

METHODOLOGY & DIAGNOSTICS APPENDIX

Gravity Model of CEE-5 Export Dependence on Germany and France

Data, Variables, Model Specifications, Diagnostic Tests, and Robustness Checks

Purpose and audience of this document

This document is the technical companion to the main report. It explains in detail the data sources used, how each variable was constructed, the logic behind each statistical model, and the full set of tests used to check the reliability of the results. It is written for readers with a quantitative background who want to verify the methodology. A Glossary at the end of this introduction defines every technical term used.

GLOSSARY OF TECHNICAL TERMS

The following table defines every technical term used in this document and in the main report, in alphabetical order. Readers unfamiliar with econometrics are encouraged to consult this glossary when they encounter unfamiliar terms.

Table 1. Glossary of key terms used in this analysis.

Term	Plain-language definition
ADF test	Augmented Dickey-Fuller test. A statistical test that asks whether a data series tends to drift away from a starting point over time (non-stationary / has a "unit root") or fluctuates around a stable average (stationary). If a series is non-stationary, simple regression can give misleading results.
Bilateral pair	One specific exporter–importer combination. This study has ten pairs: each of the five CEE countries paired with Germany and with France.
Cook's D	Cook's Distance. A measure of how much a single data point influences the regression results. Large values flag observations that, if removed, would substantially change the estimated coefficients.
DVA	Domestic Value Added. The portion of export revenue that stays in the exporting country as wages, profits, and taxes, after subtracting the cost of imported inputs used in production.
Elasticity	The percentage change in one variable caused by a 1% change in another. A supply elasticity of 0.88 means: when GDP grows by 1%, DVA exports grow by 0.88%.
FIGARO	Full International and Global Accounts for Research in Input-Output analysis. Eurostat's framework for decomposing bilateral trade into domestic and foreign value-added components.
Fixed effects	Statistical controls that absorb all differences between countries (or country pairs) that do not change over time — such as language, culture, and historical trade links — without explicitly measuring them.
GVC / Value chain	Global Value Chain. A production process spread across multiple countries, where each country specialises in one stage (e.g., mining, components, assembly, distribution).
HC3 standard errors	Heteroscedasticity-consistent standard errors (HC3 variant). A correction applied when the statistical noise in the data is not constant across observations. Makes hypothesis tests valid even when the "equal variance" assumption is violated.
I(1) / I(0)	Integration order. I(1) means the series has a unit root in levels but becomes stationary after taking first differences (year-on-year changes). I(0) means stationary in levels.
Multilateral resistance	The idea that bilateral trade between two countries depends not only on their mutual trade barriers but also on each country's average trade barriers with the rest of the world. Countries that face high barriers everywhere trade more with each other.
OLS	Ordinary Least Squares. The standard statistical method for fitting a line through data to minimise the sum of squared errors between observed and predicted values.

Term	Plain-language definition
Panel data	Data that follows the same units (here: bilateral country pairs) over multiple time periods. Combining cross-sectional and time-series variation improves statistical precision.
Pesaran CD	Pesaran Cross-sectional Dependence statistic. Tests whether the statistical errors across different country pairs move together — which they often do when all countries are hit by the same global shocks.
PPML	Poisson Pseudo-Maximum Likelihood. An estimator that models trade flows in levels rather than logarithms. It is preferred when the data are heteroscedastic (unequal variance) because it gives unbiased estimates where log-OLS would not.
Pseudo-R²	A goodness-of-fit measure for non-linear models (like PPML) that approximates the share of variation explained. Not directly comparable to the standard R ² from OLS.
R²	R-squared. The share of variation in the dependent variable (DVA exports) explained by the model. R ² = 0.93 means the model accounts for 93% of the variation.
RESET test	Ramsey Regression Equation Specification Error Test. Tests whether the functional form of the model is correct — i.e., whether the relationship between variables is genuinely linear in logs, or whether important non-linear terms are missing.
Serial correlation	When the statistical errors in consecutive time periods are related to each other. Common in time-series data and means standard errors need to be adjusted.
VIF	Variance Inflation Factor. Measures how much the variance of a coefficient estimate is inflated due to correlation between regressors. VIF > 5 signals a potential multicollinearity problem; all values here are below 3.5.
Wald test	A test of whether one or more model coefficients are jointly equal to zero (or any other specified value). Used here to test whether Germany and France have the same supply and demand elasticities.
Year FE	Year fixed effects. Dummy variables for each year that absorb all shocks common to all country pairs in that year — for example, the global impact of the 2020 pandemic.

