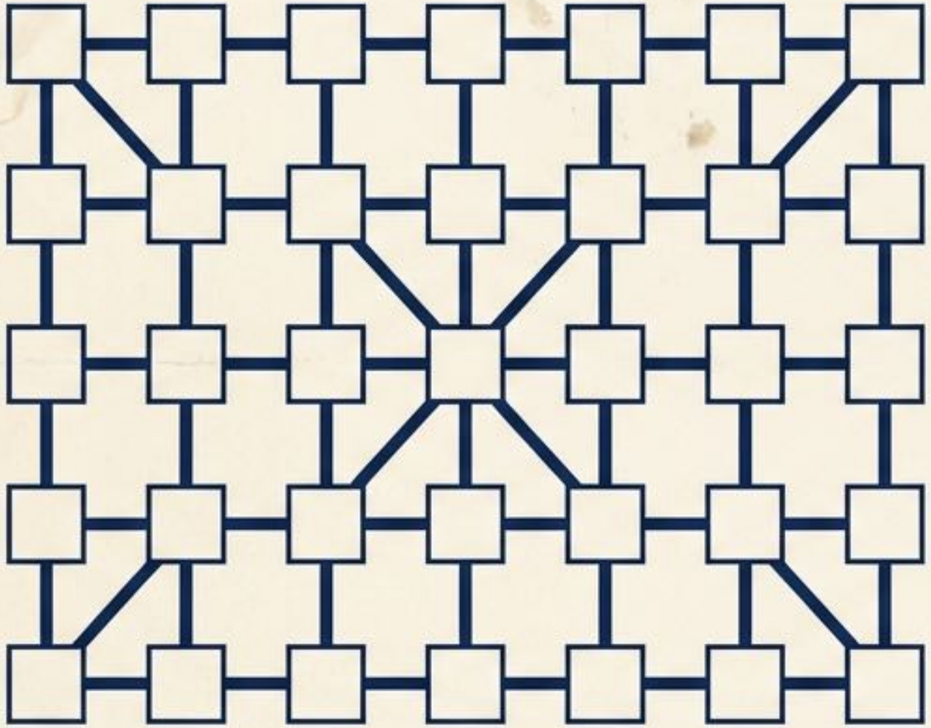
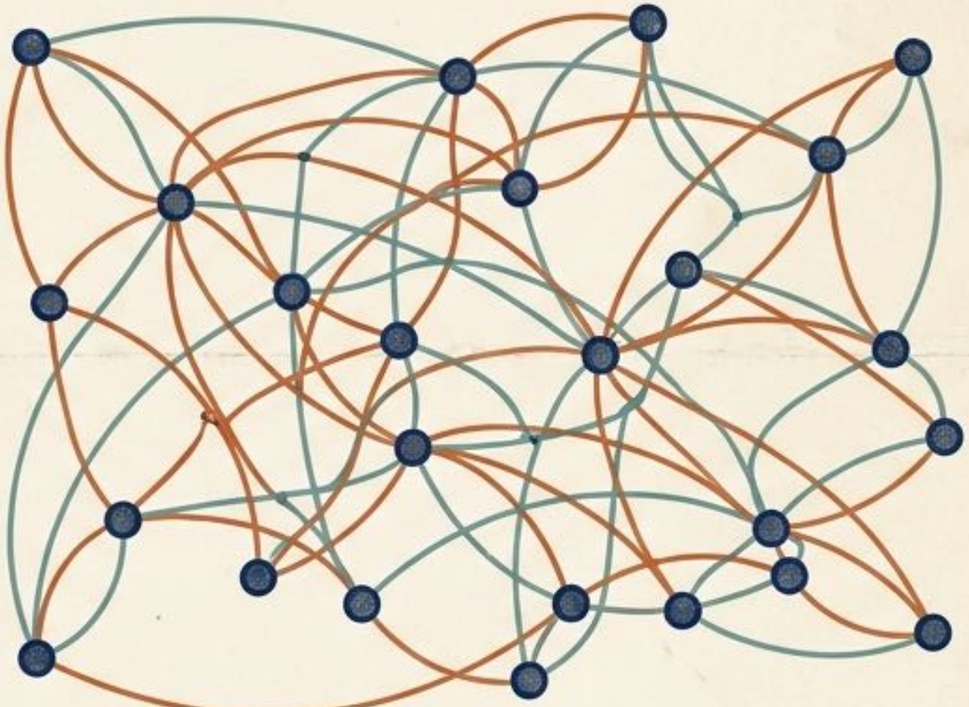
Who Are My Neighbors? Bayesian Estimation of Spatial Weight Matrices

A Methodological Study Guide

THE ASSUMPTION



THE ESTIMATION



THE STATUS QUO

Spatial econometrics usually hands you the neighborhood map before you start. If that map is wrong, every spillover estimate is wrong too.

THE INTERVENTION

This tutorial does not choose the map. It estimates it.

