The spatial extent of growth spillovers

Rouven Beiner

February 7, 2025

Abstract

This study examines how far gross regional product (GRP) spills over from one location to another. I apply the market price test by [Stigler and Sherwin, 1983] to regional growth and extend it to explicitly account for distance decay, adjust for factors like population and find how far growth spills over. I find the estimated spillover extent to be approximately 600 km, suggesting that economic growth is interdependent between nearly all regions within Germany.

Contents

1 Introduction and Motivation

Understanding the reach of growth spillovers and the extent of a common market between regions provides significant advantages for policy, business, and research. From a policy standpoint, knowledge is crucial for evaluating potential spillover effects and externalities. For businesses, knowing the market boundaries is fundamental for pricing strategies and for understanding the competitive landscape relevant to new market entrants. In research, it is essential for accurately modeling economic systems. To design robust quasi-experiments, researchers must either select relatively independent regions or appropriately adjust for spillover effects. Moreover, knowing the market's extent is particularly relevant for agent-based models, where the interactions between agents within and across regions need to be properly understood and modeled.

In this study, I analyze the spatial economic interdependencies of German regions using 22.5 km × 22.5 km grid cells and gross regional product (GRP) per capita data from 2000 to 2020. I examine correlations between regions for long-term growth paths (GRP levels) and short-term fluctuations (first differences in GRP) to derive a distance decay curve, showing how correlations vary with distance. The concept of using price correlations as an indicator for the extent of a common market was originally proposed by [Stigler and Sherwin, 1983]. Growth correlations can indicate if a common market between regions exists when prices (CPI) are approximated by GRP ($GRP = CPI \times Q$).

However, [Stigler and Sherwin, 1983] left open an important question: there is no universally agreed-upon correlation cutoff (e.g., 0.8, 0.5) to signal a common market. I argue that as long as distance continues to affect correlation, spillovers exist. Only when further distance no longer affects correlation, the hypothesis of spillovers is rejected. To this end, I employ a bounds testing procedure to determine whether the correlation-distance curve (which decreases with distance) has effectively converged to a baseline correlation level. This baseline likely arises from shared federal or national policies that correlate regional economies regardless of distance.

2 Testing for a common market

The dynamics of market behavior and equilibrium price formation across various market structures—from efficient to inefficient markets, as well as monopolistic and perfectly competitive settings—are well established in economics. Much of this knowledge can be translated to the study of growth spillovers.

In classical theory, a unifying characteristic of a market is price uniformity, emerging because suppliers aim to maximize profits via price competition. As Cournot stated, "A market for a good is the area within which the price of a good tends to uniformity." Similarly, [Marshall, 2013] noted, "The more nearly perfect a market is, the stronger is the tendency for the same price to be paid for the same thing at the same time in all parts of the market." The spatial extent of a market is defined by the range over which prices co-move. Analogously, growth spillovers between two regions can be measured by how their GRP moves together.

A seminal and influential study on the extent of the market was conducted by [Stigler and Sherwin, 1983], who proposed to test whether a market exists between two locations by analyzing the similarity of price movements. Earlier, [Horowitz, 1981] also suggested the use of price data in market definition. According to [Stigler and Sherwin, 1983], no adjustment for transportation costs is necessary, as these costs are generally either minor or stable contributors to price differences. They initially analyze price levels but note that spurious correlations may arise. To address this, they recommend testing correlations in changes of prices (first differences). Finding significant correlations in price changes is often more challenging as competition primarily equalizes price levels across a market and only indirectly harmonizes price changes. As noted by [Horowitz, 1981], irrelevant short-term coincidences may occur, but in the long run, price levels within the same market cannot deviate beyond transportation costs.

Furthermore, correlations in price changes depend on the observation frequency. For example, analyzing daily price changes for bread may reveal no correlation, as bakeries rarely adjust prices daily. By contrast, correlations in bread price changes may become evident at lower frequencies, such as monthly observations, or of course at price levels. If autocorrelation is present, it can bias correlation estimates, and autocorrelation-adjusted correlations should be calculated.

3 Determinants of growth Spillovers

A large body of empirical research has identified significant growth spillovers among neighboring countries or regions. For instance, [Moreno and Trehan, 1997] analyze the relationship between a country's growth rate and the economic growth of neighboring countries, finding evidence of spillovers between geographically close countries. Similarly, [Elhorst et al., 2024] examined the spatial reach of growth spillovers across 266 NUTS-2 EU regions from 2000 to 2018, finding distances ranging from 700 to over 1500 km, depending on growth determinants.

Strong growth spillovers between two nearby regions can arise from several underlying factors. First, strong trade linkages—typically more pronounced between neighboring regions— support growth spillovers and indicate the presence of common goods markets. Additional contributing factors include strong investment relationships, which suggest a shared capital market, and Marshallian externalities, which imply specialized labor clusters or shared supplier networks within a region. While investment relationships (capital markets) are generally non-localized, bilateral trade (goods markets) and Marshallian externalities tend to be geographically localized. Therefore, I expect trade linkages and Marshallian externalities to be the primary drivers of growth spillovers between nearby regions.

The strength of bilateral trade between two regions typically depends on factors such as geographical distance, cultural proximity, and trade agreements, as initially explained by the gravity model proposed by [Tinbergen, 1963]. However this relationship is interrupted when trade barriers exist. For example, in the 1970s and 1980s Berlin, despite its geographical and cultural proximity, two markets existed and no trade occurs between the eastern and the western part. It is reasonable to expect that closer trade ties would result in more correlated GRP, as stronger trade linkages create greater interdependencies on the demand and supply sides across regions. Empirical studies ([Frankel and Rose, 1998], [Clark and van Wincoop, 2001]) confirm that. However, in theory, there is an opposing effect of trade linkages on GRP synchronization. According to classical Ricardian theory, increased trade leads to greater specialization, making it less likely that sector-specific shocks in one economy will affect others. The specialization of trading countries results from comparative cost advantages, which occur when regions possess different characteristics. This is especially true for regions separated by large distances, where cultural, ecological, and other factors contribute to differing specializations. [McCallum, 1995] found that U.S. states with strong trade linkages are more specialized than comparable independent nations. But if countries are similar in their characteristics, intra-industry trade based on increasing returns to scale and differentiation will be dominant and strong spillovers exist. In this analysis, small German regions with similar structural characteristics are considered. I can expect strong spillover effects between neighboring regions because of both strong trade linkages and similar industry specializations.

Another mechanism for growth spillover and business cycle synchronisation are financial linkages and foreign direct investments as found by [Dabla-Norris et al., 2010] and [Hsu et al., 2011]. However this mechanism is less dependend on distance, and is therefore not an appropriate explanation for the occurrence of growth spillovers of neighborign regions.

Furthermore, growth spillovers between regions in close proximity can arise from various non-priced externalities initially proposed by Alfred Marshall ([Marshall, 1920]). Within Marshallian externalities, [Fujita et al., 1999] distinguishes between knowledge spillovers (technological externalities) and market-based advantages (pecuniary externalities).

Pecuniary externalities can occur when firms in neighboring regions benefit from shared access to a specialized labor market, a network of input suppliers, or market size effects. For example, it may be profitable for some suppliers to locate in a neighboring region with a lower degree of agglomeration while still taking advantage of proximity to a saturated region. For pecuniary externalities, geographic proximity is particularly important ([Puga and Venables, 1996]), and they are likely a significant factor determining the strength of GRP correlation between regions.