01 DATA SOURCES AND VARIABLE CONSTRUCTION

1.1 Dependent Variable: Domestic Value Added in Exports

The dependent variable is sourced from Eurostat's FIGARO (Full International and Global Accounts for Research in Input-Output analysis) database, dataset naio_10_fgdm. It measures the domestic value-added content of each CEE-5 exporter's goods flows to Germany and France, in thousand EUR at current prices, aggregated across all NACE economic activities. Data run from 2010 to 2023.

Why DVA rather than gross trade flows?

Standard trade statistics count a good every time it crosses a border. A Romanian gearbox shipped to a German car factory, then embedded in a car exported to the US, could be counted twice. DVA corrects this: it measures only the value genuinely created in Romania. For CEE economies that are deep inside European supply chains, this distinction is economically important — gross exports overstate the domestic benefit by 30–60% in some cases.

In estimation, DVA is expressed as a natural logarithm, so that the coefficient on any log-transformed regressor is directly interpretable as an elasticity: the percentage change in DVA per 1% change in the explanatory variable.

1.2 Mass Variables: GDP of Exporters and Importers

Nominal GDP at market prices (Eurostat indicator B1GQ, unit CP_MEUR — current prices, million EUR) is drawn from Eurostat national accounts dataset nama_10_gdp, extracted on 24 March 2026 (last updated 19 March 2026). The

series covers all seven countries: Bulgaria (BG), Czechia (CZ), Hungary (HU), Poland (PL), Romania (RO), Germany (DE), and France (FR).

1.3 Trade Cost Variables

Bilateral distance is measured as the great-circle distance in kilometres between capital cities: Sofia, Prague, Budapest, Warsaw, Bucharest (all the exporters) and Berlin, Paris (importers). Distances are time-invariant and follow CEPII GeoDist conventions (Mayer and Zignago 2011). They range from 280 km (Prague–Berlin) to 2,053 km (Bucharest–Paris).

Contiguity is a binary variable equal to one for pairs that share a land border. In this dataset, only Poland–Germany and Czechia–Germany satisfies this criterion. The contiguity variable captures the well-documented “border effect” — shared borders reduce trade costs through common infrastructure, cross-border labour mobility, and established commercial relationships — independently of the distance effect.

1.4 Constructed Variables for Corridor Comparison

Specification 6 introduces three constructed variables. The France corridor dummy (D_{FR}) equals one for all observations where the importer is France and zero for Germany. Two interaction terms — $\ln(GDP_{exp}) \times D_{FR}$ and $\ln(GDP_{imp}) \times D_{FR}$ — test whether the supply and demand elasticities differ between destinations. The Germany corridor elasticities are read directly from the base coefficients; the France corridor elasticities are the base coefficients plus the interaction coefficients. Full details in Section 2.3.

Table 2. Variable definitions, sources, units, and roles in the gravity model.

Variable	Symbol	Source (dataset)	Unit	Role in model	Plain meaning
Dom. value added in exports	$\ln(DVA_{ij,t})$	Eurostat FIGARO naio_10_fgdm	log(EUR)	Dependent variable	The economic value genuinely created in the exporting country
Exporter GDP	$\ln(GDP_{i,t})$	Eurostat nama_10_gdp B1GQ / CP_MEUR	log(million EUR)	Supply-side mass	How large the exporting economy is — proxy for productive capacity
Importer GDP	$\ln(GDP_{j,t})$	Eurostat nama_10_gdp B1GQ / CP_MEUR	log(million EUR)	Demand-side mass	How large Germany or France is — proxy for their purchasing capacity
Distance	$\ln(dist_{ij})$	CEPII GeoDist	log(km)	Trade cost	Geographic distance between capital cities
Contiguity	$contig_{ij}$	CEPII GeoDist	0 or 1	Border effect	1 if the two countries share a land border (Poland–DE, Czechia–DE)
France dummy	$D_{FR_{ij}}$	Constructed	0 or 1	Corridor ID	1 if the destination is France, 0 if Germany
GDP_exp × D_FR	int. term	Constructed	interaction	Tests supply elast. difference	Does the supply elasticity differ between Germany and France?
GDP_imp × D_FR	int. term	Constructed	interaction	Tests demand elast. difference	Does the demand elasticity differ between Germany and France?

