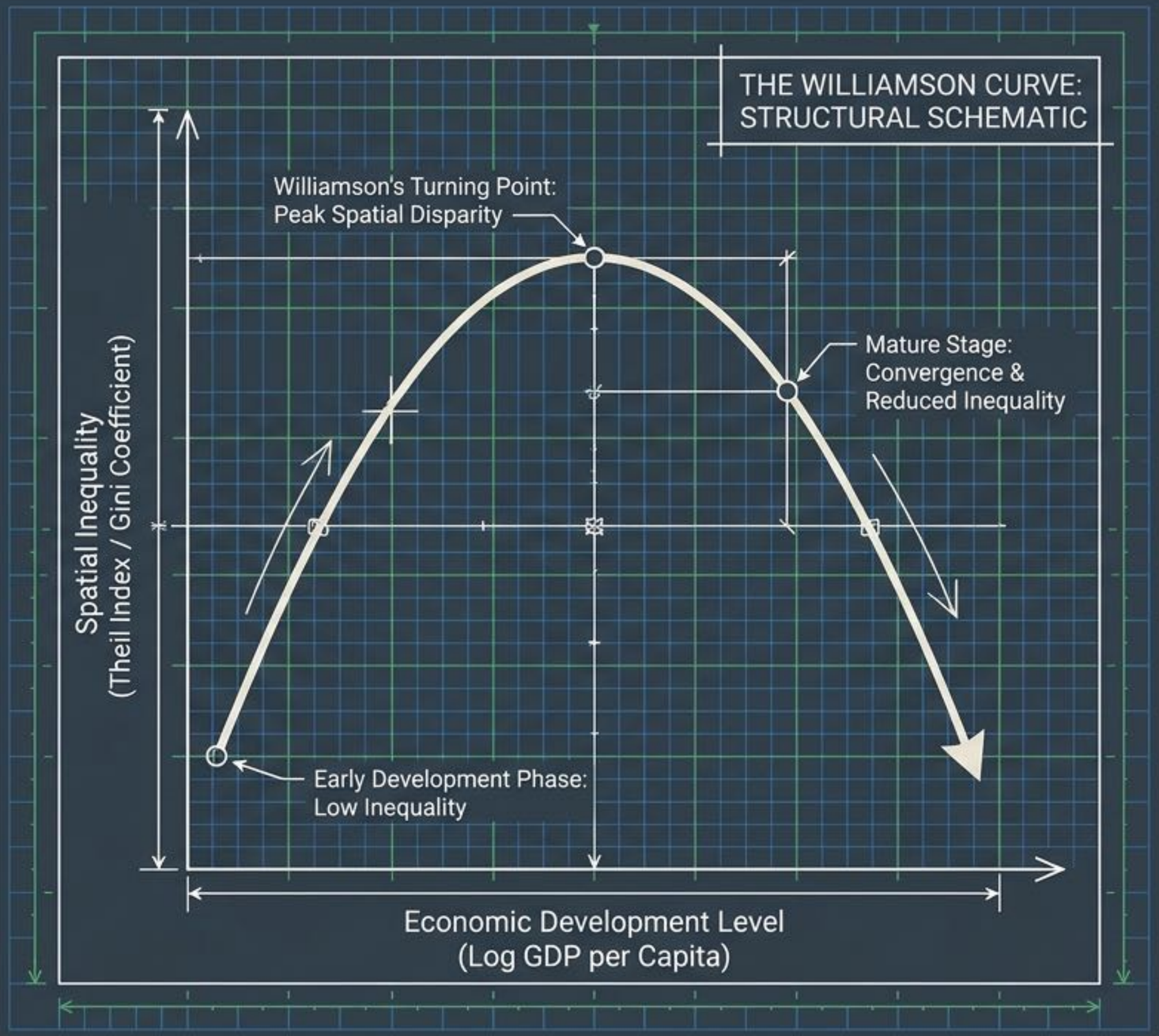
Spatial Inequality and Economic Development: Testing the Inverted-U Hypothesis

An Econometric Workbench for Estimating the Williamson Curve

Synthesizing cross-sectional and panel data methodologies to evaluate regional disparities across 56 countries (1980–2009).