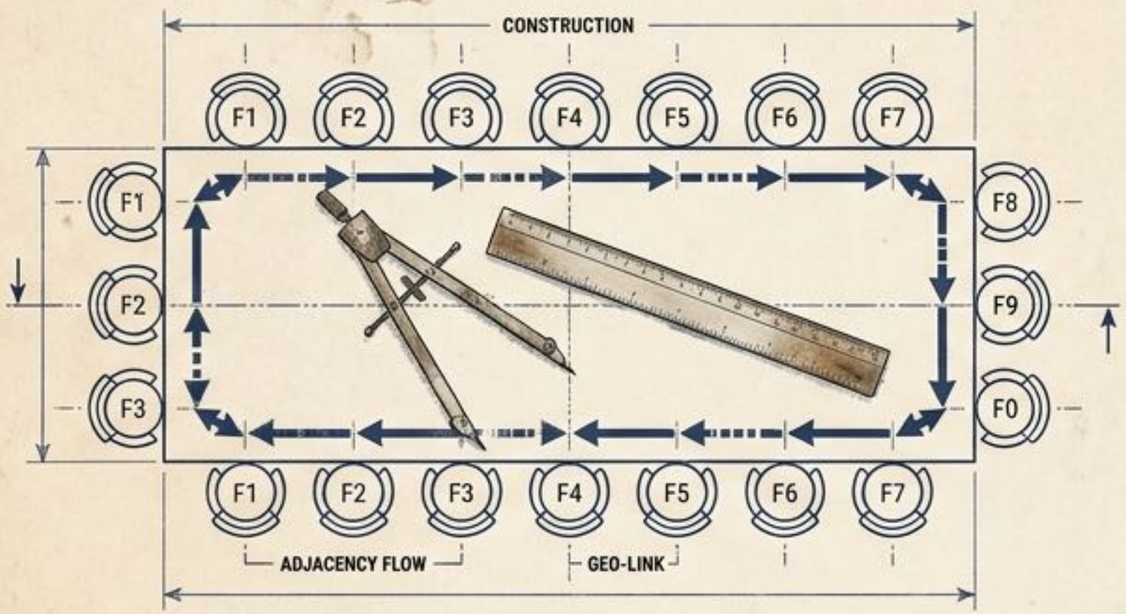
THE APPLICATION

90 European NUTS-1 regions, 19 years (2001–2019), 1,710 observations.

THE FUNDAMENTAL LIMITATION OF ASSUMED GEOGRAPHY

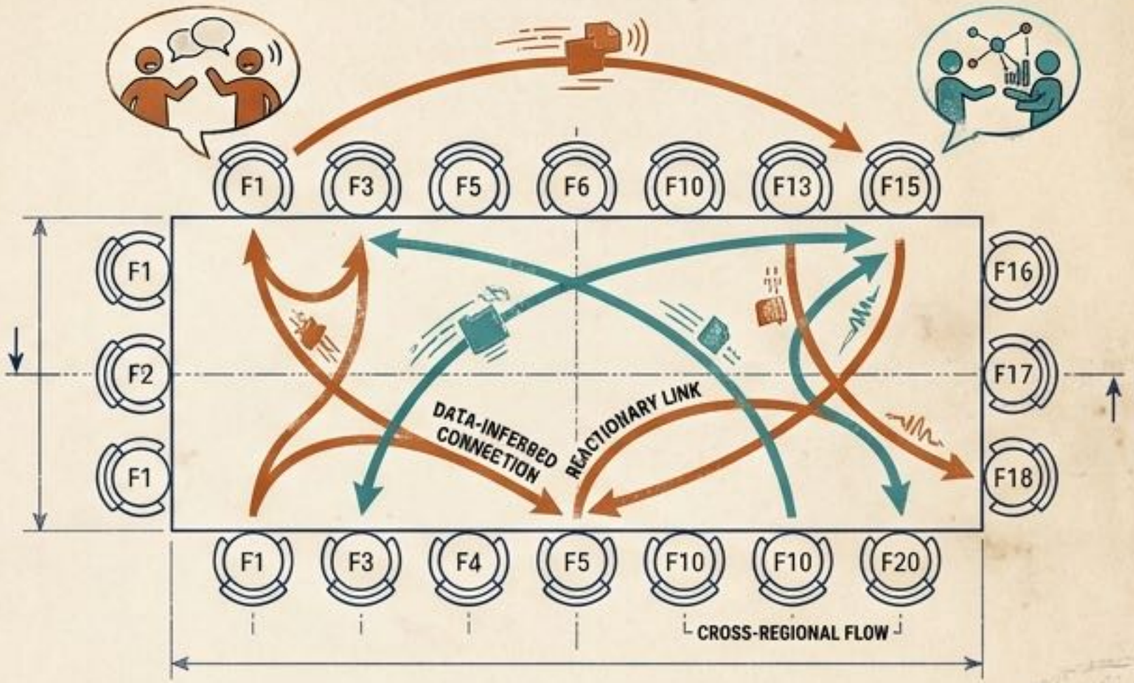
THE SEATING CHART ANALOGY

A STANDARD PRACTICE



Assume spillovers based on who sits next to whom (Geography).

B THIS METHODOLOGY

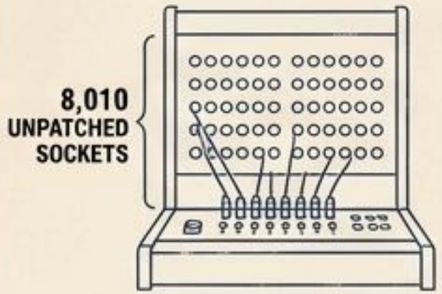


Infer the neighborhood from who reacts together (Data).