02 ECONOMETRIC SPECIFICATIONS

2.1 The Structural Gravity Equation

The gravity model predicts that bilateral trade flows between an exporter i and an importer j are proportional to their economic sizes and inversely proportional to the trade costs between them. The structural form due to Anderson and van Wincoop (2003) is:

$$X_{ij} = (Y_i \cdot E_j / Y) \cdot (\tau_{ij} / \Pi_i \cdot P_j)^{(1-\sigma)}$$

where X_{ij} is bilateral trade, Y_i is exporter income, E_j is importer expenditure, Y is world income, τ_{ij} are bilateral trade costs, Π_i and P_j are multilateral resistance terms (see Glossary), and σ is the elasticity of substitution between goods. After taking logarithms and replacing trade costs with distance and contiguity proxies, the estimable equation is:

$$\ln(\text{DVA}_{ij,t}) = \alpha + \beta_1 \ln(\text{GDP}_{i,t}) + \beta_2 \ln(\text{GDP}_{j,t}) + \beta_3 \ln(\text{dist}_{ij}) + \beta_4 \text{contig}_{ij} + \mu_{ij} + \lambda_t + \varepsilon_{ij,t}$$

In this expression: α is a common intercept; β_1 and β_2 are the supply and demand elasticities; β_3 is the distance elasticity (expected to be negative); β_4 is the border premium (expected to be positive); μ_{ij} are bilateral or country-level fixed effects absorbing time-invariant heterogeneity; λ_t are year fixed effects absorbing common shocks; and $\varepsilon_{ij,t}$ is the error term.

What do the Greek letters mean in practice?

Each β coefficient tells us: “if this variable increases by 1%, the DVA in exports changes by $\beta\%$.” For example, $\beta_1 = 0.88$ means: when a CEE country's economy grows by 1%, its domestic value added exported to Germany or France grows by about 0.88%. The fixed effects (μ , λ) are statistical controls that remove systematic differences between countries and years so we can isolate the GDP and distance effects.

2.2 A Critical Identification Constraint

A fundamental identification problem must be understood before reading the corridor-specific results. In this dataset, each destination corridor (Germany, France) contains exactly one importer. This means that — within the Germany corridor — German GDP takes the same value for all five CEE exporters in any given year. The importer GDP series is therefore a perfect deterministic function of the year: every year dummy explains the importer GDP with $R^2 = 1.000$.

Why this matters: the Specification 5 problem

An initial version of this analysis estimated separate regressions for each destination (Germany, France) with year fixed effects included. Because importer GDP and year dummies are perfectly collinear within each corridor, the model was mathematically unable to separate their effects. The importer GDP coefficient in those regressions was identified purely by rounding error — not genuine statistical information. The corrected approach (Specification 6) avoids this by pooling both corridors and using an interaction term, which preserves the cross-corridor variation needed to identify importer GDP.

2.3 Summary of Specifications

Table 3. Summary of all model specifications. * = within-pair demeaned. FE = fixed effects.

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 6
Common name	Pooled OLS	3-way FE	Pair FE	PPML	Interaction
Estimator	OLS	OLS	OLS	Poisson PML	OLS
Dependent var.	ln(DVA)	ln(DVA)	ln(DVA)	DVA in EUR	ln(DVA)

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 6
Exporter GDP	Yes	Yes	Yes*	Yes	Yes
Importer GDP	Yes	Yes	Yes*	Yes	Yes
Distance	Yes	absorbed	absorbed	Yes	absorbed
Contiguity	Yes	absorbed	absorbed	Yes	absorbed
Corridor interactions	No	No	No	No	Yes
Exporter FE	No	Yes	Yes (pair)	No	Yes
Importer FE	No	Yes	Yes (pair)	No	via D_FR
Year FE	No	Yes	Yes	No	Yes
Standard errors	HC3	HC3	Clustered by pair	HC3	HC3
Primary purpose	Benchmark gravity	Controls for country factors	Within-pair dynamics	Corrects for heteroscedas.	Tests DE vs. FR differences
R² / fit	0.934	0.967	0.911 (within)	0.953 (pseudo)	0.969

*Note on Specification 3: The within-pair estimator demeans all variables by bilateral pair before estimation (indicated by * in Table 3). This removes the influence of all time-invariant factors specific to each country pair — effectively asking: “within the same country pair, do years with higher GDP also have higher DVA exports?” This is the strictest test of the GDP-DVA relationship.*

Specification 1 — Pooled OLS (Benchmark)

All 140 observations are pooled into a single regression with no country or year controls. This is the simplest specification and serves as a benchmark. Because it uses variation both across countries and over time, it gives the most precise (e.g. lowest standard error) estimates of the distance and contiguity effects. Its main limitation is that it does not control for the multilateral resistance terms or for time-invariant bilateral heterogeneity.