NotebookLM

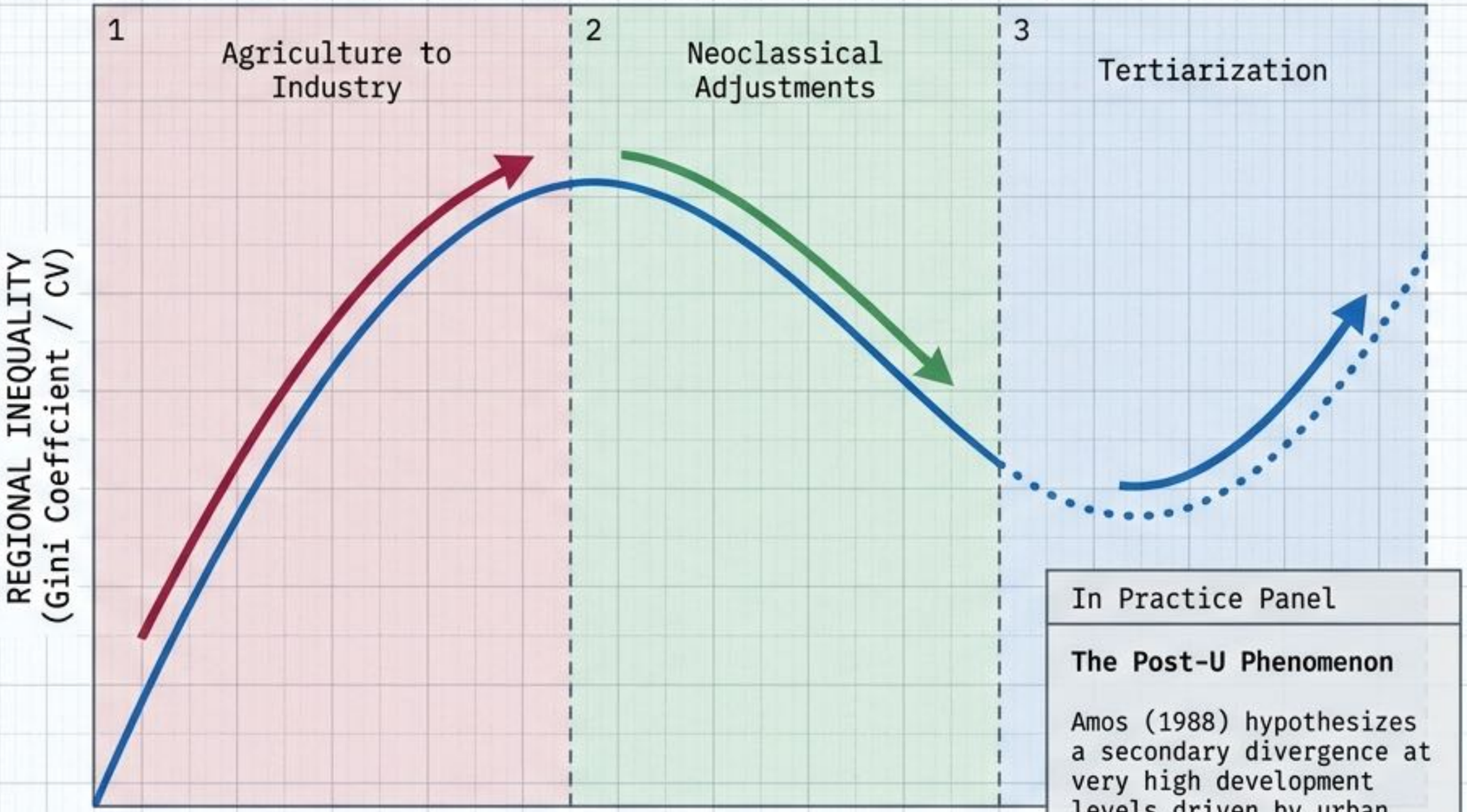
The Williamson Hypothesis: A Structural Trajectory

Methodology Callout

The Kuznets/Williamson Mechanism:

Industrialization is driven by point-source natural resources (e.g., coal/iron), causing initial spatial divergence.

Factor market adjustments (migration, capital mobility) later drive natural convergence.



In Practice Panel

The Post-U Phenomenon

Amos (1988) hypothesizes a secondary divergence at very high development levels driven by urban decay and the shift to modern service sectors (tertiarization).

NotebookLM

Quantifying Spatial Inequality: The Weighted Coefficient of Variation

Methodology Callout

$$WCV = \frac{1}{\bar{y}} \left[\sum_{i=1}^{n} p_i (\bar{y} - y_i)^2 \right]^{1/2} \qquad (1)$$

National average GDP p.c.

Population share

Regional GDP p.c.

Interpretation Key

Why Population Weights (p_i) Matter:

Unweighted measures (like standard deviation)
Unweighted measures (like standard deviation)
overstate inequality in sparsely populated, extreme-income regions.

In Practice Panel

The Canadian Territories Anomaly

The northern territories are extremely poor but house only 100,000 people. An unweighted measure would falsely signal severe nevere national economic fragmentation; WCV corrects this by weighting by demographic density.