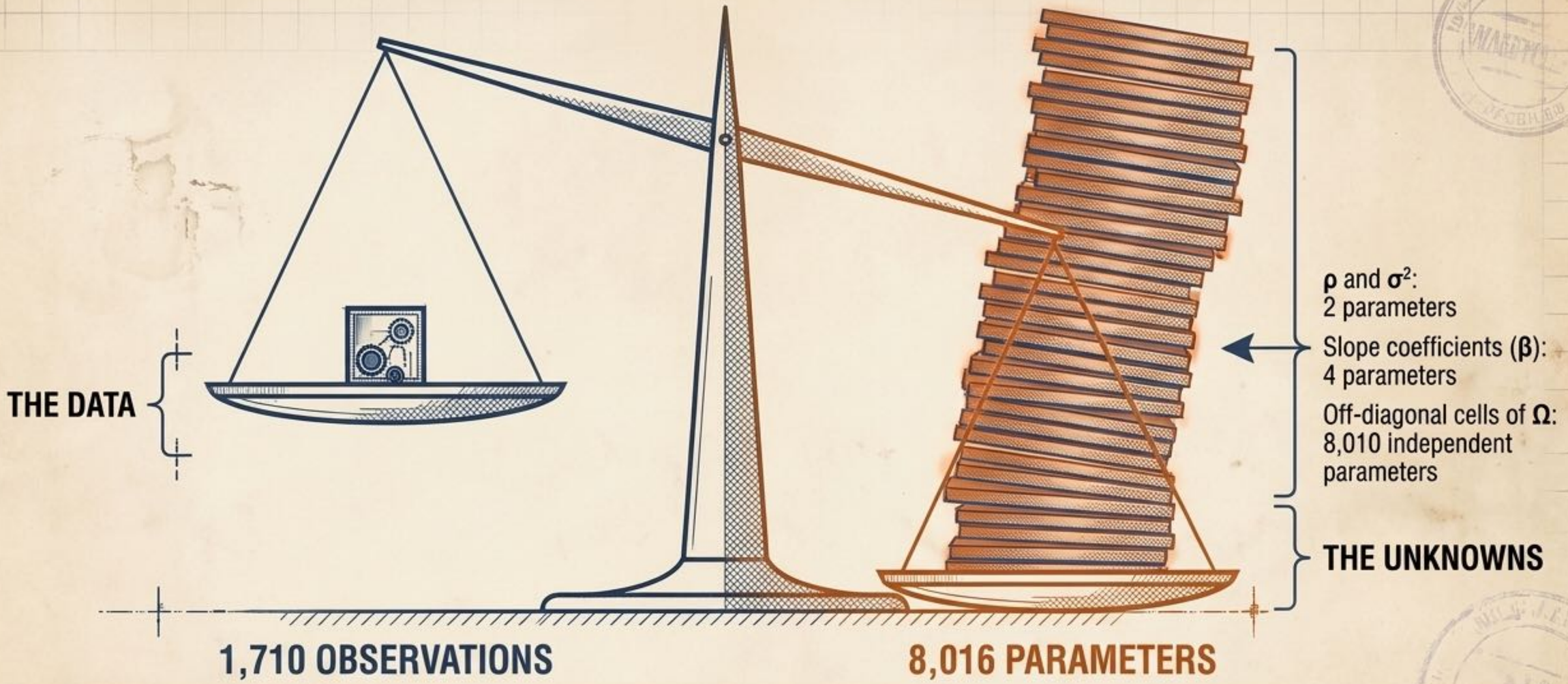
THE PROBLEM: Standard models make all conclusions about spillovers conditional on the assumed map being right.

THE QUESTION: Which European regions actually behave as each other's neighbors for productivity growth—and does it matter if we assume the map or estimate it?

VOCABULARY FOR ENDOGENOUS SPATIAL NETWORKS

CONCEPT & MATH	RIGOROUS DEFINITION	INTUITIVE ANALOGY
Adjacency Matrix (Ω) & Spatial Weight Matrix (W)	**Ω** is an n × n binary table (1 = neighbor, 0 = not). **W** is its row-standardized counterpart, where row i sums to 1.	 An unlabelled phone switchboard with 8,010 unpatched sockets. Row-standardization splits your daily call minutes evenly among everyone you call.
Spatial Autoregressive Parameter (ρ)	A single system-wide scalar tracking how strongly neighbor outcomes feed into a region's. Stability requires $\|\rho\| < 1$.	 The gain knob on a public-address system. Turn it up, and the room gets louder from the same original sound.
Spatial Multiplier: $(I - \rho W)^{-1}$	The device that translates one local shock into a system-wide outcome through infinite feedback rounds.	 A rumor retold office to office. Each retelling reaches more people but gets slightly quieter.