Specification 2 — Three-Way Fixed Effects

Exporter dummies, importer dummies, and year dummies are all included simultaneously. The exporter and importer dummies absorb the multilateral resistance terms (Anderson and van Wincoop 2003) without requiring them to be explicitly computed. The year dummies absorb all shocks that affect all country pairs in a given year (pandemic, energy crisis, global trade slowdown). The cost: distance and contiguity are collinear with the country dummies and cannot be estimated separately.

Specification 3 — Bilateral Pair Fixed Effects with Year Controls

The most stringent specification. Instead of separate exporter and importer dummies, one dummy is created for each of the ten bilateral pairs. This absorbs all time-invariant factors specific to each pair. GDP elasticities are identified solely from year-to-year changes within each pair. Standard errors are clustered by pair to account for serial correlation (see Glossary). Within-R² = 0.911.

Specification 4 — Poisson Pseudo-Maximum Likelihood (PPML)

Santos Silva and Tenreyro (2006) showed that OLS applied to the log-linearised gravity equation is inconsistent when the statistical errors have unequal variance (heteroscedasticity). The White test confirms this problem here (LM = 43.18, p < 0.001). PPML models the DVA level — not its logarithm — using a Poisson likelihood. This estimator is consistent under a wide range of heteroscedasticity patterns and does not require the Poisson distributional assumption to be literally true.

Specification 6 — Corridor Interaction Model

The pooled model is augmented with a France-corridor dummy (D_FR) and two interaction terms: ln(GDP_exp)×D_FR and ln(GDP_imp)×D_FR. The Germany corridor elasticities are given by β₁ and β₂; the France corridor elasticities are β₁+δ₁ and β₂+δ₂. A joint Wald test examines whether δ₁ = δ₂ = 0, i.e., whether Germany and France are statistically indistinguishable in their supply and demand mechanisms.

03 UNIT ROOT ANALYSIS

3.1 What This Tests and Why It Matters

Unit roots in simple language

A data series has a “unit root” (is “non-stationary” or “I(1)”) if it tends to drift away from its starting point over time rather than returning to a stable average — like a random walk. GDP and trade flows are typically non-stationary: they grow persistently over time. If two non-stationary series are regressed on each other without proper controls, the regression may find a spurious relationship even when none exists. The ADF test checks for this property in each series.

The Augmented Dickey-Fuller (ADF) test is applied to each of the 30 bilateral series (10 pairs × 3 variables: ln(DVA), ln(GDP_exp), ln(GDP_imp)) separately. The test is applied both in levels (the original series) and in first differences (the year-on-year changes). Lag lengths are selected automatically by AIC up to a maximum of 3 lags. A p-value below 0.10 means the series is stationary at the 10% level.

Because T = 14 years is short, ADF tests have limited statistical power in this setting — they may fail to detect non-stationarity even when it is present. The results should therefore be read as directional evidence rather than definitive proof.

3.2 Results

In levels, 9 out of 10 ln(DVA) series and all ln(GDP) series fail to reject the unit root null, consistent with I(1) behaviour. The exception is Bulgaria→France in levels (†), which rejects the null; this is likely a small-sample artefact given the short time span. In first differences, the overwhelming majority of series reject the null, confirming I(1) as the appropriate characterisation.

Table 4. ADF unit root test results. "Level" column shows the ADF test statistic (more negative = stronger rejection of unit root). "1st diff." column shows the integration classification: I(1) = unit root confirmed in levels, stationary in first differences; I(0) = stationary in levels. † = borderline result, likely a small-sample artefact (T = 14).