Technological knowledge is theoretically non-localized and transferable globally. In Romer's (1986) influential model, knowledge remains non-localized: a firm's knowledge acts as a public good that other firms can freely access, and once discovered, it spreads instantaneously across the entire economy.

However, recent research has shown that local technology spillovers play a significant role in growth models and have been confirmed empirically. [Chua, 1993], for instance, incorporate neighboring countries' physical and human capital, while [Ertur and Koch, 2007] demonstrate that technological spillovers diminish with distance. Empirical studies, such as [Bottazzi and Peri, 2003], [Funke and Niebuhr, 2005], and [Rodríguez-Pose and Crescenzi, 2008], confirm the spatial limits of R&D spillovers. [Bottazzi and Peri, 2003] find that R&D spillovers are highly localized, typically limited to a radius of 300 km. They highlight that while patents function as a "public good"—widely accessible and internationally available—another component of knowledge is tied to the experience of scientists and is "attached" to individuals. This localized knowledge spreads primarily through personal contacts and face-to-face interactions, effectively making it a "local public good." [Rodríguez-Pose and Crescenzi, 2008] further highlight proximity's role in knowledge diffusion across the EU regions, with clear distance decay effects.

In conclusion, correlations in growth between neighboring German regions can be attributed to distance and the underlying mechanisms associated with it. Key factors include strong trade linkages and similar sectoral specializations. These sectoral specializations often emerge from local knowledge spillovers and market-based externalities, such as access to specialized labor markets or networks of input suppliers.

4 Data

To estimate distance-decay effects, previous studies often relied on administrative units (e.g., NUTS-2 or NUTS-3), but this approach poses at least two problems. First, these units vary in size, complicating efforts to isolate distance effects. Second, their irregular shapes make distance estimation difficult. I address these issues by using uniform grid cells of equal size.

I rely on the global gridded GRP dataset from [Wang and Sun, 2022], which provides a high spatial resolution of 30 arc-seconds and 0.25 degrees for the years 2000–2020 at annual intervals. This dataset integrates Gross Regional Product (GRP) data for more than 800 provinces (states) in 48 countries, downscaled through Nighttime Light (NTL) imagery and gridded population data. For Germany, the dataset draws on NUTS-2 regions covering 16 states. The authors combine NTL and population data to create fine-gridded GRP estimates. Relying on NTL alone can underestimate urban centers and overestimate rural areas because of image saturation, while population-only approaches assume a uniform GRP per capita within an administrative boundary—an assumption that often proves unrealistic.

To calculate GRP per capita, I use the gridded population data from WorldPop, which covers Germany and its neighboring countries at a $1km \times 1km$ resolution from 2000 to 2020.

I reproject both datasets to EPSG:3857 to ensure spatial consistency. I adopt a target resolution of about 22.5 $km \times 22.5$ km, matching the average cell size of the original 0.25-degree GRP grid in Germany. First, I reproject the GRP data—covering regions inside and outside Germany—to the target coordinate reference system and resolution. Then I resample the reprojected population data to align with the $22.5km \times 22.5km$ grid, applying summation for all projection and resampling steps. Because the population data includes Germany's neighboring countries, grid cells that partially extend beyond Germany's borders remain intact. The appendix (Figures 4a and 5) presents aggregated population data as well as the GRP and population time series for a sample region.

I precisely define the distances between any two cells to construct a distance-decay curve. Because the cells are fixed and square, the distance to the center of the nearest neighboring cell equals 22.5 km. Relying solely on the n-th nearest cells would yield uneven real distances, since diagonal distances exceed horizontal or vertical ones. To create a consistent band of surrounding cells, I apply a circular approximation. For each target radius r, I define the distance band as

$$\mathrm{r} - 0.5 \leq \text{distance} \leq \mathrm{r} + 0.5,$$

where the distance from the origin cell (x, y) to another cell (i, j) is

$$\text{distance} = \sqrt{(x-i)^2 + (y-j)^2}.$$

Figure 4b in the appendix illustrates how cells at different distances are selected. Figure 6 shows the total number of grid-cell pairs at each specified distance across Germany. As the distance among cells grows, the total number of cell pairs for each distance also increases.

5 Descriptive Statistics

Figure 1 displays the distributions of GRP, population, and GRP per capita, averaged annually from 2000 to 2020. I use a logarithmic scale for clearer visualization.

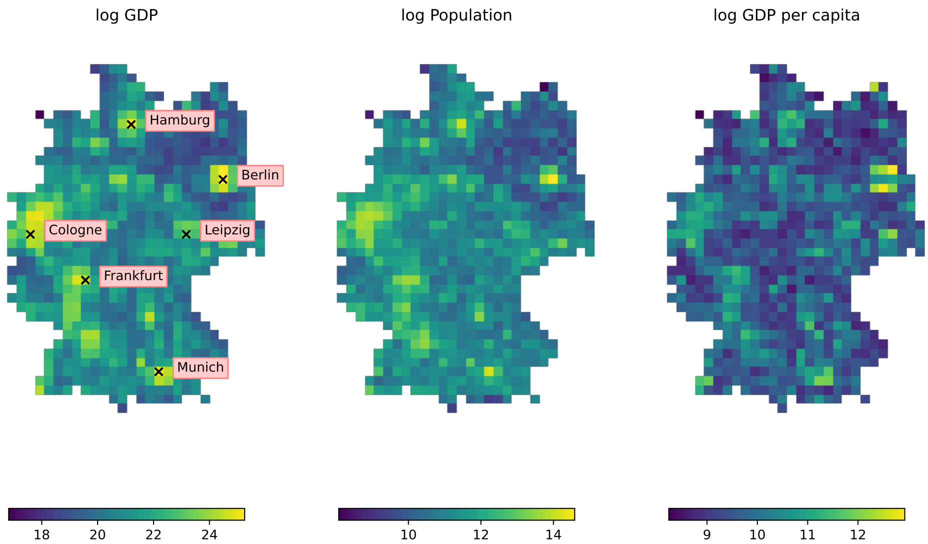


Figure 1: Rasterized and EPSG:3857-projected grid cells over Germany, showing GRP, population, and GRP per capita.

Particularly high GRP per capita values appear around large cities, such as Berlin, Munich, or Dresden. One likely explanation is that firms locate production or logistics facilities on urban outskirts, where they benefit from the city's labor market but also encounter relatively low property costs. At

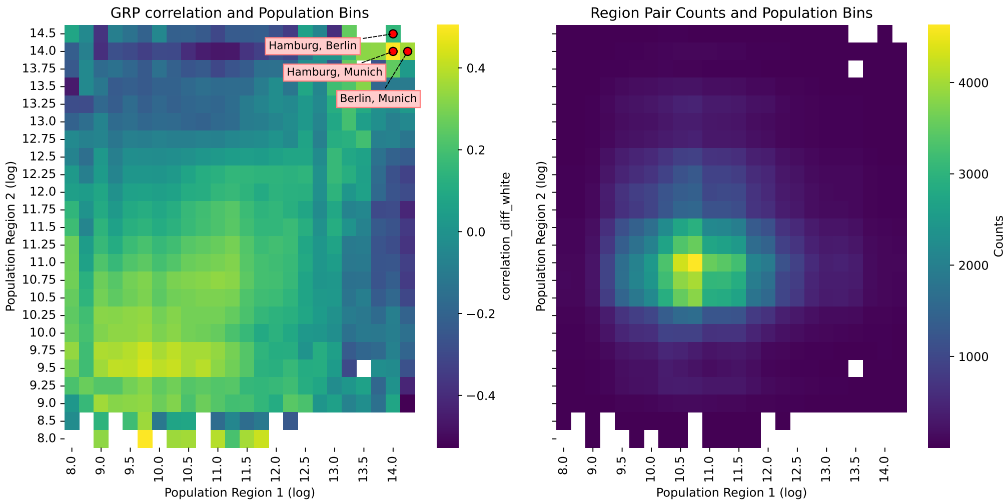


Figure 2: GDP correlation and population of both regions

the same time, population density remains low in these outskirt areas, raising GRP per capita.