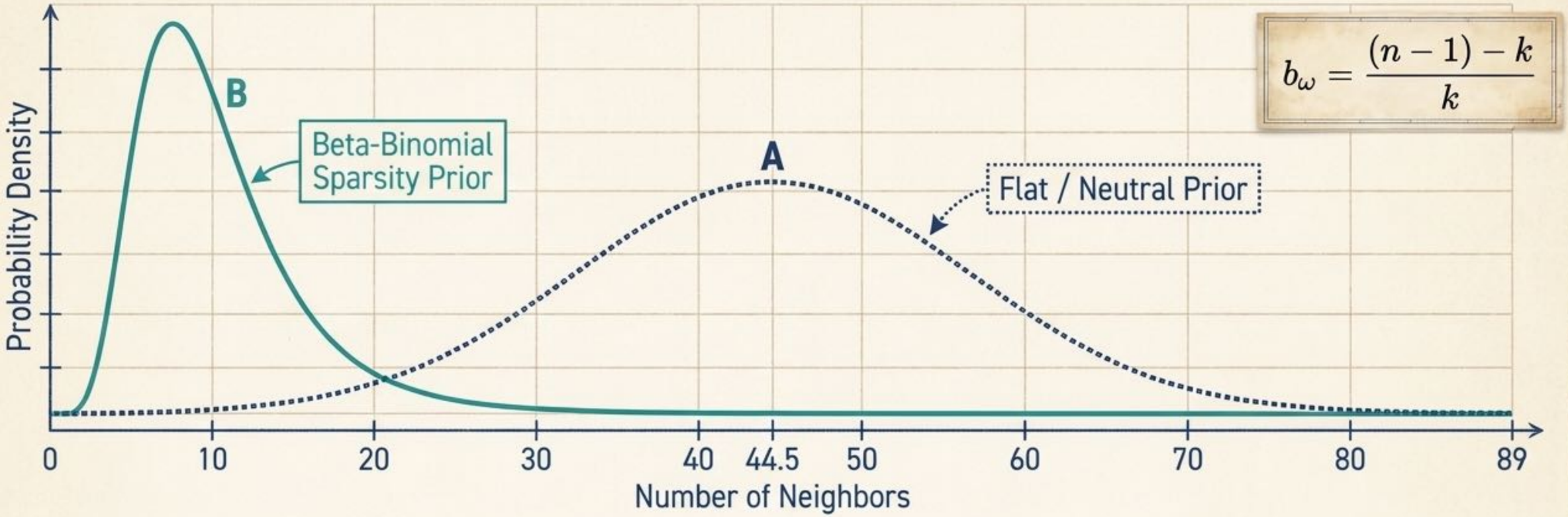
THE CURSE OF DIMENSIONALITY MANDATES A PRIOR



THE RATIO: 4.7 unknowns per observation.

THE IMPLICATION: A model with 4.7 parameters per observation cannot be estimated by likelihood alone. The likelihood does not have a unique peak to climb. The Bayesian prior is not a caveat about the method; it is the method.

SHRINKAGE: MAKING THE IMPOSSIBLE TRACTABLE



$$b_\omega = \frac{(n-1)-k}{k}$$

The Danger of "Neutral"

A flat weight, mathematically, asserts a Binomial (n-1, 0.5) prior. In a 90-region panel, this quietly expects every region to be connected to 44.5 neighbors—half of Europe.

The Solution

Anchor the expected degree at a sparse target (e.g., $k = 7$). For 90 regions and a target of 7, the prior probability of any individual link is just 7/89 = 7.87%.

THE MECHANICS OF ESTIMATION: THE GIBBS SAMPLER

ORDINARY REGRESSION

Textbook Gaussian and Inverse-Gamma updates for β and σ^2.

LINK BY LINK (The 8,010)

Element-wise Bernoulli updates. Score the model with link ij=1, then ij=0. Flip a weighted coin.

THE GRIDDY-GIBBS STEP

Updating ρ. Visually scoring the posterior density on a fine 60-point grid.

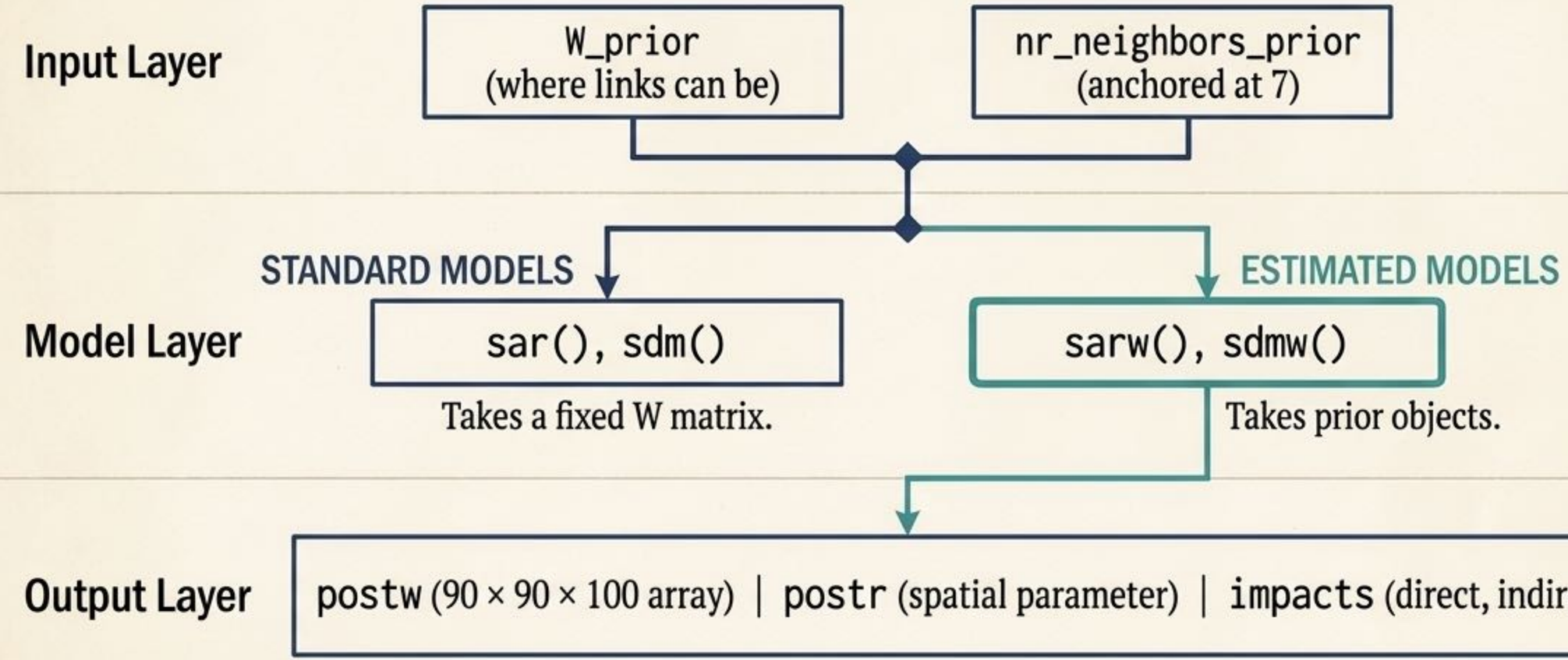
THE METAPHOR

Tuning a piano string by string. Fix one string given how the others currently sound, move to the next, and repeat thousands of times.