NotebookLM

Measurement Challenges: Nominal vs. Real Sub-National Prices

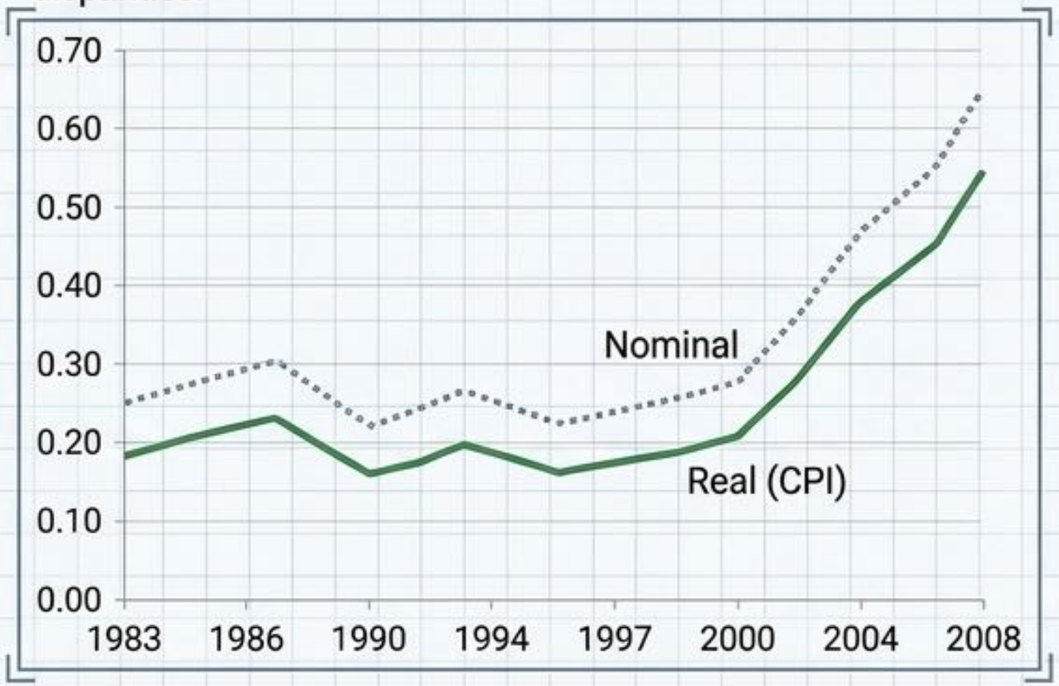
Interpretation Key

Nominal measures overstate real inequality because non-tradables (housing, food) adjust to local incomes (Engel & Rogers, 1996).

Price Disparity Matrix

Time-Series Adjustment (Canada)

Uses provincial CPIs (base 2002) to adjust panel data over time. Eliminates inflation noise but cannot measure baseline level disparities.