Figure 2 shows the relationship of the Populations of a regions pair and their GRP correlation (first differences, whitened). Population is log transformed and clustered into discrete bins. Since the GRP point process is assumed to be spatially inhomogeneous—specifically, first-order non-stationary (i.e., the mean GRP varies by location)—only region pairs with a distance of at least $d \geq 22.5 \times 4$ are considered. This threshold ensures that the correlations in the heatmap are not primarily driven by neighboring regions, which typically exhibit similar population counts and highly correlated GRP values. The right heatmap shows the number of pairs for each population bin, as a robustness check.

Figure 2 illustrates the relationship between the populations of a pair of regions and their GRP correlation (first differences, whitened). I log-transform the population values and cluster them into discrete bins. Because I assume the GRP point process to be spatially inhomogeneous—specifically, first-order non-stationary (i.e., the mean GRP varies by location)—I only consider region pairs with a distance of at least $d \geq 22.5 \times 4$. This threshold ensures that the correlations in the heatmap are not dominated by neighboring regions, which typically share similar population counts and high GRP correlations. The right heatmap shows the number of pairs for each population bin as a robustness check.

The heatmap suggests that regions with similar population counts exhibit a strong positive correlation in GRP, forming a diagonal pattern from bottom-left to top-right. By contrast, as population counts diverge substantially, the correlation between regional GRPs becomes negative. Interestingly, even regions with smaller populations demonstrate common movements in GRP, possibly due to shared structural characteristics or sector specializations. Rural regions may generate significant GRP from agriculture, mid-sized cities from manufacturing, and large urban areas from service industries. A sector-specific shock in one region can therefore affect another region specializing in the same sector. For example, Munich, dominated by startup-driven tech industries might be strongly influenced by changes in investor risk aversion, and similarly specialized regions (Berlin) could exhibit parallel responses.

6 Correlation Analysis

Since there are multiple observations over time, the straightforward temporal correlation between two regions' GRP levels can be estimated. The covariance between two locations ℓ_i, ℓ_j is given by:

$$\hat{\sigma}_{GRP}(\ell_i, \ell_j) = \sum_{t=1}^{T} \left[GRP(\ell_{it}) - \mu_G RP(\ell_i)\right] \left[GRP(\ell_{jt}) - \mu_G RP(\ell_j)\right], \tag{1}$$

and the correlation is given by:

$$\hat{\rho}(\ell_i, \ell_j) = \frac{\hat{\sigma}_{GRP}(\ell_i, \ell_j)}{\sqrt{\hat{\sigma}_{GRP}(\ell_i, \ell_i)}\sqrt{\hat{\sigma}_{GRP}(\ell_j, \ell_j)}}. \tag{2}$$

The distance decay curve is a function of distance d that estimates the mean correlation between any two points ℓ_i and all points ℓ_j that are d distances apart:

$$\hat{\rho}(d) = \frac{\sum_{i=1}^{L} \sum_{j=1}^{L} 1(\|\ell_i - \ell_j\| = d)\hat{\rho}(\ell_i, \ell_j)}{\sum_{i=1}^{L} \sum_{j=1}^{L} 1(\|\ell_i - \ell_j\| = d)}. \tag{3}$$

As suggested by [Stigler and Sherwin, 1983], estimating correlations in first differences, particularly in autocorrelation-adjusted first differences, is a more robust method. The covariance of first differenced GRP between two regions is given by:

$$\hat{\sigma}_{\Delta GRP}(\ell_i, \ell_j) = \sum_{t=1}^{T} \left[\Delta GRP(\ell_{it}) - \mu_{\Delta GRP}(\ell_i)\right] \left[\Delta GRP(\ell_{jt}) - \mu_{\Delta GRP}(\ell_j)\right], \tag{4}$$

and the covariance of autocorrelation-adjusted first differences (white noise errors) between two regions is:

$$\hat{\sigma}_\epsilon(\ell_i, \ell_j) = \sum_{t=1}^{T} \left[\epsilon(\ell_{it}) - \mu_\epsilon(\ell_i)\right]\left[\epsilon(\ell_{jt}) - \mu_\epsilon(\ell_j)\right], \tag{5}$$

where the white noise errors ϵ_{it} are derived from an AR(p) model applied to the first differences:

$$\Delta GRP_{it} = \alpha + \sum_{p=1}^{P} \phi_p \Delta GRP_{i,t-p} + \epsilon_{it}. \tag{6}$$

Here, P is the number of optimal lags chosen by minimizing the Akaike Information Criterion (AIC) with a maximum of two lags. Correlations and distance decay curves are similarly estimated using Equations (2) and (3).

The mean of correlations for specific distances in Equation (3) is estimated with a regressogram including dummy variables for each distance:

$$\hat{\rho}(\ell_i, \ell_j) = \sum_{d=1}^{D} \beta_{dij} D_{dij} + \epsilon_{ij}, \tag{7}$$

where $\sum_{d=1}^{D} \beta_{dij} D_{dij}$ represents the set of distance dummies indicating the distance between cells i, j. The regression function can be expressed as:

$$E[\hat{\rho}(\ell_i, \ell_j) \mid d_1, d_2, \ldots, d_D] = \sum_{d=1}^{D} \beta_{dij} D_{dij}, \tag{8}$$

Since for a given observation only one $d_k = 1$ while all others are zero, the conditional expectation simplifies to:

$$E[\hat{\rho}(\ell_i, \ell_j) \mid d_k = 1] = \beta_k. \tag{9}$$

Thus, the OLS coefficient β_k for dummy d_k corresponds to the mean correlation $\hat{\rho}(k)$ for distance k, as it captures the average value within the group defined by $d_k = 1$.

As a baseline method, Equation (7) does not account for other factors such as population. To address population-related influences, regressions in levels, first differences, and AC-adjusted first differences are estimated with control variables:

$$\hat{\rho}(\ell_i, \ell_j) = \sum_{d=1}^{D} \beta_{dij} D_{dij} + \sum_{d=1}^{D} \gamma_{dij} D_{dij} \times \mathrm{Pop}_{dij} + \epsilon_{ij}, \tag{10}$$

where $\hat{\rho}(\ell_i, \ell_j)$ is the correlation of GRP between regions i, j, and $\sum_{d=1}^{D} \gamma_{dij} D_{dij} \times \text{Pop}_{dij}$ include the average population of both regions interacted with each distance dummy. To find the marginal effect (ME) of distance d from the control model, I take the partial derivative of $\hat{\rho}(\ell_i, \ell_j)$ with respect to D_{dij}:

$$\frac{\partial \hat{\rho}(\ell_i, \ell_j)}{\partial D_{dij}} = \beta_{dij} + \gamma_{dij} \times \text{Pop}_{dij}. \tag{11}$$

The ME of distance d depends on β_{dij} and on the interaction term $\gamma_{dij} \times \text{Pop}_{dij}$, introducing heterogeneity in the ME. I average the ME for each distance to get the distance decay curve.