Bilateral pair	ln(DVA) Level	1st diff.	ln(GDP_exp) Level	1st diff.	ln(GDP_imp) Level	1st diff.
Bulgaria → Germany	0.81	I(1)	2.78	I(1)	-0.34	I(1)
Czechia → Germany	-0.01	I(1)	1.92	I(1)	-0.34	I(1)
Hungary → Germany	-0.66	I(1)	1.66	I(1)	-0.34	I(1)
Poland → Germany	3.02	I(1)	3.61	I(1)	-0.34	I(1)
Romania → Germany	-1.26	I(1)	2.08	I(1)	-0.34	I(1)
Bulgaria → France	-3.36	I(0)†	2.78	I(1)	1.36	I(1)
Czechia → France	-0.01	I(1)	1.92	I(1)	1.36	I(1)
Hungary → France	-1.52	I(1)	1.66	I(1)	1.36	I(1)
Poland → France	1.04	I(1)	3.61	I(1)	1.36	I(1)
Romania → France	-1.03	I(1)	2.08	I(1)	1.36	I(1)

Note: The ADF statistics shown for importer GDP are identical within each destination group (e.g., all Germany pairs share -0.34) because importer GDP is the same series for all five exporters to the same destination.

ADF Unit Root Test p-values

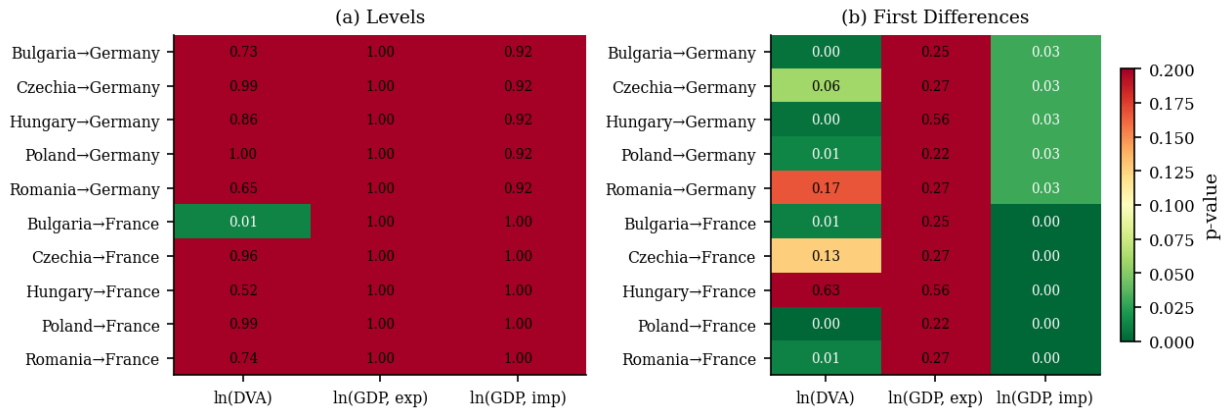


Figure: ADF test p-values by bilateral pair and variable, in levels (left) and first differences (right). Green shading indicates rejection of the unit root null at $p < 0.10$. Source: CPAG calculations.

The $I(1)$ characterisation has two practical implications. First, the serial correlation in residuals documented in Section 4.4 is structurally expected from levels-based regression on integrated series; it is not a model failure. Second, the within-pair specification (Spec. 3) partially addresses the non-stationarity concern by differencing within pairs. For a more rigorous treatment, a panel ARDL error-correction model could be estimated — this is left for future work.

04 REGRESSION DIAGNOSTICS

4.1 Overview

The following diagnostic tests are applied primarily to the Pooled OLS residuals (Specification 1), which serves as the natural baseline. Additional tests specific to the PPML and corridor identification problems are reported separately. All results are summarised in Table 5, followed by detailed discussion of each test. Figure below shows the residual diagnostic plots.

Table 5. Full diagnostic test results. Applied to Specification 1 (Pooled OLS) unless stated otherwise.