THE SPEED SECRET (SHERMAN-MORRISON)

Flipping one link requires recomputing a 90×90 determinant. The Sherman-Morrison identity calculates this rank-one change instantly, reducing computation from weeks to minutes.

SOFTWARE ARCHITECTURE AND IMPLEMENTATION

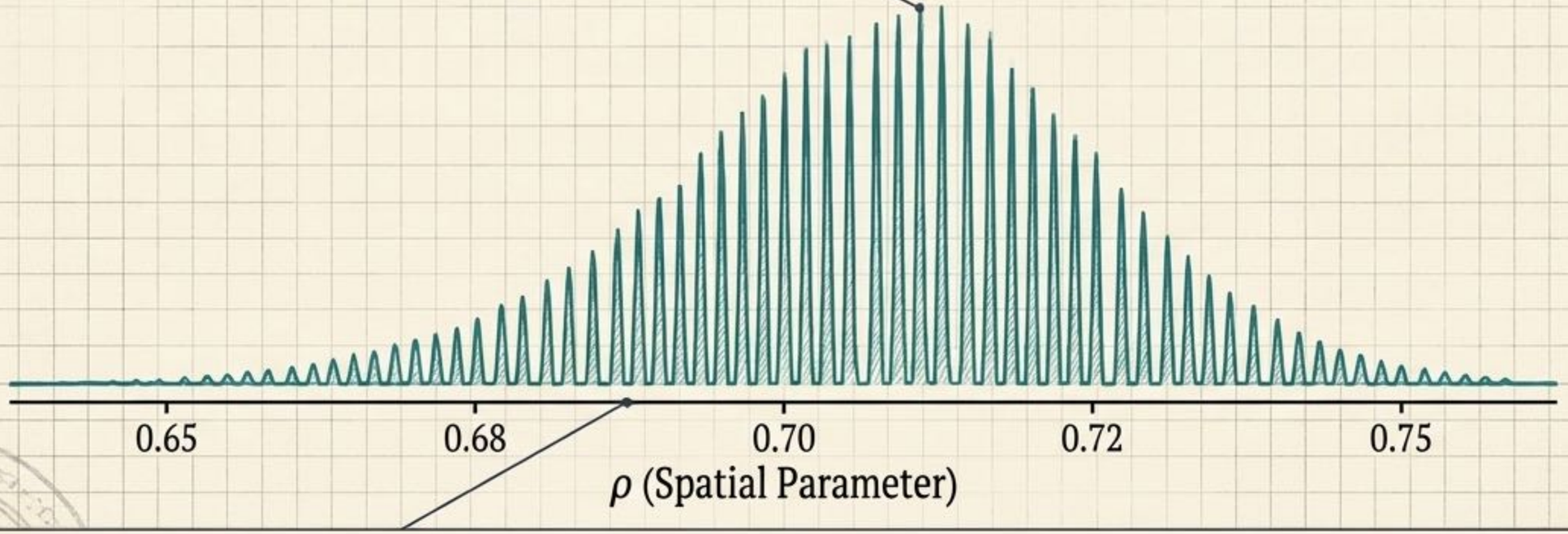


The Rule: The plain function name takes a neighborhood map you supply. The 'w' suffix estimates it.

Implementation Note: In Durbin specifications, the X and Z matrices must be completely disjoint to prevent rank deficiency.

INTERPRETING THE STRUCTURAL PARAMETER OUTPUT

The Result: $\rho = 0.71322$ (Posterior SD: 0.01574).
Interpretation: Spatial dependence is strong and precisely estimated. A one-point rise in neighbor-average growth moves own-growth by 0.71 points before feedback.

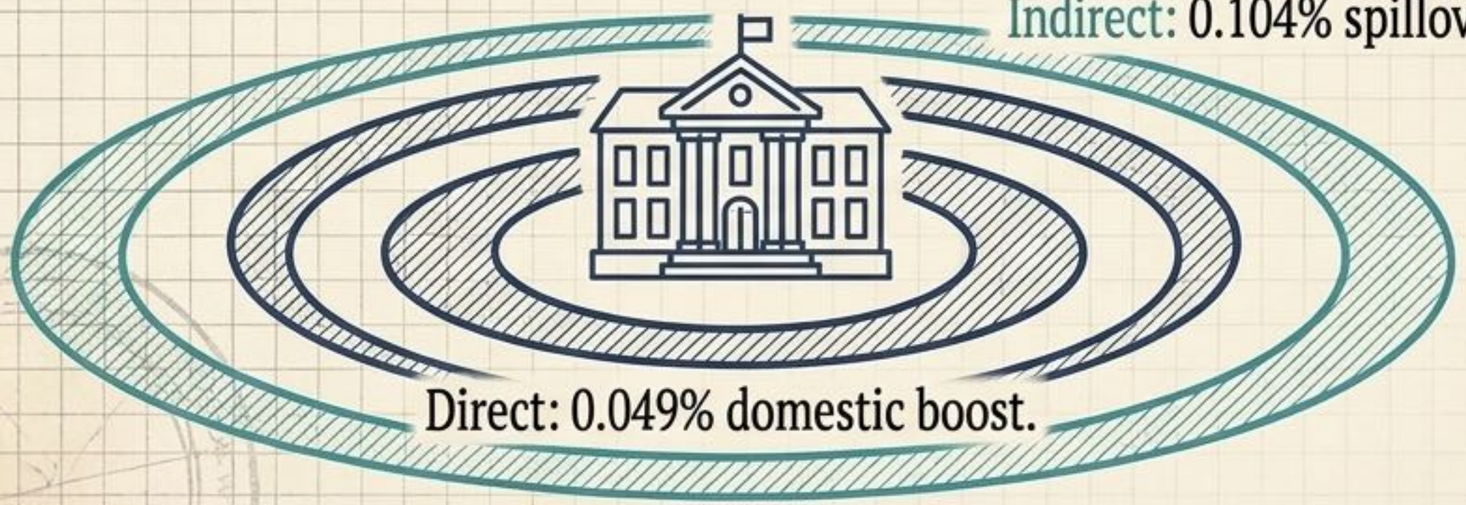


The Visual Artifact: Why the comb structure? This is the Griddy Gibbs sampler made visible. Because ρ is sampled from a 60-point grid, it can only ever land on one of 60 exact values. It is a feature of the sampler design, not a broken model.

