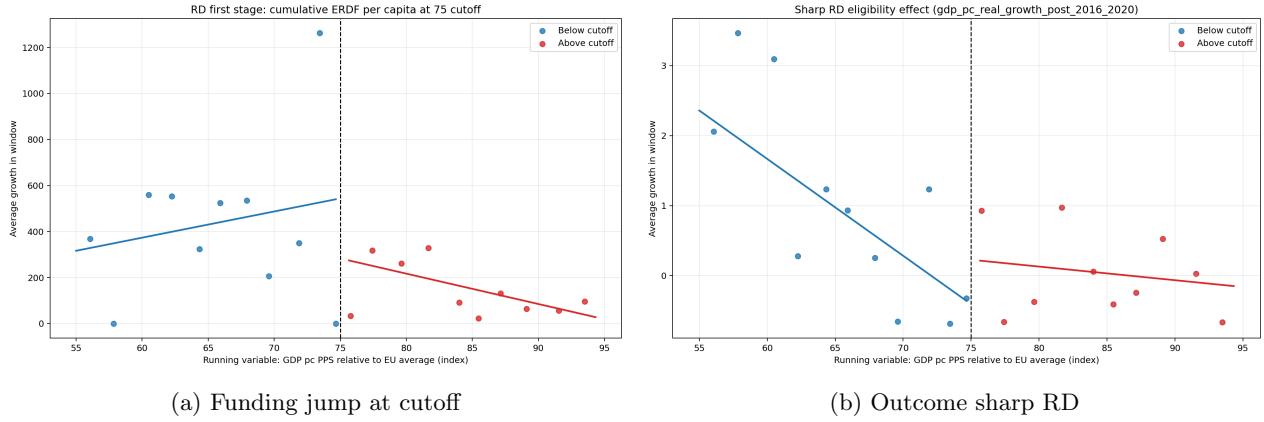


Figure 2: RD diagnostics and reduced-form eligibility discontinuity

Table 6: Sharp RD outcome discontinuities (eligibility ITT)

Outcome	Window	BW	Coef	SE	p	N
PPS growth	post_2016_2020	10.000	-1.971	1.536	0.200	53
PPS growth	post_2021_2023	10.000	2.840	1.202	0.018	53
Real growth	post_2016_2020	10.000	-0.159	1.375	0.908	53
Real growth	post_2021_2023	10.000	2.420	0.975	0.013	53

Table 7: RD placebo pre-trend test

Window	BW	Coef	SE	p	N
pre_2010_2013	10.000	-1.942	2.456	0.429	53

### 4.3 IV diagnostics and headline IV

Table 8 summarizes first-stage diagnostics across candidates. The selected headline design is cross-sectional 2SLS using cumulative exposure instrumented by `eligible_lt90`. Table 9 reports those estimates; panel IV is reported in Table 10.

Table 8: IV first-stage candidate diagnostics

ID	Sample	Endogenous	Instrument	F	Partial R2	F (headline)	Status
Z1	panel	erdf.eur.pc.ll1	z1_post75				failed
Z4_75	cross_section	erdf.eur.pc.cum.2014..2020	eligible_lt75	10.71	0.04	4.60	ok
Z4_75	cross_section	erdf.eur.pc.cum.2015..2020	eligible_lt75	10.71	0.04	4.60	ok
Z4_90	cross_section	erdf.eur.pc.cum.2014..2020	eligible_lt90	9.61	0.04	15.65	ok
Z4_90	cross_section	erdf.eur.pc.cum.2015..2020	eligible_lt90	9.61	0.04	15.65	ok
Z5	panel	erdf.eur.pc.ll1	z5_eu90	1.44	0.00		ok
Z2	panel	erdf.eur.pc.ll1	z2_eu75	1.23	0.00		ok
Z3	panel	erdf.eur.pc.ll1	z3_country75	1.12	0.00		ok

Table 9: Cross-sectional 2SLS (headline IV family)

Outcome	Window	Instrument	Coef	SE	p	F	N
Real growth	post..2016..2020	eligible_lt90	-0.004	0.002	0.088	15.654	265
Real growth	post..2021..2023	eligible_lt90	-0.005	0.003	0.069	15.654	265
PPS growth	post..2016..2020	eligible_lt90	-0.002	0.002	0.140	15.654	265
PPS growth	post..2021..2023	eligible_lt90	-0.001	0.002	0.651	15.654	265

Table 10: Panel 2SLS (selected panel candidate)

Outcome	Instrument	Coef	SE	p	F	N
Real growth	z5_eu90	-0.087	0.095	0.360	1.810	1175