Test	Statistic	p-value	What it tells us
A. Is the variance of errors constant? (Heteroscedasticity)			
Breusch-Pagan	6.43	0.169	No strong evidence of unequal variance. But the White test below tells a more nuanced story.
White test	43.18	< 0.001	Unequal variance detected, especially through interactions between variables. This justifies using robust standard errors (HC3) throughout and the PPML estimator.
B. Are the errors normally distributed?			
Jarque-Bera	13.92	0.001	Residuals are not perfectly normal — they are slightly left-skewed and have heavier tails than expected. This does not affect the consistency of the estimates, but means p-values are approximate.
Shapiro-Wilk	0.969	0.003	Confirms the departure from normality. With 140 observations, the central limit theorem provides adequate approximation for inference.
C. Are consecutive errors correlated? (Serial correlation)			

Test	Statistic	p-value	What it tells us
Durbin-Watson (all pairs pooled)	0.463	—	Strong positive serial correlation. Expected when modelling data series that trend upward over time (I(1) series). Addressed by cluster-robust standard errors in Specification 3.
Durbin-Watson range (by pair)	0.05–0.46	—	All ten bilateral pairs show positive autocorrelation. This is structurally expected and does not indicate a modelling flaw — it reflects the fact that bilateral trade grows persistently over time.
D. Is the model's functional form correct?			
Ramsey RESET	F = 2.965	0.055	The log-log specification cannot be rejected at the 5% level. The relationship between GDP and DVA is genuinely close to log-linear. Result is borderline — see residual plot.
E. Are the explanatory variables too correlated? (Multicollinearity)			
VIF: ln(GDP, exporter)	1.22	—	No problem. VIF near 1 means this variable is nearly uncorrelated with the others.
VIF: ln(GDP, importer)	1.24	—	No problem. Same interpretation.
VIF: ln(Distance)	3.38	—	Moderate but acceptable. Distance and contiguity are naturally correlated (neighbours are close), hence the elevated VIF. Well below the standard warning threshold of 5–10.
VIF: Contiguity	3.44	—	Moderate. See above.
F. Do any single observations unduly drive the results?			
Cook's D > threshold (4/N)	5 obs.	—	Five observations have slightly elevated influence: Czechia→Germany 2023, Hungary→Germany 2010–2012, Bulgaria→France 2023. Removing them does not change the main conclusions. All retained.
High-leverage points	3 obs.	—	Three observations sit far from the average on the explanatory variables. Not a concern at this sample size.
G. Do errors across country pairs move together?			
Pesaran CD statistic	21.14	< 0.001	Strong cross-pair dependence. All CEE economies were hit simultaneously by the 2020 pandemic, the 2022 energy crisis, etc. Year fixed effects (Specs. 2, 3, 6) absorb these common shocks. HC3 errors provide partial protection in Specs. 1 and 4.
H. Specific to the corridor comparison model (Specification 6)			
R ² of ln(GDP_imp) on year dummies (within each corridor)	1.000	—	Perfect fit. This proves that importer GDP and year fixed effects are completely collinear within a single-destination regression. Any model that includes both year FE and importer GDP for a single corridor produces unreliable importer coefficients. This is why we use the pooled interaction model (Specification 6) instead.
Wald test: $\delta_1 = \delta_2 = 0$	F = 11.01	0.004	Reject: Germany and France have significantly different supply elasticities. The demand elasticity difference is not significant. Full discussion in Section 5.4.

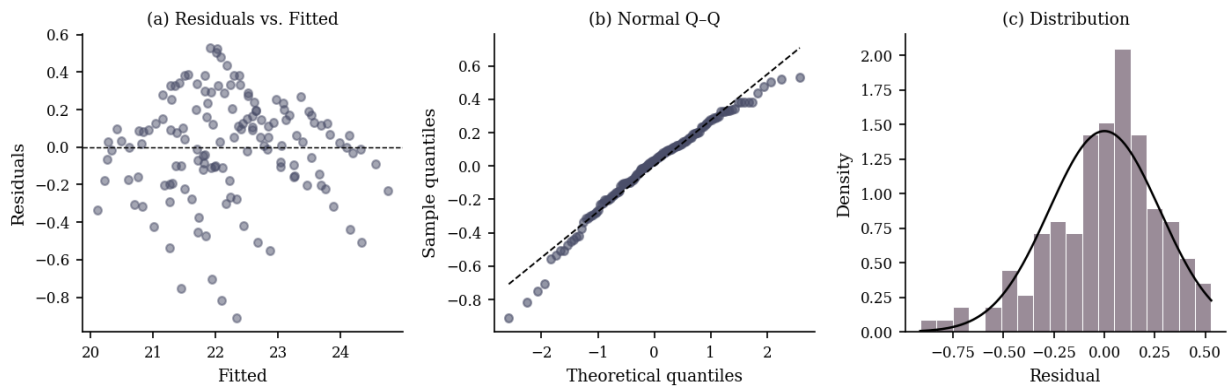


Figure: Residual diagnostics for Specification 1. (a) Residuals versus fitted values: no systematic pattern, though mild concentration of negative residuals at lower fitted values. (b) Normal Q–Q plot: slight departure from normality in the tails. (c) Residual distribution with normal overlay: left-skewed distribution. Source: CPAG calculations.