7 Results and Discussion

Because of federal policies in Germany, I expect a baseline correlation among regions, regardless of geographical distance or shared market structures. The distance decay curve exhibits what I define as the inter-regional base correlation —the correlation level to which the curve converges asymptotically as distance d approaches infinity. I interpret the extent of the market as the distance where the distance decay curve converges to this inter-regional base correlation.

Figure 3 shows the distribution of the correlation coefficients (left), along with the slopes of the smoothed correlation coefficients without control variables (middle) and with control variables (right). I estimate these from GRP levels (first row), GRP first differences (second row), and autocorrelation-adjusted GRP first differences (last row).

The boxplots (left subfigures) represent the distribution of correlations at each distance —i.e., the distance decay curves. The box spans from the first quartile (Q1) to the third quartile (Q3), with a line indicating the median. Whiskers extend 1.5 times the interquartile range (IQR), and I plot outliers as individual points beyond the whiskers. Each point corresponds to a pair of regions at a specific distance and the estimated correlation coefficient between their GRP time series.

The middle and left subplots show the slopes of the smoothed correlation coefficients along with bootstrapped percentiles (5% and 95%) for each specification (levels, first differences, autocorrelation-adjusted first differences) without control variables (middle) and with control variables (right). The correlation coefficients represent the distance dummy coefficients of the regressograms, smoothed over time using a LOWESS with a bandwidth of 7. I estimate both bounds using the bounds test procedure explained in Section B with a significance level of 5%.

The effect of distance on the correlation of short-term growth shocks (GRP differences) and long-term growth trends (GRP levels) is similar. All three distance decay curves exhibit a downward trend, with slopes converging toward zero at around 600 km. I summarize the estimated market-extent bounds in Table 2. I consider the positive slope beyond 600 km as a fluctuation driven by randomness and a low number of observations at these distances (see Figure 6). These extended distances represent straight air distances, and they align with findings from similar studies. For illustration, 600 km is approximately the distance between Hamburg and Munich or roughly the length of Germany's west–east diagonal.

The convergence levels of short-term growth shocks (GRP differences) and long-term growth trends (GRP levels) differ. The GRP-level correlations (top row) start near 0.75 and stabilize at about 0.4. As expected, the correlations for both first-differenced series appear much smaller: they begin around 0.6 and converge close to zero with increasing distance. These convergence levels of GRP levels and GRP differences indicate an inter-regional base correlation above zero in the long run. Federal policy exerts a structural long-run effect on German regions but does not affect short-run growth among them.

	Lower bound (km.)	**Upperbound (km.)**
No control variables		
Levels	540	540
Differences	630	630
AC. adj. Differences	630	630
With control variables		
Levels	540	562.5
Differences	630	630
AC. adj. Differences	630	630
Average	600	603.75

Table 1: Estimated Market Extent bounds.

I find that nearly all German regions exhibit economic interdependence, except for those separated by distances greater than 600 km. From a policy perspective, these results suggest that regions should consider not only their immediate neighbors but also more distant ones. These findings have important implications for economic research. First, theoretical growth models should incorporate the idea that economic growth extends beyond direct neighbors. Additionally, the strong interconnectivity among German regions challenges assumptions of regional independence in quasi-experiments, indicating that researchers should treat distance as a crucial factor. The AC-

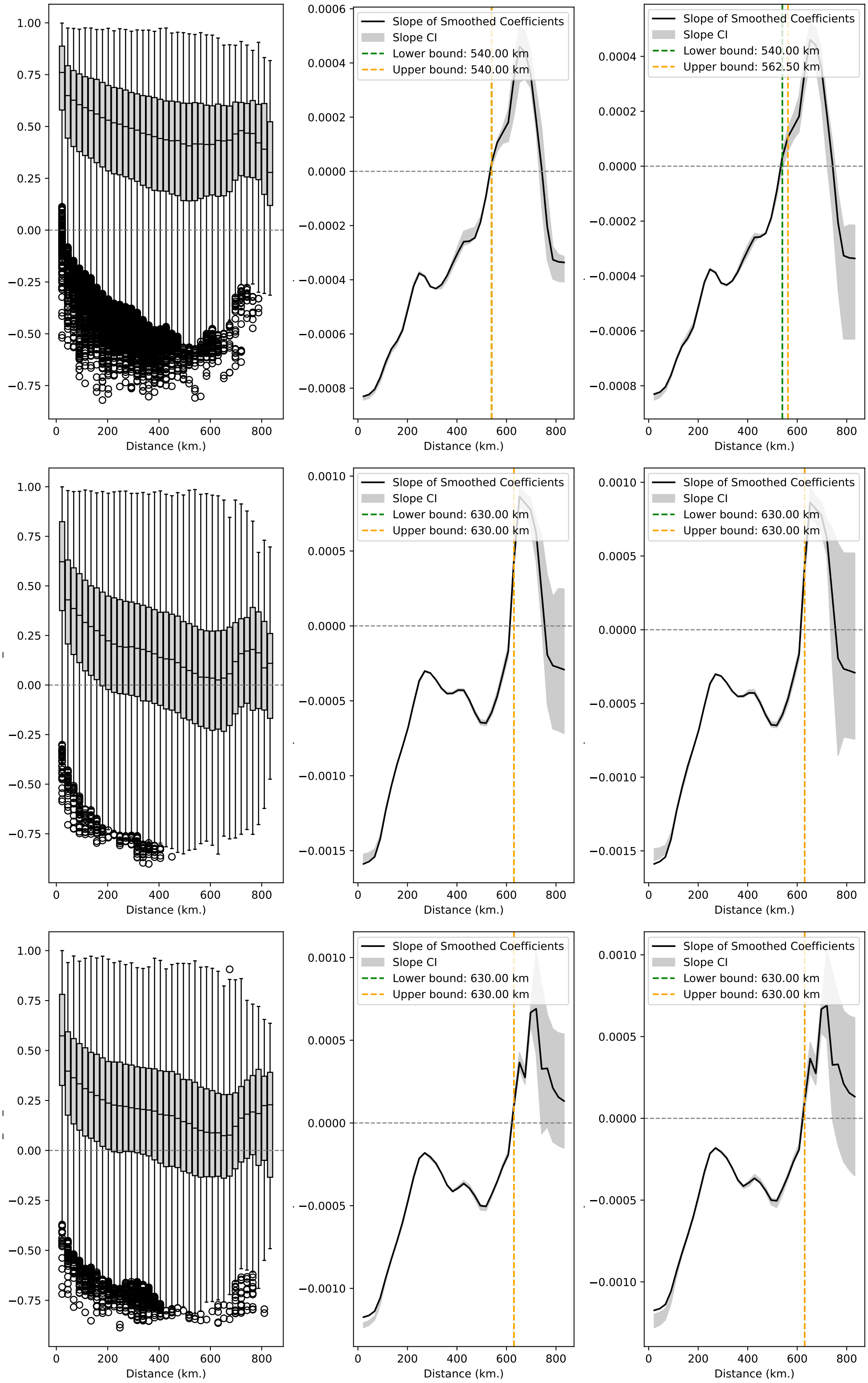
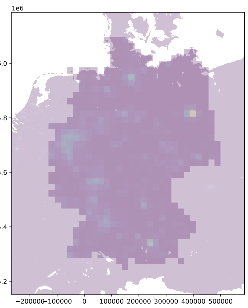


Figure 3: Correlations, Slopes and estimated bounds

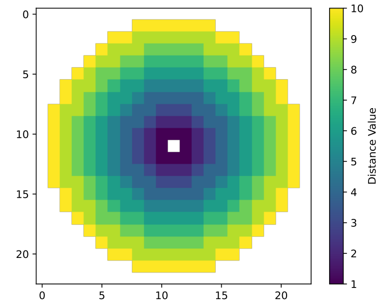
Top row: GRP Levels, Middle row: GRP Differences, Bottom row: AC. adj. GRP Differences. Columns represent correlations (left), slopes without control (middle), and slopes with control (right).

adjusted method in first differences controls for lagged effects of GRP within the same region but cannot capture cross-sectional lagged effects. Future studies could extend the analysis by incorporating cross-sectional lag effects, thereby modeling economic spillovers among German regions across both time and space.