DECOMPOSING IMPACTS: THE SPATIAL MULTIPLIER AT WORK

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \rho^3 W^3 \ldots$$

Round 1 (Yourself) | Round 2 (Neighbors) | Round 3 (Neighbors' Neighbors)



THE MATHEMATICAL STRAITJACKET

In a basic SAR model, every variable shares the exact same indirect-to-direct ratio.

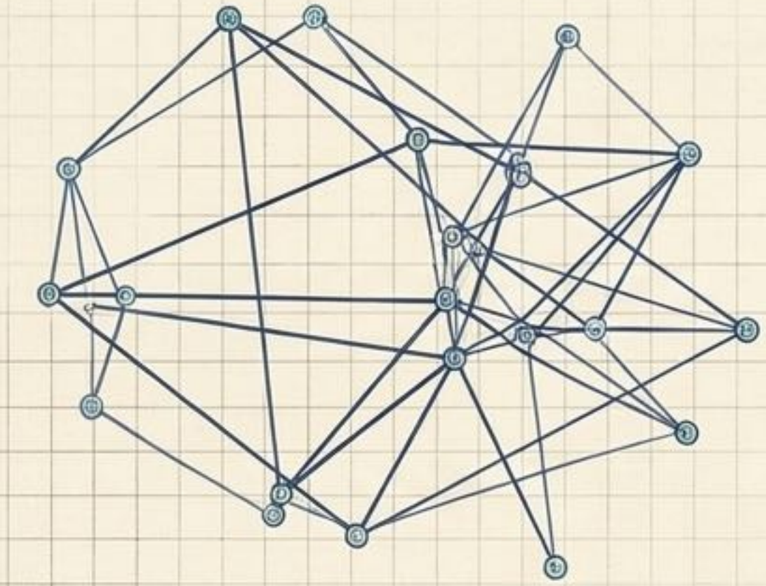
Here, spillovers are uniformly 2.11x the own-region effect.

ECONOMIC IMPLICATION

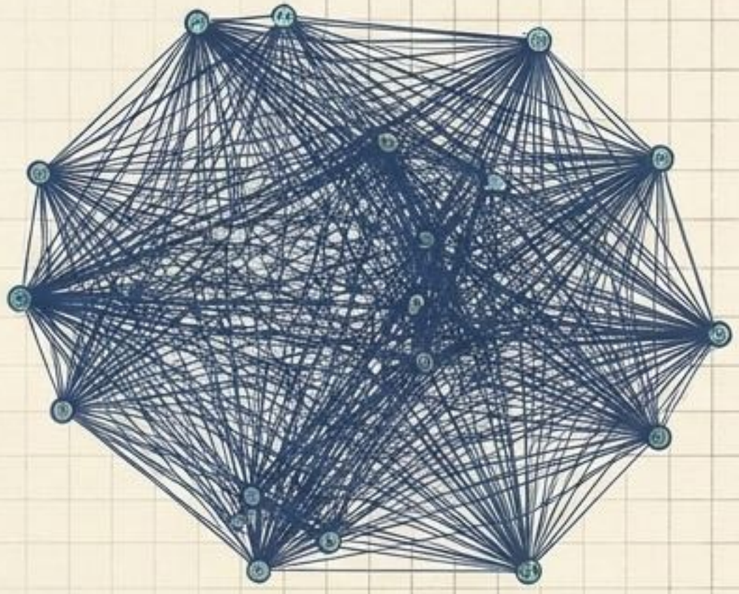
Whoever pays for the universities captures roughly a third of the measured return. The rest leaks across borders.

Reading the Estimated Network Distribution

Direct Network (W)



Multiplier Network



Posterior Link Probability

The share of the 100 retained draws where a specific link is switched on. A link seen in every single draw has a probability of 1.0.

Direct W vs. Multiplier

The direct network (W) averages 6.47 neighbors per region. The Multiplier network eventually connects almost everyone, because third-order neighbors still carry a weight of 0.362.