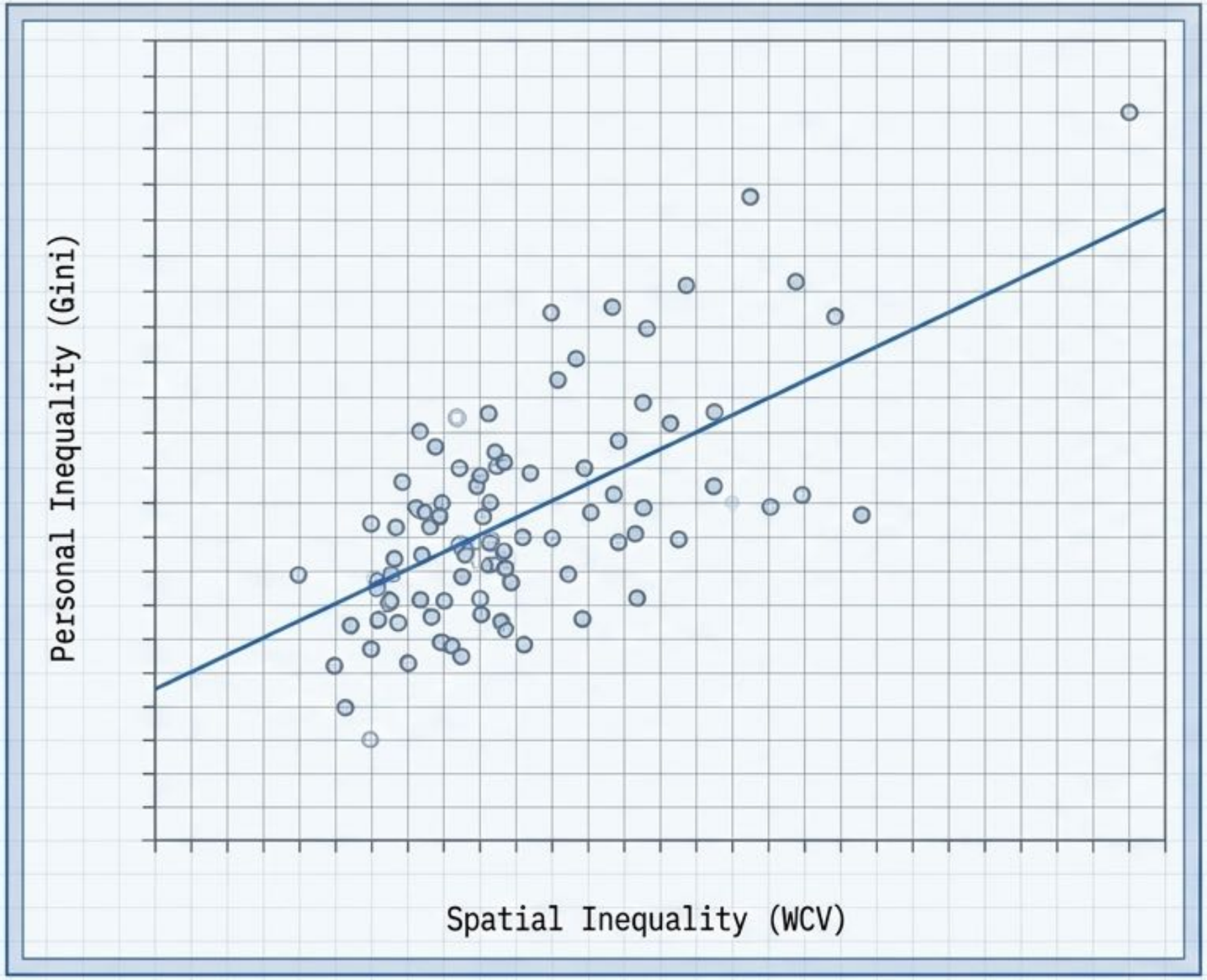


Level Adjustment (United States)

Uses experimental Regional Price Parities. Adjusting for real prices shrinks the WCV by roughly one-third, proving nominal data consistently upward-biases real spatial inequality.

NotebookLM

Horizontal vs. Vertical Disparities: Spatial vs. Personal Inequality



Methodology Callout

OLS Fit:
GINI_i = 0.313 + 0.152 * WCV_i
(Correlation: rho = 0.324)

Interpretation Key

Spatial inequality explains roughly one-third of personal inequality variance.

In Practice Panel

The Variance Paradox
High GINI does not guarantee high WCV.

USA: High personal inequality (GINI=0.47) but low spatial inequality (WCV=0.17).

Slovenia: Low personal inequality (0.27) but severe regional disparity (0.58).

NotebookLM

Empirical Strategy: The Methodological Choice Matrix