A Additional data visualization



(a) Population projection and aggregation



(b) Cell distance estimation

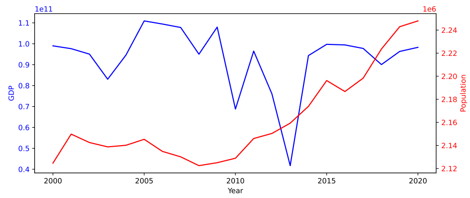


Figure 5: Population and GDP time series for a sample region (Berlin)

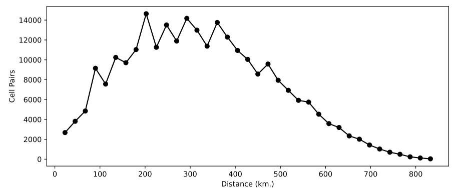


Figure 6: Counts of cell pairs for each distance

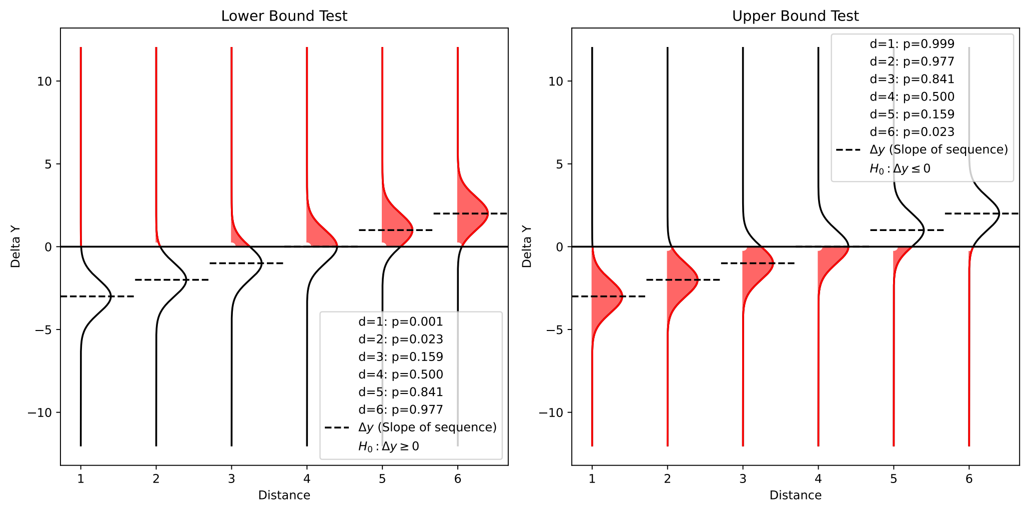


Figure 7: Bounds test (theory)

Illustration of the lower bound (left) and upper bound (right) testing procedure in theory.

B Bounds Testing approach

In the following I generalize the approach to estimate the market extend from the distance decay curve. The correlation-distance curve is a sequence (a_n) that converges to a finite limit (the inter-regional base correlation) L as $n \to \infty$. Formally, a sequence (a_n) is defined as: $a : \mathbb{N} \to S, \quad n \mapsto a_n$ where a_n represents the n-th term of the sequence and S is the set of real numbers $\mathbb{R}$. Convergence is given at $n = \inf$ where the slope of the sequence is equal zero:

$$|a_{n+1} - a_n| = \Delta a_n = 0$$

. There are several techniques to determine the point at which a sequence can be considered to have practically converged. The term *practically* is used here because, strictly speaking, a sequence converges only asymptotically. The usual criterion for sequences is to check whether the term differences (the slope) are smaller than a specific threshold

$$|a_{n+1} - a_n| = |\Delta a_n| < \epsilon$$

for sufficiently large n. However, this requires specifying ϵ, introducing an additional degree of subjectivity in the analysis.

Since this analysis involves a data distribution rather than a single deterministic sequence, the distribution of the sequence at any specific n can be used to establish two new criteria for convergence of a decreasing sequence.

The series is considered to have practically converged at n if the probability mass of the sequence slopes Δa_n above zero is greater than a given proportion p (e.g., $p = 0.05$). This can be mathematically expressed as:

$$P(\Delta a_n > 0) > p \quad \Leftrightarrow \quad F_{\Delta a_n}(0) < 1 - p$$

This criterion will be referred to as the *lower bound*, as it serves as an initial progressive indicator of convergence, with a tendency to favor the detection of convergence.

As a more conservative indicator, referred to as the *upper bound*, the following criterion is proposed: The series is considered to have converged if the probability mass of Δa_n below zero is smaller than a given proportion p. Mathematically, this can be written as:

$$P(\Delta a_n < 0) < p \quad \Leftrightarrow \quad F_{\Delta a_n}(0) < p$$

These two criteria together provide a practical bounds testing framework for assessing convergence based on the distribution of the sequence slopes. The upper bound approach implicitly assumes that the smoothed slope transitions from negative to positive at some point. Due to the stochastic nature of the data, the smoothed slope is likely to fluctuate around zero at its convergence level. This phenomenon is well-documented in Slutzky's seminal work, "The Summation of Random Causes as the Source of Cyclic Processes" ([Slutzky, 1937]) published in Econometrica in 1937. If the slope turns positive, it can be concluded that the convergence level has been reached at n. This lower and upper bounds represents the **minimum extent of the market** and the **maximum extent of the market**.

If the convergence level L would be known or observable, $a_n = L$ can be tested directly. An alternative procedure outlined in the appendix tries to identify L and test if $a_n = L$.

For a sequence y_d that decreases towards its convergence level, the bounds testing procedure consists of the following steps (To be conducted separately for each distance $d \in D$):

1. **Lower Bound Testing:**

   $$H_0 : \Delta y_d \geq 0 \quad \text{(Convergence)}, \quad H_1 : \Delta y_d < 0 \quad \text{(Non-convergence)}$$

   The null hypothesis is that the slope at d is greater or equals zero (the series has converged), while the alternative is that the slope is smaller then zero (series has not converged). If the proportion of $f_{\Delta y_d}$ above zero is more than a given p-value, the hypothesis of convergence ($\Delta y_d \geq 0$) is rejected.

2. **Upper Bound Testing:**

$$H_0 : \Delta y_d \leq 0 \quad \text{(Non-convergence)}, \quad H_1 : \Delta y_d > 0 \quad \text{(Convergence)}$$

The null hypothesis is that the slope at d has not changed its sign, indicating that the series has not converged, while the alternative hypothesis states that the slope has become positive, indicating convergence. If the proportion of $f_{\Delta y_d}$ below zero is less than a given significance level (p-value), the hypothesis of non-convergence ($\Delta y_d < 0$) is rejected.

Figure 8 visualizes the testing procedure for the lower and the upper bound applied on GDP data. The theoretical testing procedure is visualized in the appendix (Figure 7).