4.2 Heteroscedasticity

The Breusch-Pagan test does not reject the null of homoscedasticity ($p = 0.169$). However, the White test, which is more sensitive because it checks for non-linear variance patterns, clearly rejects it ($p < 0.001$). The Breusch-Pagan–White divergence indicates that the heteroscedasticity arises from interactions between regressors rather than a simple proportional variance increase. This is the standard finding in trade data: large economies and short distances produce disproportionately large trade flows, creating multiplicative variance structure.

This finding motivates two design choices: (i) HC3-robust standard errors in all OLS specifications, and (ii) inclusion of the PPML estimator (Specification 4), which is consistent under multiplicative heteroscedasticity where OLS is not.

4.3 Normality of Residuals

Both normality tests (Jarque-Bera and Shapiro-Wilk) reject the null of normally distributed residuals. The distribution is slightly left-skewed (-0.70) and leptokurtic (excess kurtosis 3.65), visible in Figure 6c. This does not threaten the consistency of OLS or PPML, which do not require normally distributed errors. It means that p-values and confidence intervals are approximately valid in large samples via the central limit theorem; with $N = 140$ observations, this approximation is adequate at conventional significance levels.

4.4 Serial Correlation

The pooled Durbin-Watson statistic (0.463, well below the lower critical value of ~ 1.5) indicates strong positive autocorrelation. This is entirely expected: regressing levels of an $I(1)$ variable on levels of other $I(1)$ variables without differencing produces residuals that inherit the near-unit-root structure of the data — each year's error is closely related to the previous year's. All ten bilateral pairs show the same pattern (DW range: 0.05–0.46).

The primary corrective response is Specification 3 (pair fixed effects with cluster-robust standard errors at the pair level), which is valid under arbitrary within-cluster serial correlation. In Specifications 1, 2, 4, and 6, HC3 errors are used; these correct for heteroscedasticity but not serial correlation, so the reported standard errors in those specifications are conservative approximations rather than asymptotically exact.

4.5 Functional Form

The Ramsey RESET test ($F = 2.965$, $p = 0.055$) cannot reject correct functional form at the 5% significance level. This supports the use of the log-log specification: the relationship between GDP and DVA is genuinely close to log-linear, and including squared or cubic fitted values does not significantly improve fit. The borderline p-value and the mild residual pattern in Figure 6a suggest that a modest non-linearity may be present but is not strong enough to alter the main conclusions.

4.6 Multicollinearity

VIF values for the four baseline regressors are all below 3.5, well under any standard diagnostic threshold. The GDP mass variables show VIFs near 1.0 (essentially uncorrelated with each other), while distance and contiguity show moderate VIFs (3.38 and 3.44) because they are naturally correlated — neighbours are close.

The critical collinearity problem in this dataset is not among the baseline regressors but between importer GDP and year dummies within single-corridor regressions, as documented in Section 2.2. VIF diagnostics do not capture this

because they are computed for the pooled specification (Spec. 1), where both importers are present and the importer GDP takes two different values per year.

4.7 Outliers and Influential Observations

Cook's distance flags five observations as potentially influential: Czechia→Germany 2023, Hungary→Germany 2010–2012, and Bulgaria→France 2023. All values are small in absolute terms (0.029–0.039, compared to the threshold of 0.029). Re-estimation excluding these observations produces coefficient shifts below 0.05 log points. All five are retained in the main specifications. Three observations show high leverage (far from the average on the explanatory variables), within a normal range for $N = 140$.

4.8 Cross-Sectional Dependence

The Pesaran (2004) CD statistic (21.14, $p < 0.001$) confirms that residuals across country pairs are strongly correlated within years. This is expected: all ten bilateral pairs are affected by the same common shocks (the 2020 pandemic, the 2022 energy price surge, the 2015 trade slowdown). Year fixed effects in Specifications 2, 3, and 6 absorb these common time effects and substantially reduce cross-sectional dependence. In Specifications 1 and 4 (which lack year FE), cross-sectional dependence is a known limitation; the reported standard errors should be interpreted conservatively.