	Parametric (Polynomials)	Semiparametric (Flexible/Kernels)
Cross-Sectional Data (Between-country)	**OLS Regression (Levels)** Tests long-run structural changes across different countries at a fixed point in time.	**Robinson (1988) Estimator** Controls for parameter heterogeneity without enforcing a global functional form.
Panel Data (Within-country)	**Fixed-Effects OLS** Isolates short-to-medium-term within-country dynamics (1980–2009).	**Baltagi & Li (2002) Estimator** Combines fixed effects with non-parametric smoothing (B-splines).

NotebookLM

Parametric Specification: Modeling the Kuznets Curve

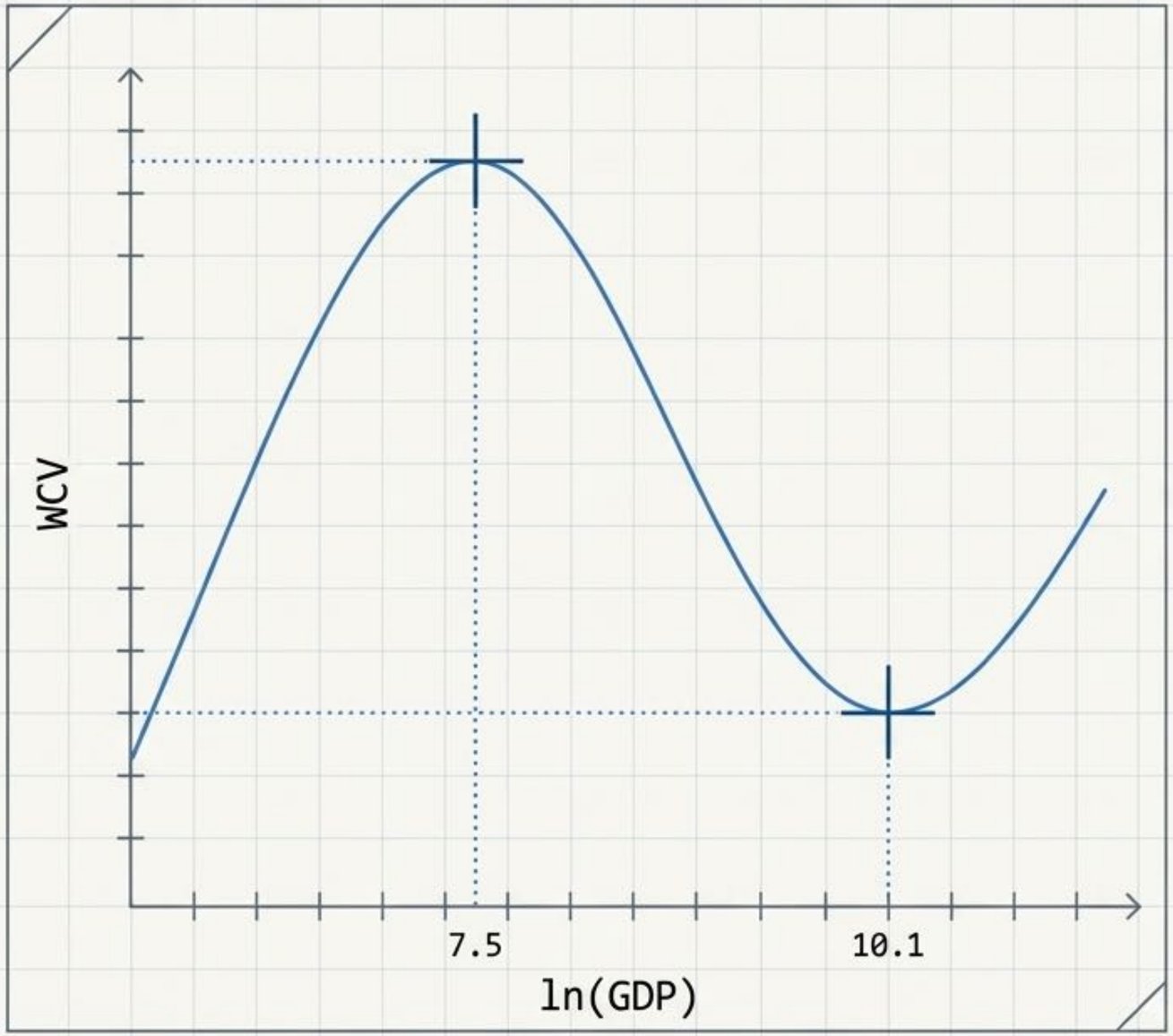
$$WCV_{i,t} = \alpha_i + \beta_1 Y_{i,t} + \beta_2 Y_{i,t}^2 + \beta_3 Y_{i,t}^3 + \sum \gamma_m X_{m,i,t} + \mu_t + \varepsilon_{i,t}$$