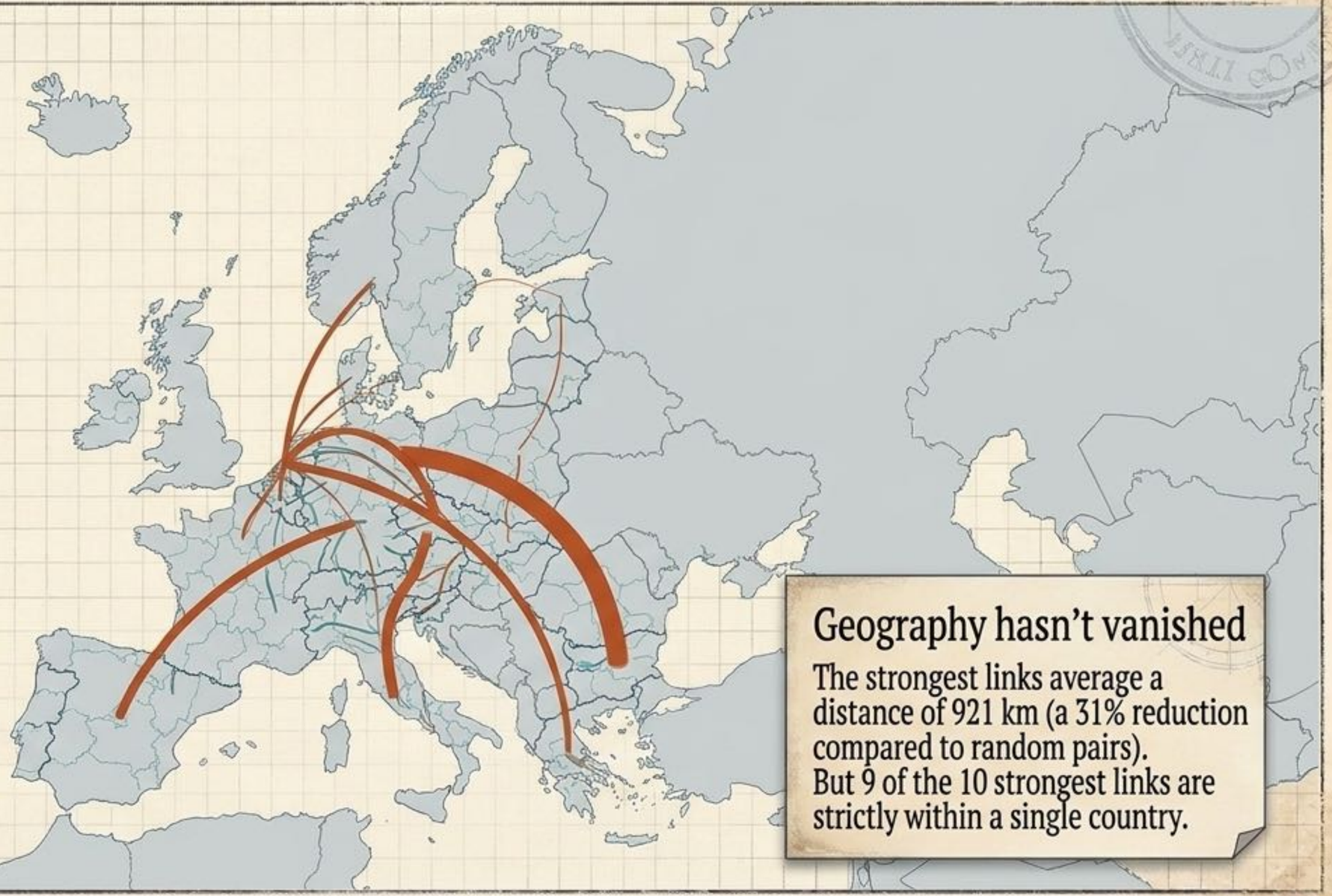
Warning

The posterior mean is a summary of uncertainty. Averaging over 100 uncertain networks makes the average look denser than any single model actually entertained.

Application Results: The True European Map

The Core Finding

The estimated network organizes by nationality and supranational bloc, not just physical adjacency.



The Data

Regions put 5x more weight on compatriots than random chance predicts (35.6% vs 7.1%), and twice as much on supranational blocs.

Geography hasn't vanished

The strongest links average a distance of 921 km (a 31% reduction compared to random pairs).
But 9 of the 10 strongest links are strictly within a single country.

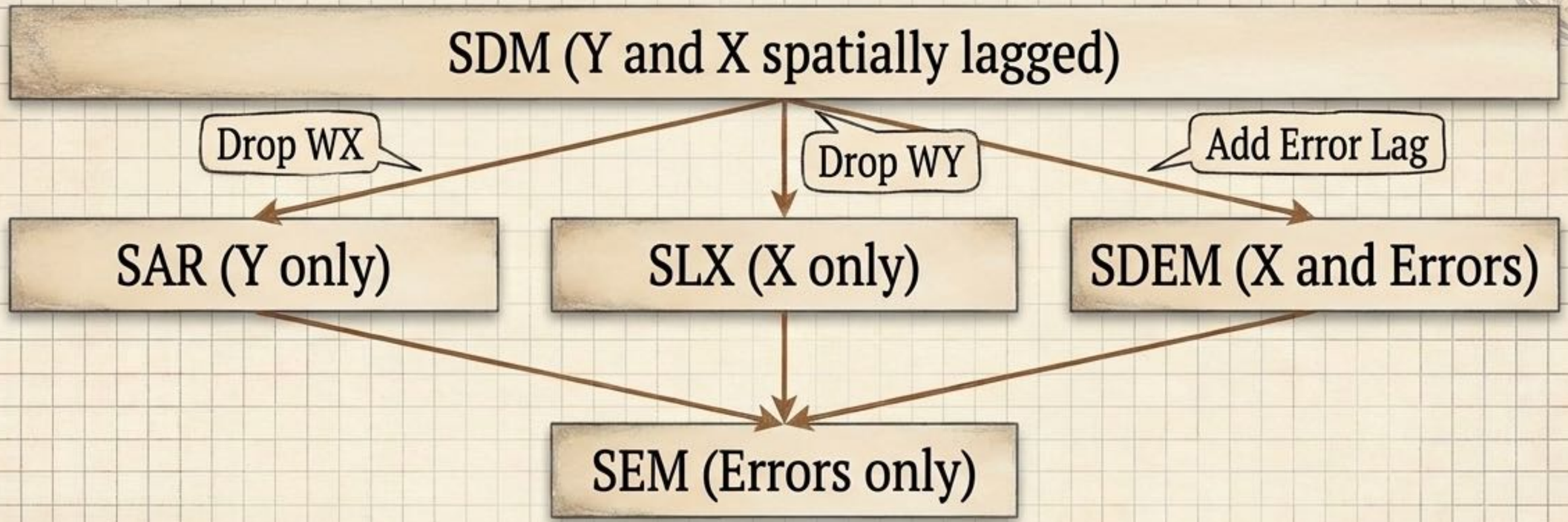
Does Changing the Map Change the Answer?