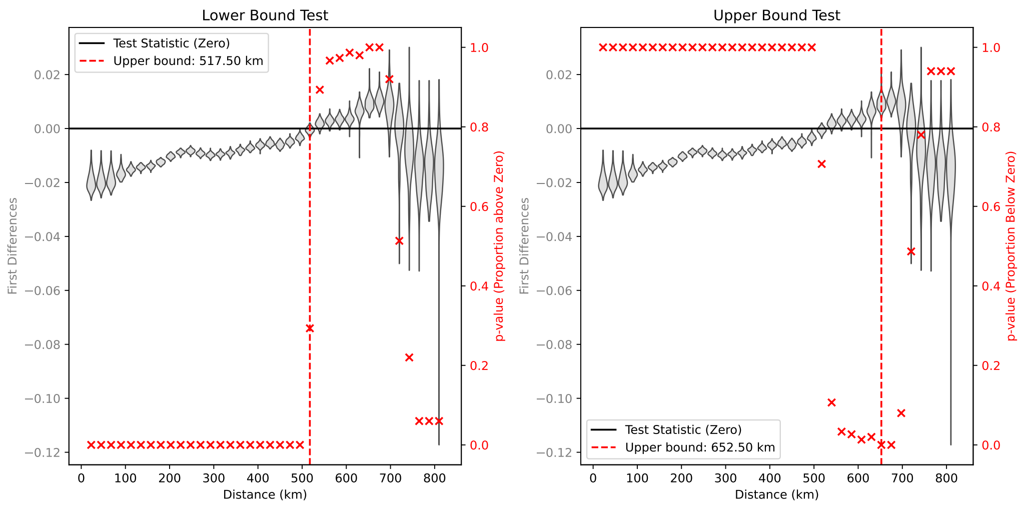


Figure 8: Bounds test

Illustration of the lower bound (left) and upper bound (right) testing procedure applied on data.

For both the lower and upper bounds tests, confidence intervals are constructed by inverting around Δy_d based on the bootstrapped distribution $f_{\Delta y_d}$. This bootstrapped distribution serves as the null distribution. Both tests are one-sided, and for a significance level of 95%, the 95% or 5% percentile is used to determine the bounds. The occurrence of a zero or positive slope within the observed data is not guaranteed. In this case no bounds can be estimated.

C Alternative Bounds Test

The alternative bounds test estimates the convergence level by first identifying the distance where the slope shifts from negative to positive, corresponding to the upper bounds test in Section B, which defines the **maximum extent of the market**. The **minimum extent of the market** is determined as the distance where the correlation first intersects the inter-regional base correlation. The base correlation is estimated as the weighted mean of the distance decay curve values beyond the maximum extent of the market, with weights based on the number of observations at each distance.

For a series that is decreasing towards its convergence level, the alternative bounds testing procedure involves the following steps:

1. **Upper Bound Testing:**

   $$H_0 : \Delta y_d \leq 0 \quad \text{(Non-convergence)}, \quad H_1 : \Delta y_d > 0 \quad \text{(Convergence)}$$

   The null hypothesis is that the slope at d has not changed its sign, indicating that the series has not converged, while the alternative hypothesis states that the slope has become positive, indicating convergence. If the proportion of $f_{\Delta y_d}$ below zero is less than a given significance level (p-value), the hypothesis of non-convergence ($\Delta y_d < 0$) is rejected.

2. **Estimation of the Convergence Level $\tilde{y}$:** The convergence level $\tilde{y}$ is estimated as the mean of the distance decay curve for distances greater than the highest identified bound (i.e. the upper bound if identified). To account for variations in the number of observations across distances, a weighted mean should be applied, ensuring an unbiased estimation of $\tilde{y}$: $\tilde{y} = \frac{\sum_{d \in D} w_d \cdot y_d}{\sum_{d \in D} w_d}$ where w_d is a weight based on observation counts.

3. **Lower Bound Testing:**

   $$H_0 : y_d \leq \tilde{y} \quad \text{(Convergence)}, \quad H_1 : y_d > \tilde{y} \quad \text{(Non-convergence)}$$

   $$E[y_d] \sim f_{y_d} \quad \text{(Test statistic)}$$

   The null hypothesis states that the series has converged at distance d if $y_d \leq \tilde{y}$, while the alternative suggests that the series has not yet converged, i.e., $y_d > \tilde{y}$. The null distribution $f_{\tilde{y}}$ is derived from bootstrapping the estimated values $\tilde{y_d}$. If the proportion of the null distribution exceeding the test statistic is below the chosen significance level α, the null hypothesis is rejected, indicating non-convergence at that distance. The lower bound is defined as the first distance where the p-value exceeds the chosen significance level.

Upper bound and Lower bound testing using this alternative method is illustrated in Figure 9 and Figure 10.

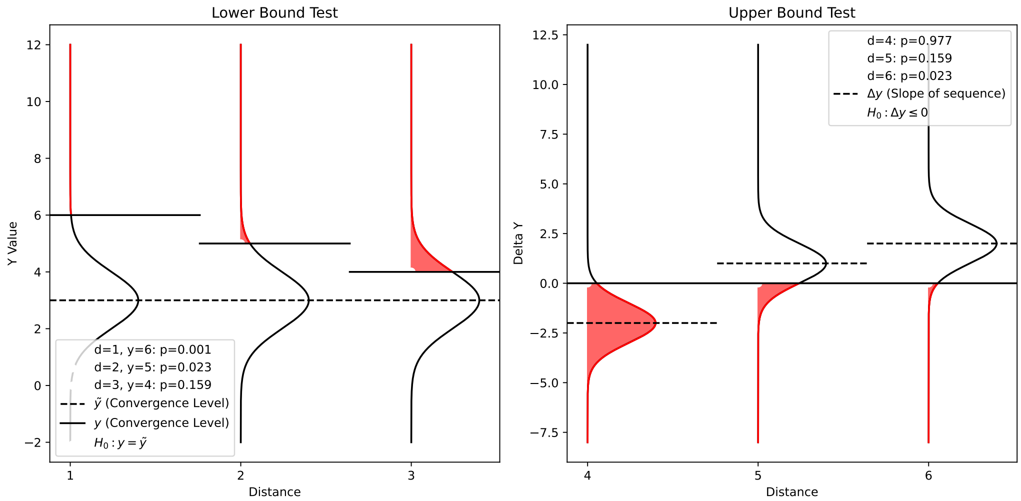


Figure 9: Alternative bounds test (theory)

Illustration of the lower bound (left) and upper bound (right) testing procedure in theory. For the lower bound, the Null distribution is the distribution of the convergence level, for the upper bound, The Null distribution is the distribution of the Sequences' slope. The red filled area represents the p-value.

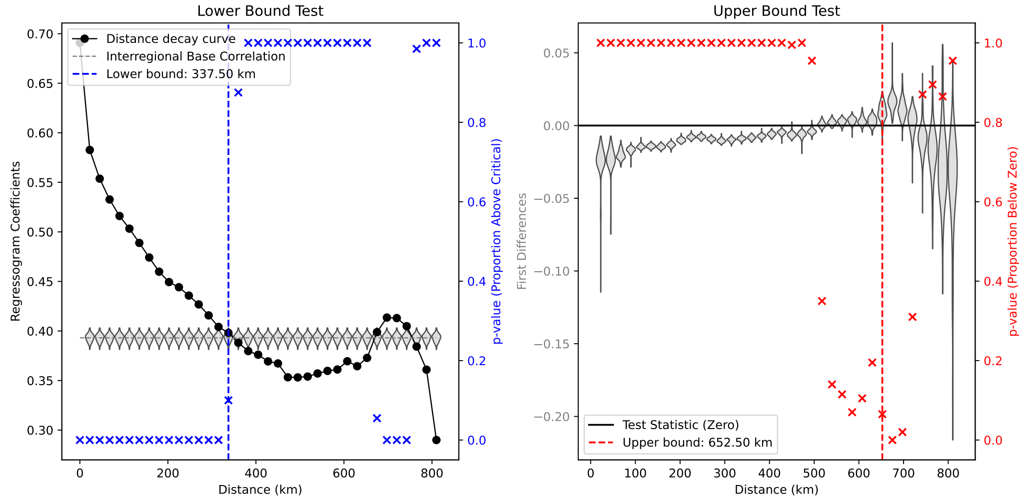


Figure 10: Alternative bounds test (real data)

Illustration of the lower bound (left) and upper bound (right) testing procedure applied on data. The violin plots show the null distributions $(f_{\tilde{y}}, f_{\Delta y_d})$ for each distance. The proportion of the distribution above $E[y_d]$ (left) and below zero (right) are the respective p-values, shown as red x's (right axis scale).