Interpretation Key	Interpretation Key	Interpretation Key
$\beta_1 > 0$: Initial divergence (Industrialization)	$\beta_2 < 0$: Convergence phase (Neoclassical adjustments)	$\beta_3 > 0$: Tertiary divergence (Modern service agglomeration)

Control Variables (X_m): Fixed effects (α_i, μ_t), spatial units, land area, ethnic fractionalization, trade/GDP, and urbanization.

NotebookLM

The Calculus of Turning Points



Methodology Callout:

Finding the Extrema:

Taking the first derivative with respect to Income (Y) and setting to zero:

$$\frac{\partial WCV}{\partial Y} = \beta_1 + 2\beta_2 Y + 3\beta_3 Y^2 = 0$$

Interpretation Key

Based on Table 2, Column 5 Estimates:

- **Local Maximum (Peak Inequality):** Occurs at ln(GDP) = 7.5 (approx. $2,000 GDP p.c.).
- **Local Minimum (Convergence Floor):** Occurs at ln(GDP) = 10.1 (approx. $24,000 GDP p.c.).

Beyond this $24k threshold, tertiarization drives a renewed divergence.

NotebookLM

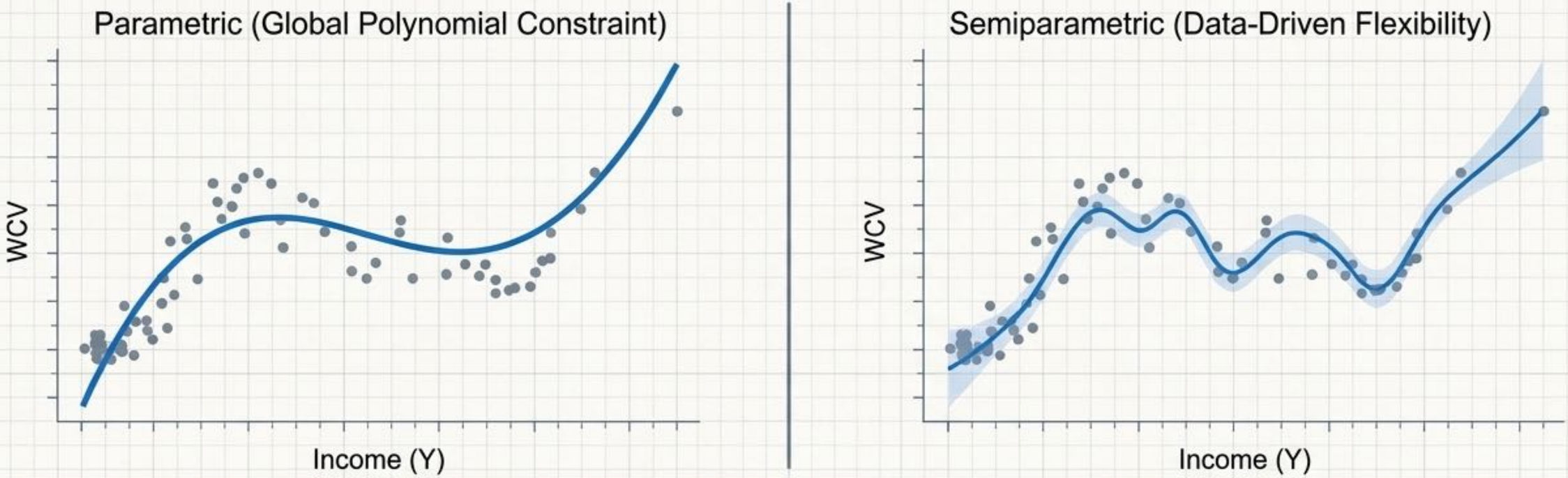
Parametric Results Diagnostics

Result Synthesis Matrix			
Coefficient	Cross-Section (Between Country)	Panel Annual	Panel 5-Year Average
beta_1 (Y)	✓	✓	✓
beta_2 (Y^2)	✓	✓	✓
beta_3 (Y^3)	✓	✗	✗

Interpretation Key	
The Inverted-U Holds	The Cubic Tail is Context-Dependent
beta_1 and beta_2 are statistically significant with expected signs across all model specifications. The Kuznets mechanism is robust in both long-run and short-run dynamics.	beta_3 is only positive and significant in the cross-section. Panel data (within-country short-run) fails to capture the long-run 'tertiarization' tail, particularly when income is log-transformed.

NotebookLM

The Semiparametric Rationale: Parameter Heterogeneity



Methodology Callout:

The Limits of Polynomials: Durlauf (2001) argues that forcing a global polynomial function places too strong a restriction on growth empirics.

Interpretation Key:

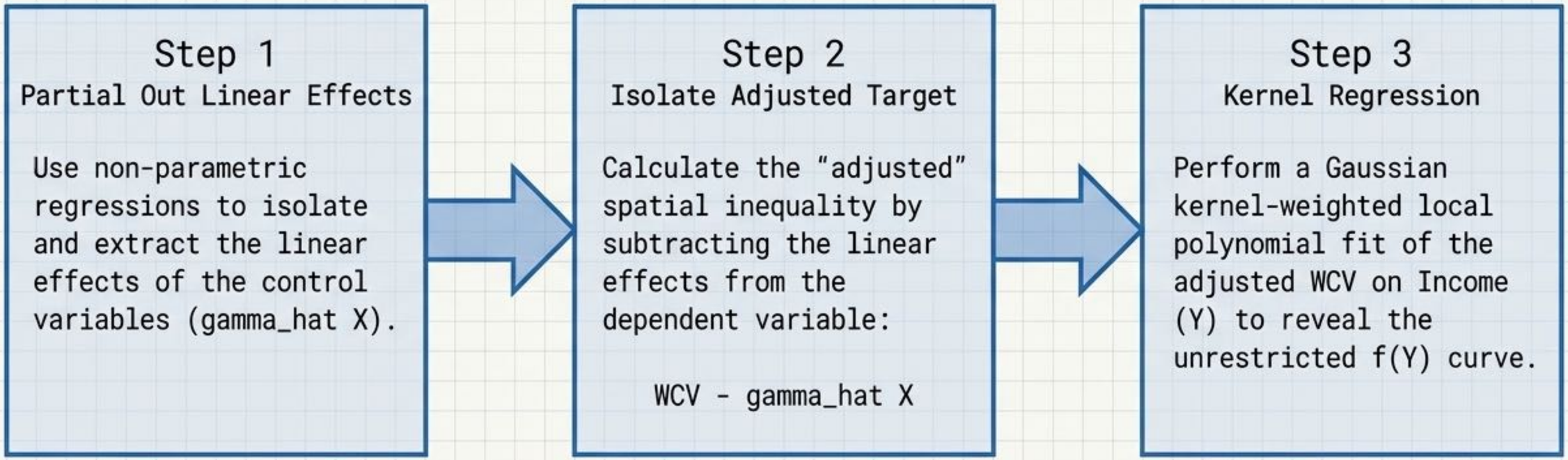
The Semiparametric Solution:

$$WCV = \alpha + f(Y) + \gamma X + \epsilon$$

This allows the linear controls (γX) to remain parametric, while the income effect $f(Y)$ is estimated non-parametrically, letting the data determine the functional form without pre-imposed algebraic constraints.

NotebookLM

Process Flow: The Cross-Sectional Semiparametric Estimator



Methodology derived from the Robinson (1988) estimator.

NotebookLM

Process Flow: The Panel Data Semiparametric Estimator

Step 1: First Differences
Remove unobservable country fixed effects by taking the first difference of the equation.

Step 2: Spline OLS
Conduct OLS regressions using spline functions to approximate the non-linear elements.

Step 3: Fixed Effect Recovery
Calculate and re-insert the specific fixed effects based on the estimated parameters from Step 2.