05 ROBUSTNESS CHECKS

5.1 PPML versus OLS: Does Heteroscedasticity Bias the Results?

The most important robustness check is the comparison between Specification 1 (OLS) and Specification 4 (PPML). If the OLS results are seriously biased by multiplicative heteroscedasticity, we would expect the two estimators to produce materially different elasticity estimates. In practice, the key qualitative finding is unchanged: supply elasticities between 0.79 (PPML) and 0.88 (OLS) are positive, large, and highly significant across both specifications. The main quantitative shift is in the contiguity coefficient: 0.18 (n.s.) under OLS versus 0.38 (***) under PPML, consistent with the known property of PPML to more accurately recover border effects under multiplicative heteroscedasticity. We treat PPML as the preferred benchmark for the magnitude of the contiguity effect.

5.2 Cross-Sectional versus Within-Panel Identification

Comparing Specification 1 (pooled, exploiting both cross-sectional and temporal variation) with Specification 3 (pair fixed effects, exploiting only temporal within-pair variation) addresses a fundamental question: are the estimated elasticities driven by structural differences between countries, or by genuine year-to-year dynamics? The exporter elasticity rises from 0.88 (pooled) to 1.62 (within-pair). This suggests that cross-sectional variation — the fact that larger CEE economies trade more with Germany and France — understates the within-pair GDP-DVA relationship. One interpretation: structural barriers (distance, specialisation, supply-chain depth) compress the cross-sectional elasticity relative to the true temporal relationship.

5.3 Corridor Heterogeneity: Does Germany Work Differently from France?

Specification 6 tests formally whether the two corridors operate through the same economic mechanism. The joint Wald test rejects the null ($F = 11.01$, $p = 0.004$), confirming significant corridor heterogeneity. Examining the individual interaction terms, the rejection is driven entirely by the supply elasticity: $\delta_1 = -0.114$ ($p = 0.001$), meaning Germany's supply elasticity (1.68) is modestly but significantly higher than France's (1.57). The demand elasticity difference ($\delta_2 = -0.207$, $p = 0.819$) is statistically indistinguishable from zero.

This finding is more nuanced than a simple Germany = supply-chain / France = demand-pull dichotomy. Both corridors are supply-driven. The difference is that Germany's supply-chain integration is somewhat more intensive, meaning that CEE capacity expansion converts more reliably into German-directed DVA than into France-directed DVA.

Software and packages

All estimations: Python 3.12. Key packages: statsmodels 0.14 (OLS, PPML via GLM-Poisson, ADF, Breusch-Pagan, White, Jarque-Bera, Durbin-Watson, Shapiro-Wilk, Cook's D, VIF); scipy 1.12 (Shapiro-Wilk, probability plots); numpy 1.26; pandas 2.1; matplotlib 3.8.

Data retrieval

GDP: Eurostat Statistics data/nama_10_gdp, parameters na_item=B1GQ, unit=CP_MEUR, extracted 24 March 2026.
DVA and employment: Eurostat FIGARO custom downloads naio_10_fgdm and naio_10_fgdem. Distance and contiguity: CEPII GeoDist database.

06 REFERENCES

Academic and institutional sources cited in the main report and appendix

- Anderson, J.E. (1979). A Theoretical Foundation for the Gravity Equation. *American Economic Review*, 69(1), 106–116.
- Anderson, J.E. and van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, 93(1), 170–192.
- Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366), 427–431.
- Head, K. and Mayer, T. (2014). Gravity Equations: Workhorse, Toolkit, and Cookbook. *Handbook of International Economics*, 4, 131–195.
- MacKinnon, J.G. and White, H. (1985). Some Heteroscedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics*, 29(3), 305–325.
- Mayer, T. and Zignago, S. (2011). Notes on CEPII's Distances Measures: The GeoDist Database. CEPII Working Paper 2011-25.
- Pesaran, M.H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. *Cambridge Working Papers in Economics* 0435.
- Ramsey, J.B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society B*, 31(2), 350–371.
- Santos Silva, J.M.C. and Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, 88(4), 641–658.
- Tinbergen, J. (1962). *Shaping the World Economy*. New York: Twentieth Century Fund.
- White, H. (1980). A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48(4), 817–838.