

S2 Text: Model setup and sensitivity analysis

Socio-economic distance matrix

In the real model setup, socio-economic distance between countries c_{ij} is constructed using fixed effects coefficients from the statistical model. In Table 1 in main text, we estimated several statistical models, including some with fixed effects by pairs of countries. The socio-cultural distance between two countries is then taken as the opposite of the fixed effect coefficient for this pair, for the full statistical model. Summary statistics for the 29x29 matrix (28 EU countries and Jersey which is in the database a distinct country from UK but has an important total turnover since many British companies are located there for fiscal advantage reasons) are shown below in Table S2-A. We observe that distances are rather localized but with large outliers, and an important proportion are not defined (339 out of 841), corresponding to couple of countries having no exchange at all in the dataset.

Table S2-A. Summary statistics of the socio-cultural distance estimated with a fixed effects model.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-2.327	1.440	2.451	2.538	3.447	11.797	339

Synthetic model setup

The procedure for the synthetic sector setup is the following:

- Sectors distributions follow log-normal densities with most mass in $[0; 1]$
- Large cities are more innovative and more diverse. Assuming that sectors are ordered by innovativity (the larger the sector index the larger the innovativity), this assumption is translated by taking a mode and variance of 0.5 for the largest city and a mode and variance of $1/K$ for the smallest, and a linear interpolation

between the two. For each city, we define

$$\log(e_i) = \frac{(\log(E_i) - \log(E_{imin}))}{(\log(E_{imax}) - \log(E_{imin}))} * (1/2 - 1/K) + 1/K$$

- Log-normal parameters (μ_i, σ_i) for each city are then fixed by $\mu_i - \sigma_i^2 = \log(e_i)$ and $-3\sigma_i^2 - 2\log(\exp(\sigma_i^2) - 1) = \log(e_i)$
- σ_i^2 is the unique positive root of $f(X) = 0$ with $f(X) = -3X - 2\log(\exp(X) - 1) - \log(e_i)$ and $\mu_i = \log(e_i) + \sigma_i^2$.

Sensitivity analysis

The table S2-B gives the full results for the Global Sensitivity Analysis, for all model indicators and free parameters. We give the first order indices and the total indices.

Table S2-B. Saltelli sensitivity indices, for indicators in rows and parameters in columns. We give for each pair the first order index (F) and the total order index (T).

	γ_G		γ_D		γ_S		γ_W		γ_O		γ_D	
	F	T	F	T	F	T	F	T	F	T	F	T
Internationalisation	0.2	0.3	0.7	0.7	0.001	0.009	$4 \cdot 10^{-4}$	0.007	0.03	0.04	0.02	0.04
Metropolisation	0.02	0.1	0.02	0.2	0.002	0.1	0.001	0.09	0.2	0.6	0.3	0.6
Modularity	0.3	0.4	0.6	0.6	0.004	0.02	$3 \cdot 10^{-4}$	0.01	0.005	0.03	0.002	0.03
Avg. com. size	0.008	0.09	0.01	0.1	0.002	0.07	0.003	0.04	0.3	0.6	0.4	0.6
Degree entropy	0.006	0.02	0.003	0.02	0.006	0.03	0.008	0.02	0.5	0.5	0.5	0.5
Weight entropy	0.04	0.1	0.03	0.1	0.008	0.08	0.01	0.07	0.4	0.5	0.4	0.5