Step 4: Error Component Smoothing
Estimate f(Y) via an error component model utilizing B-splines (order k=4).

In Practice Panel

Baltagi & Li (2002) Context:

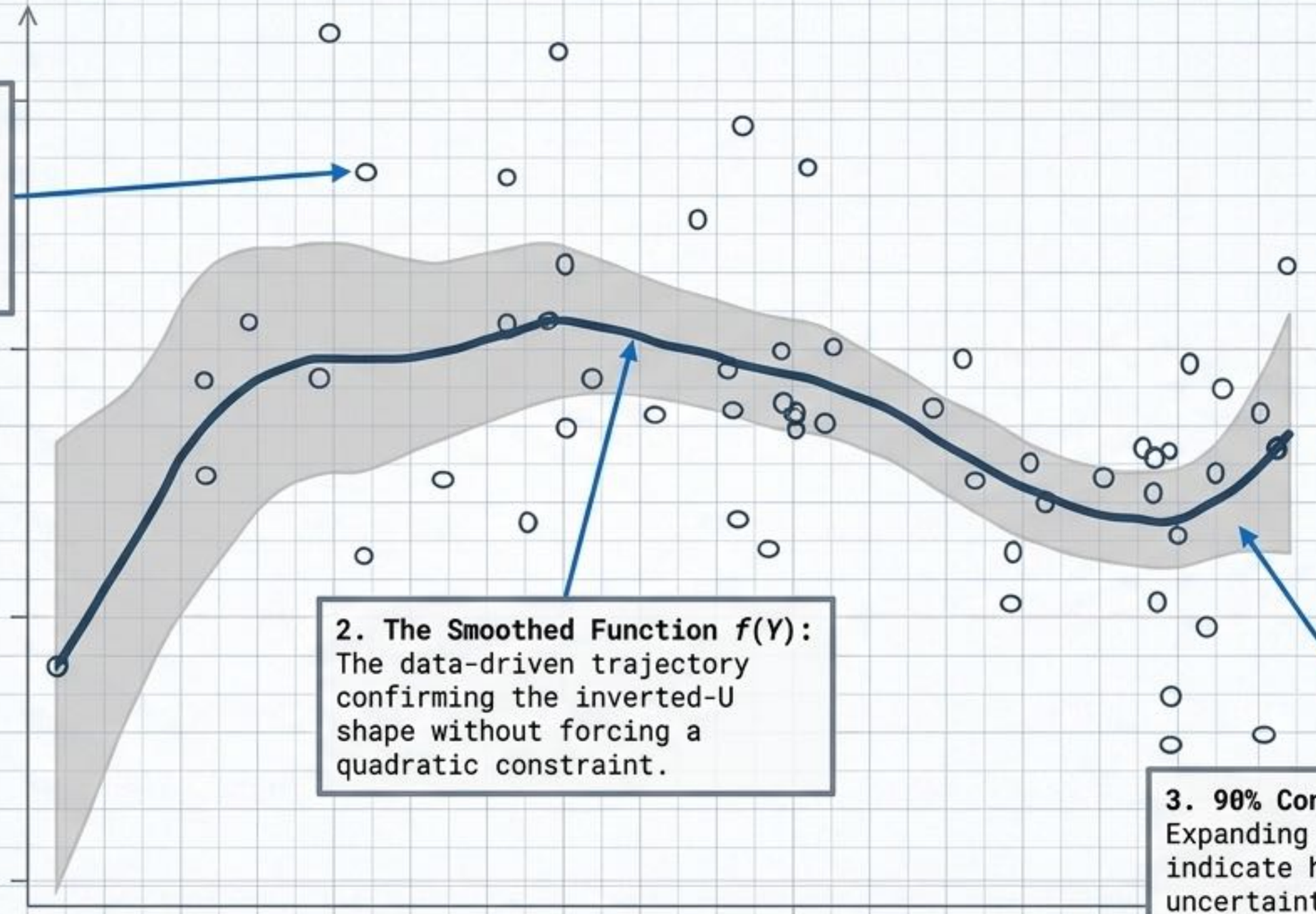
Baltagi & Li (2002)

Panel data requires specialized non-parametric handling because traditional fixed-effects models force a singular unobservable constant on each country.

This 4-step process circumvents the limitation, preserving parameter heterogeneity across time.