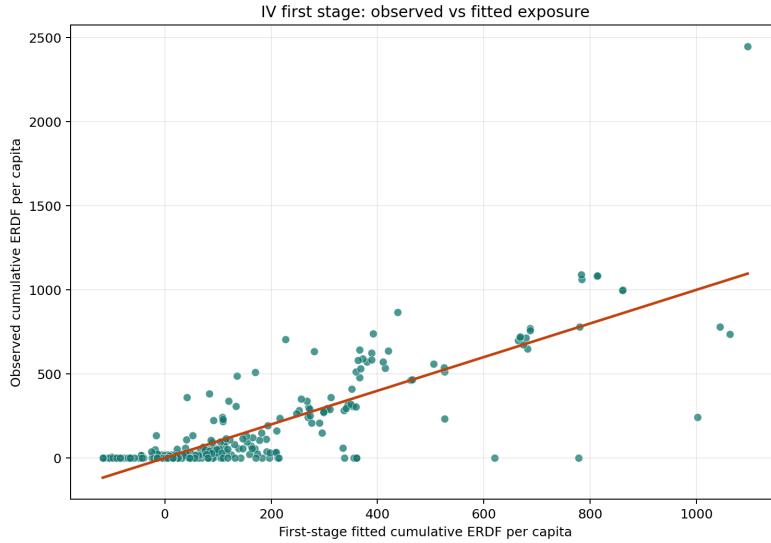


Figure 3: IV first stage: fitted vs observed cumulative exposure

#### 4.4 Comparison and convergence

Table 11 compiles the main estimator families. Table 12 summarizes beta convergence evidence.

Table 11: Model comparison summary (V3.1)

Family	Model	Window	Coef	SE	p	F	Headline
IV (cross-section)	IV 2SLS cross-section (selected candidate)	post_2016_2020	-0.004	0.002	0.088	15.654	True
IV (panel)	IV 2SLS panel (selected candidate)	panel	-0.087	0.095	0.360	1.810	False
RD (sharp ITT)	Sharp RD bw=10.0	post_2016_2020	-0.159	1.375	0.908		False
TWFE	Model A	panel	0.006	0.003	0.060		False
TWFE	Model B	panel	0.003	0.003	0.244		False
TWFE-dynamic	Model C (l1 term)	panel	0.012	0.005	0.013		False

Table 12: Beta convergence (Model D)

Outcome	Term	Coef	SE	p
Real growth	log real GDP pc lag	-5.267	0.602	0.000
PPS growth	log PPS GDP pc lag	-7.725	0.436	0.000

## 5 Discussion and limitations

- **Weak fuzzy RD:** at the 75 cutoff, first-stage funding-jump F-statistics remain below 2 across bandwidths, so fuzzy RD is not credible for ERDF LATE.
- **IV assumptions:** the headline cross-sectional IV depends on exclusion and monotonicity assumptions tied to policy thresholds; estimates should be interpreted as design-based but assumption-dependent.
- **Coverage mismatch:** eligibility/running-variable mapping does not cover all region-year observations in the full panel because of NUTS revision overlap.
- **Outcome construction:** real GDP per capita is reconstructed from volume indices anchored to nominal levels, not direct chain-linked euro series.

## 6 Conclusion

With the currently available policy structure and data, V3.1 supports a cautious interpretation: convergence patterns are robustly negative in beta terms and declining in sigma terms, while direct ERDF growth effects remain sensitive to estimator choice. RD funding-jump diagnostics indicate weak support for fuzzy RD identification at the 75 threshold. The strongest available design in this pipeline is cross-sectional IV with cumulative exposure instrumented by `eligible_lt90`, which achieves a stronger first stage but still yields imprecise headline effects. The defensible claim is therefore limited: evidence is consistent with non-zero policy impacts in some specifications, but large causal claims are not yet supported without stronger quasi-experimental variation.

## References

- European Commission, Cohesion Data Platform. Finance implementation details (2014–2020). <https://cohesiondata.ec.europa.eu/2014-2020-Finances/Finance-Implementation-Details/99js-gm52>, 2026a. Dataset ID: 99js-gm52; Accessed 2026-02-08.
- European Commission, Cohesion Data Platform. Historic eu payments regionalised and modelled time series. <https://cohesiondata.ec.europa.eu/Historic-Historic-EU-payments-regionalised-and-modelled-/tc55-7ysv>, 2026b. Dataset ID: tc55-7ysv; Accessed 2026-02-08.
- European Commission, DG REGIO. Data for research and analyses. [https://ec.europa.eu/regional-policy/information-sources/maps/data-for-research\\_en](https://ec.europa.eu/regional-policy/information-sources/maps/data-for-research_en), 2026. Accessed 2026-02-08.

Eurostat. Gross domestic product (gdp) at current market prices by nuts 2 regions. [https://ec.europa.eu/eurostat/databrowser/view/nama\\_10r\\_2gdp/default/table](https://ec.europa.eu/eurostat/databrowser/view/nama_10r_2gdp/default/table), 2026a. Dataset ID: nama\_10r\_2gdp; Accessed 2026-02-08.

Eurostat. Gross value added by nuts 2 regions. [https://ec.europa.eu/eurostat/databrowser/view/nama\\_10r\\_2gvagr/default/table](https://ec.europa.eu/eurostat/databrowser/view/nama_10r_2gvagr/default/table), 2026b. Dataset ID: nama\_10r\_2gvagr; Accessed 2026-02-08.