References

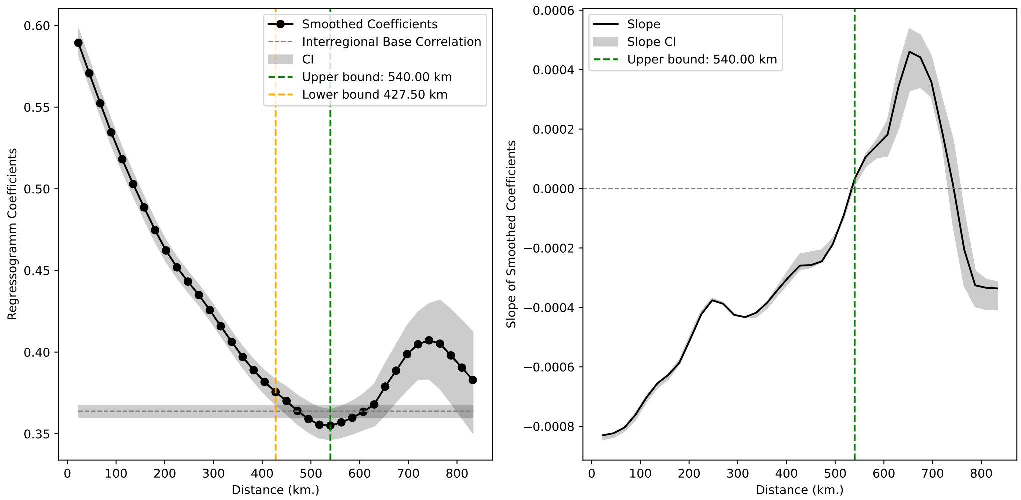
[Bottazzi and Peri, 2003] Bottazzi, L. and Peri, G. (2003). Innovation and spillovers in regions: Evidence from european patent data. *European Economic Review*, 47(4):687–710.

[Chua, 1993] Chua, H. (1993). Regional spillovers and economic growth. Harvard University, unpublished.

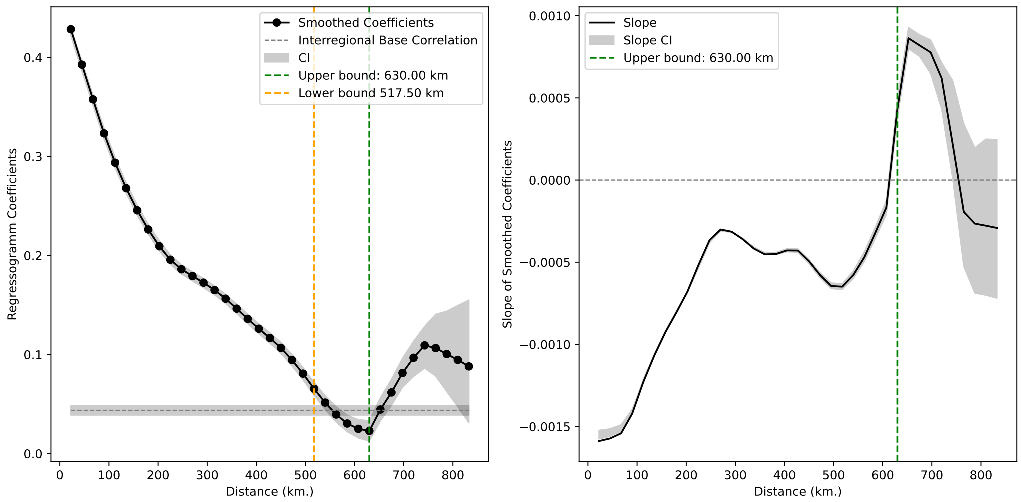
[Clark and van Wincoop, 2001] Clark, T. E. and van Wincoop, E. (2001). Borders and business cycles. *Journal of International Economics*, 55(1):59–85. Intranational & International Economics.

[Dabla-Norris et al., 2010] Dabla-Norris, E., Honda, J., Lahreche, A., and Verdier, G. (2010). Fdi flows to low-income countries: Global drivers and growth implications. IMF Working Paper WP/10/132, International Monetary Fund. Strategy, Policy, and Review and African Departments.

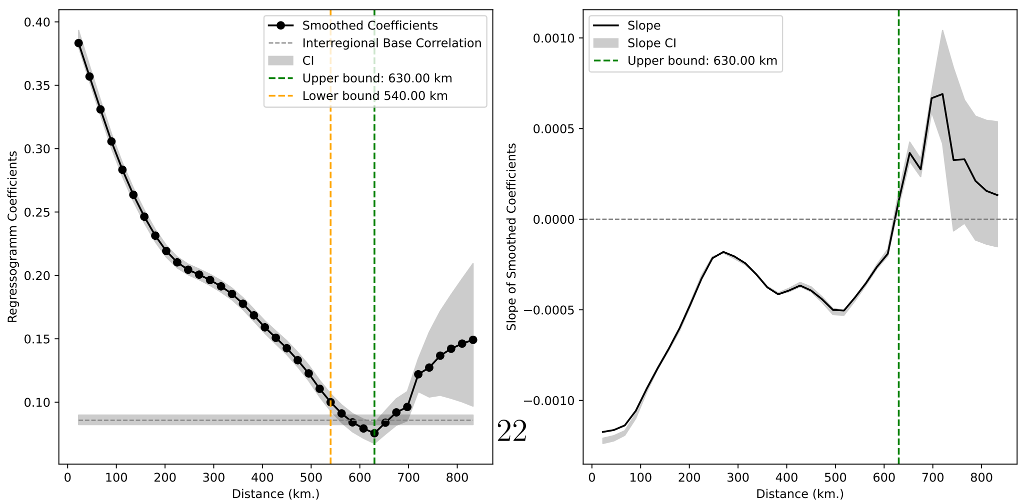
[Elhorst et al., 2024] Elhorst, J. P., Tziolas, I., Tan, C., and Milionis, P. (2024). The distance decay effect and spatial reach of spillovers. *Journal of Geographical Systems*, 26(2):265–289.