NotebookLM

Anatomy of a Partial Fit Graph



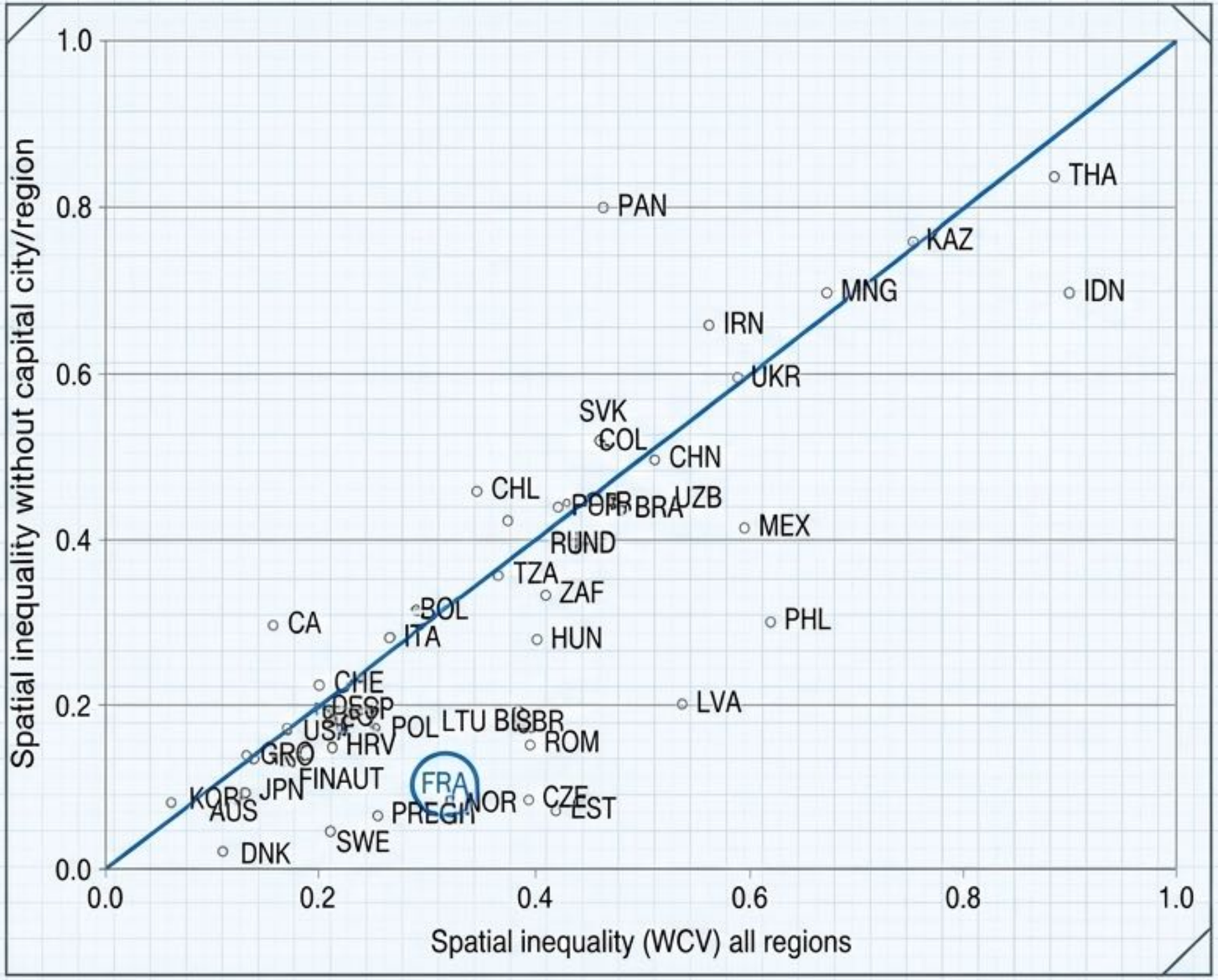
1. Centered Partial Residuals: These points represent regional inequality after the linear effects of controls (like urbanization and trade) have been partialed out.

2. The Smoothed Function *f*(Y): The data-driven trajectory confirming the inverted-U shape without forcing a quadratic constraint.

3. 90% Confidence Bands: Expanding bands at the tails indicate higher statistical uncertainty where data points (extreme low or high incomes) are sparse.

NotebookLM

Robustness Check: The Capital City Distortion



In Practice Panel

The Agglomeration Effect (Ades & Glaeser, 1995)

Centralized political institutions artificially concentrate resources in capital cities, creating massive outliers.

Interpretation Key

Reading the Bisection Graph:

- Below the line: Inequality is higher when the capital is included.
- Above the line: Inequality is higher without the capital.

The French Distortion:

Including Paris spikes French inequality (WCV = 0.29). Excluding it reveals a highly homogenous countryside (WCV = 0.08).

The Core Result:

The overarching inverted-U hypothesis survives even when these dominant capital outliers are excluded.

NotebookLM

Synthesis: Sectoral Shifts as the Driving Mechanism

Methodology Callout

The Proxy Problem:
Income growth (ln GDP) is merely a proxy. Replacing it with Sectoral Data (Non-agricultural GVA / GDP) directly tests the Williamson mechanism.

Mechanism Matrix

Interpretation Key

Final Validation:
The regression confirms the precise transmission channel. The empirical blueprint validates the structural theory: the inverted-U is caused by the mechanics of sectoral shifts, not just aggregate wealth creation.

NotebookLM