Map Source	**Estimated W (Data)**	Queen Contiguity (Borders)	7-Nearest-Neighbor (Distance)
Neighbors per Region	**6.47 neighbors**	3.80 neighbors	7.00 neighbors
ρ Parameter	**ρ = 0.713**	ρ = 0.607	ρ = 0.719
Total Impact (Education)	**Impact: 0.00153**	Impact: 0.00066	Impact: 0.00074
Indirect-to-Direct Ratio	**Ratio: 2.11**	Ratio: 1.20	Ratio: 2.20

Qualitative Survival:
Signs agree across all maps. Convergence is negative; education spillovers are positive.

Quantitative Failure:
Magnitudes do not survive. The data-chosen map yields more than **twice** the total economic effect of either assumed geographic map.

The Model Taxonomy & Spatial Bias



Omitted Variable Bias

If the true data generation process involves spatially lagged regressors (SDM), estimating a SAR model will artificially shrink ρ to compensate.

The Rule

Getting the neighborhood map right does not rescue a researcher from getting the specification wrong.

Identification Conditions (De Paula et al., 2025)

Editorial Checklist

- ✓ I: $[W]_{ii} = 0$
- ✓ II: Row sum constraints
- ✓ III: $\rho\beta_1 + \beta_2 \neq 0$
- ✓ IV: Row standardization
- ✓ V: Heterogeneous second-order structure

⚠ Condition VI: $\rho > 0$ ⚠

The Assumption

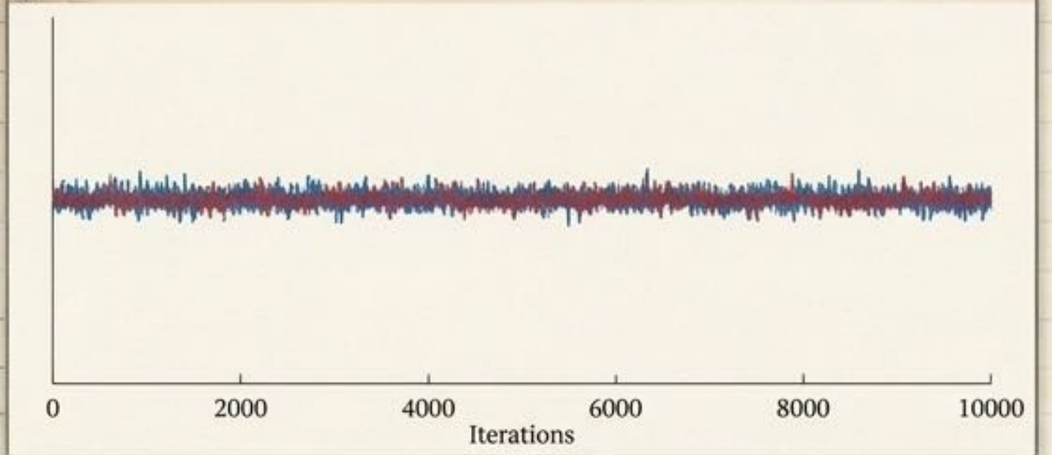
Non-negativity is imposed, not tested.

The Interpretation Caveat

The prior support restricts ρ to (0,1). Therefore, an estimate of $\rho = 0.713$ is evidence of the magnitude of positive spatial dependence, not evidence against negative spatial dependence. Negative values were excluded before the data were ever seen.

Diagnostics, Trust, and Academic Caveats

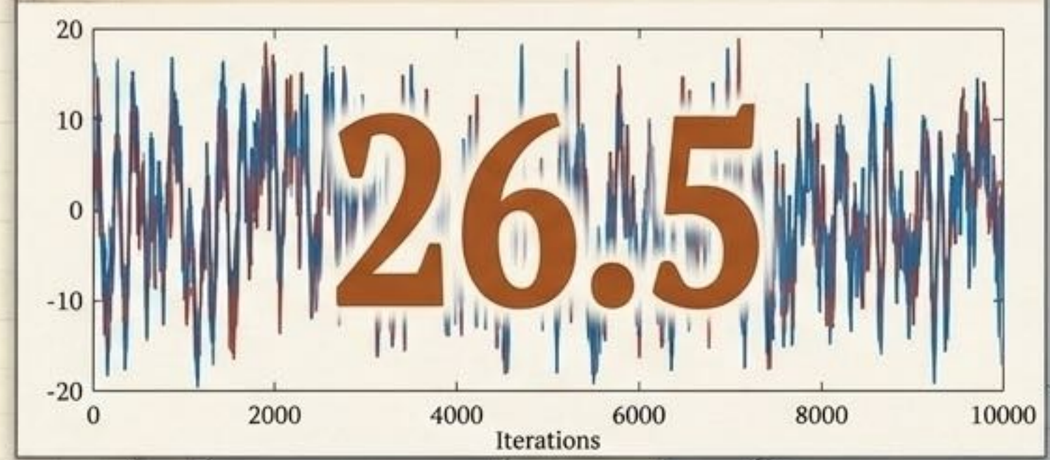
What To Trust



Posterior Means

Point estimates are stable across chains and reproduce exactly. The overall structural topology of the network is trustworthy.

What NOT To Trust



Standard Deviations & Intervals

An Effective Sample Size (ESS) of 26.5 for ρ cannot support a credible interval for a referee report. Marginal link rankings from small tutorial budgets (100 draws) are extremely noisy.

Final Caveat: Links represent residual co-movement, not proven causal mechanisms (trade, commuting). Shared exposure to shocks mimics connectivity. Naming the mechanism requires data on the mechanism.