(a) Correlation (GDP levels) and Slope (smoothed)

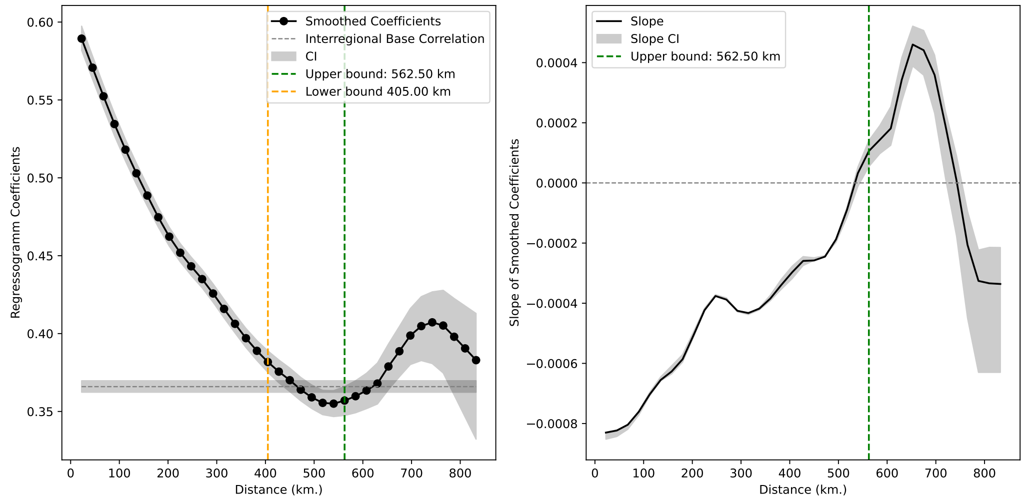


(b) Correlation (GDP differences) and Slope (smoothed)

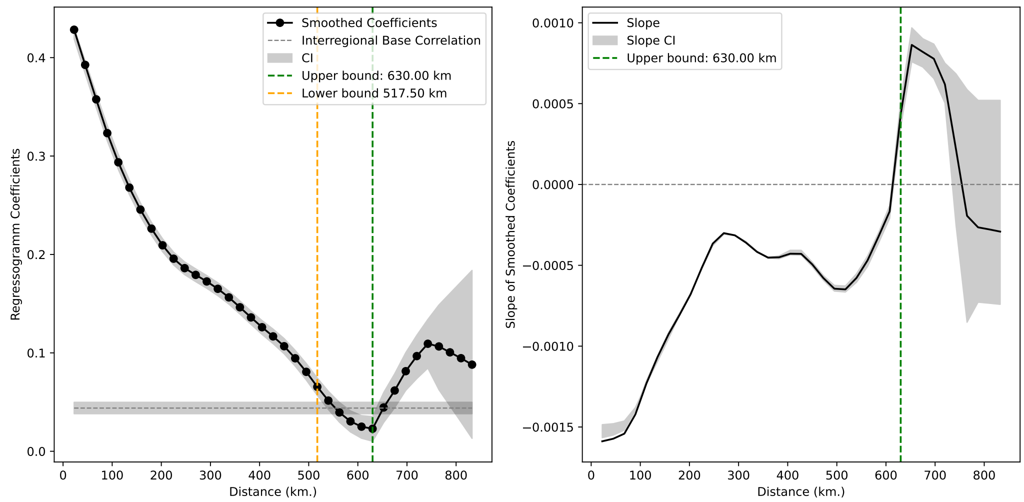


(c) Correlation (AC. adj. GDP differences) and Slope (smoothed)

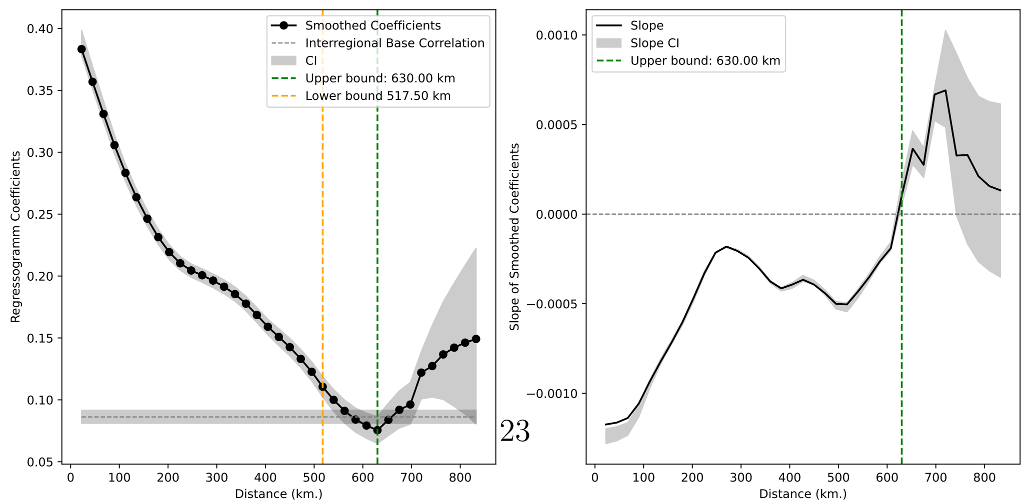
Figure 11: Estimation without Control Variables.



(a) Correlation (GDP levels) and Slope (smoothed)



(b) Correlation (GDP differences) and Slope (smoothed)



(c) Correlation (AC. adj. GDP differences) and Slope (smoothed)

Figure 12: Estimation with Control Variables.

	Lower bound (km.)	**Upperbound (km.)**
No control variables		
Levels	427.5	540
Differences	517.5	630
Prewhitened Differences	540	630
With control variables		
Levels	405	652.5
Differences	517.5	630
Prewhitened Differences	515.5	630

Table 2: Estimated Market Extent bounds (Alternative bounds test).

[Ertur and Koch, 2007] Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22(6):1033–1062.

[Frankel and Rose, 1998] Frankel, J. A. and Rose, A. K. (1998). The endogenity of the optimum currency area criteria. *The Economic Journal*, 108(449):1009–1025.

[Fujita et al., 1999] Fujita, M., Krugman, P., and Venables, A. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Cambridge, MA.

[Funke and Niebuhr, 2005] Funke, M. and Niebuhr, A. (2005). Regional geographic research and development spillovers and economic growth: Evidence from west germany. *Regional Studies*, 39(1):143–153.

[Horowitz, 1981] Horowitz, I. (1981). Market definition in antitrust analysis: A regression-based approach. *Southern Economic Journal*, 48(1):1–.

[Hsu et al., 2011] Hsu, C.-C., Wu, J.-Y., and Yau, R. (2011). Foreign direct investment and business cycle co-movements: The panel data evidence. *Journal of Macroeconomics*, 33(4):770–783.

[Marshall, 1920] Marshall, A. (1920). *Principles of Economics*. Macmillan, London.

[Marshall, 2013] Marshall, A. (2013). *Principles of Economics*. Palgrave Classics in Economics. Palgrave Macmillan London, 1 edition.

[McCallum, 1995] McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns. *American Economic Review*, 85(3):615–623.

[Moreno and Trehan, 1997] Moreno, R. and Trehan, B. (1997). Location and the growth of nations. *Journal of Economic Growth*, 2:399–418.

[Puga and Venables, 1996] Puga, D. and Venables, A. (1996). The spread of industry: Spatial agglomeration in economic development. *Journal of the Japanese and International Economies*, 10:440–464.

[Rodríguez-Pose and Crescenzi, 2008] Rodríguez-Pose, A. and Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in europe. *Regional Studies*, 42(1):51–67.

[Slutzky, 1937] Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica*, 5(2):105–146.

[Stigler and Sherwin, 1983] Stigler, G. J. and Sherwin, R. A. (1983). The extent of the market.

[Tinbergen, 1963] Tinbergen, J. (1963). Shaping the world economy. *The International Executive*, 5(1):27–30.

[Wang and Sun, 2022] Wang, T. and Sun, F. (2022). Global gridded gdp data set consistent with the shared socioeconomic pathways. *Scientific Data*, 9(1):221.