Data, Judgment Call Explanation, & Brief Literature Review

Table 1: OLS estimates for the 2023 cross-section gravity regression

Variable	Estimate	Std. Error	t-stat	p-value
Intercept	−45.410	1.2551	−36.179	6.13×10^{-277}
$\ln \mathrm{GDP}_{\mathrm{exp}}$	1.393	0.009903	140.66	$< 10^{-300}$
$\ln \mathrm{GDP}_{\mathrm{imp}}$	0.941	0.008889	105.90	$< 10^{-300}$
$\ln$ Distance	−1.692	0.027629	−61.235	$< 10^{-300}$
$\ln \mathrm{Remoteness}_{\mathrm{exp}}$	0.578	0.084026	6.883	6.04×10^{-12}
$\ln \mathrm{Remoteness}_{\mathrm{imp}}$	1.224	0.094323	12.979	2.36×10^{-38}
Border dummy (contiguity)	0.871	0.13885	6.270	3.69×10^{-10}

Notes: Dependent variable is $\ln(\mathrm{exports}_{ij} + 1)$. Sample size $N = 18{,}391$. Classic (OLS) standard errors reported.

Explanation:

For this replication task, we use data specifically pertaining to the year 2023. This is the most recent year that has complete WDI data on GDP, and for this reason, we target 2023 specifically.

For this task, we will use CEPII GeoDist, WDI, and COMTRADE. Since these three datasets include everything we need, we are able to avoid using OECD datasets.

For this reason, in contrast to the previous two replications, this analysis is done over a far wider range of countries instead of being limited to only OECD countries.

In order to avoid errors pertaining to missing data, we will conduct the analysis on countries that appear in all three data sets.

This results in a total of 182 countries being covered in our data, and they are as follows: ABW, AFG, AGO, ALB, AND, ARE, ARG, ARM, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, BWA, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COG, COL, COM, CPV, CRI, CYM, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ESP, EST, ETH, FIN, FJI, FRA, FRO, FSM, GAB, GBR, GEO, GHA, GIN, GMB, GNB, GNQ, GRC, GRD, GTM, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBR, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNG, MOZ, MRT, MUS, MWI, MYS, NAM, NER, NGA, NIC, NLD, NOR, NPL, NRU, NZL, OMN, PAK, PAN, PER, PHL, PLW, PNG, POL, PRT, PRY, QAT, RUS, RWA, SAU, SDN, SEN, SGP, SLB, SLE, SLV, SOM, STP, SUR, SVK, SVN, SWE, SWZ, SYC, TCA, TCD, TGO, THA, TJK, TKM, TTO, TUN, TUR, TUV, TZA, UGA, UKR, URY, USA, UZB, VCT, VNM, VUT, WSM, ZAF, ZMB, ZWE.

Here, as instructed, we use OLS rather than PPML, though a PPML approach would arguably be more sound.

Our script, in summary, follows the following process:

We import 2023 Comtrade trade flows, CEPII bilateral distance/contiguity data, and World-Bank GDP figures and trim each dataset to the 182 countries present in all three. We aggregates Comtrade to one export value per exporter-importer pair.

We build gravity-equation covariates—logs of exports (+1 to keep recorded zeros), GDPs, distance, remoteness, and a land-border dummy, and assembles them into a regression table.

From there we use a plain-OLS log-linear gravity model with those six predictors and report the results and basic diagnostic summaries.

Literature Review with citations:

Literature Review (PDF)

Citations (data):

Anderson, James E., and Eric van Wincoop. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, vol. 93, no. 1, 2003, pp. 170–192. doi:10.1257/000282803321455214.

CEPII. *GeoDist Database*. CEPII, n.d., `https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=6`. Accessed 9 May. 2025.

United Nations. *Commodity Trade Statistics Database (UN Comtrade)*. Department of Economic and Social Affairs, Statistics Division, 2025, `https://comtrade.un.org/`. Accessed 9 May. 2025.

World Bank. *World Development Indicators*. World Bank Group, 2025, https://data.worldbank.org/indicator. Accessed 9 May. 2025.