

Spatial Inequality

Sergio J. Rey

The study of spatial inequality has gained prominence due to its significant implications for both scientific research and policy-making. This chapter explores the concepts and computational methods used to measure spatial inequality, emphasizing a reproducible approach that social scientists can apply to their research. The analysis focuses on geographic income disparities at the sub-national level, using Mexico as a case study. By examining various a-spatial and spatially explicit approaches, the chapter highlights the complexities of measuring inequality across places and over time. The discussion includes a review of traditional inequality measures and introduces spatial decomposition methods that account for the geographical distribution of income. The findings underscore the importance of integrating spatial considerations into inequality analysis to better understand the patterns and drivers of regional disparities, thereby informing more effective and equitable policy interventions.

Geographic income inequality has risen more than 40% between 1980 and 2021.

– U.S. Department of Commerce (2023)

Introduction

The study of spatial, or geographical, disparities is crucial for both scientific and policy-oriented reasons. Scientifically, understanding these disparities allows researchers to uncover patterns and correlations that are vital for advancing knowledge in various fields such as economics (Kanbur and Venables 2005), public health (Deb Nath and Odoi 2024), and environmental science (Venter et al. 2023). From a policy perspective, recognizing and addressing geographical disparities is essential for promoting social equity and economic development. A prime example is the European Union's Cohesion Policy, where the reduction of spatial disparities between member regions takes center stage (Agnieszka 2019). Additionally, understanding spatial disparities can inform urban planning and environmental policies to create more sustainable and resilient communities.

By bridging the gap between scientific research and policy implementation, the study of geographical disparities helps in crafting evidence-based strategies that promote balanced regional

development, reduce inequalities, and improve the overall quality of life. Ultimately, this interdisciplinary approach fosters a deeper understanding of the complex dynamics at play and supports the creation of more inclusive and effective policies.

The field of spatial data science (Rey, Arribas-Bel, and Wolf 2023) provides tools to visualize and analyze spatial inequality. Thus, it is well-positioned to support such interdisciplinary research. The goal of this chapter is to introduce social scientists to the concepts and measurement of spatial inequality. The emphasis is on adopting a computationally focused and reproducible treatment that would allow researchers to apply the methods introduced here to their own investigations.

The chapter proceeds as follows. We first develop a conceptual understanding of the measurement of spatial inequality. Next, we describe the computational environment employed to analyze spatial disparities. The specific case study is then introduced. We then discuss different a-spatial approaches towards measuring inequality, followed by a detailed exploration of spatially explicit approaches for measuring geographical disparities. The chapter concludes with the identification of future research areas in the field of spatial inequality.

Inequality Concepts

The growing concern with inequality brings to mind Peter Drucker’s often-cited principle, “You can’t improve what you don’t measure.” Before we can address the technical challenges of measuring spatial inequality, it is essential to first grapple with the conceptual issues surrounding what we are measuring.

It is important to distinguish between terms that frequently appear in the inequality literature: equality (inequality) and equity (inequity). Equality refers to the state of being equal, particularly in status, rights, and opportunities. In economics, this often means distributing resources and opportunities uniformly across all individuals or groups.

Equity, conversely, involves fairness and justice in the distribution of resources and opportunities. It considers individual needs and circumstances, aiming to level the playing field. Therefore, although inequality and inequity are interconnected within the context of social justice (Sen 2004), they are not synonymous.

For instance, an equal distribution of resources, such as uniform per capita expenditure on students, can create inequities by ignoring the challenges faced by students in different contexts, such as urban versus rural districts or advantaged versus disadvantaged neighborhoods (Tine 2017). Conversely, some distributions are intentionally unequal to achieve greater equity. Progressive income tax schemes, where the tax rate increases with income, are a prime example of this approach (Ledić, Rubil, and Urban 2023).

A closely related distinction is between equality of outcomes and equality of opportunities. Inequality in outcomes refers to the unequal distribution of income, wealth, and resources among

individuals in a society, which can result from factors such as luck, effort, and inherited wealth. In contrast, inequality in opportunities focuses on the unequal access to education, healthcare, and other essential services that enable individuals to achieve their potential, irrespective of their background (Roemer 1998). Studies may differ, then, in whether they measure spatial inequality in outcomes (Khedmati Morasae et al. 2024) or spatial disparities in opportunities (Knaap 2017).

In addition to the distinction between inequality and inequity, and between outcomes versus opportunities, there is much variation in the substantive variable under focus. Income studies dominate the literature quantitatively (Gaubert et al. 2021) and are sometimes contrasted with studies of the inequality of wealth (Suss, Kemeny, and Connor 2024).¹ More granular studies examine disparities in the sources of income, such as wages, as well as employment rates (Overman and Xu 2022). Outside of economics, topics such as disparities in educational outcomes (Graetz et al. 2020), health outcomes (Khedmati Morasae et al. 2024), voting patterns (Barber and Holbein 2022), among many others, are replete across the social and life sciences.

The unit of measurement employed in inequality analysis is also an important consideration. Sala-i-Martin (2006) demonstrates that when using countries as the unit of analysis, the picture that emerges is one of large and static levels of international inequality. However, when the analysis used countries weighted by their populations, the view is one of declining inequality over time. Later, we shall see that this choice of weighting versus non-weighting of the units is an important issue in spatial inequality measurement.

There is much variation across the inequality literature in the unit of measure. Studies of personal income inequality often focus on data recorded for individuals (Piketty and Saez 2003). Other studies take the household or family as the unit of analysis (Brandolini and Smeeding 2011). In both cases, the focus is on income inequality across people. This is an essential point of departure for our study of spatial inequality, which is where the unit of measurement is a geographical area. In other words, in spatial inequality analysis, the focus is inequality across places.

Still further, some studies examine the spatial distribution of personal income inequality (Partridge, Partridge, and Rickman 1998; Frank 2009)—that is, how inequality between individuals within a state varies across states. In the spatially oriented inequality studies, the geographical unit of analysis can range from countries (Milanović 2018), to intra-national regions (Ganong and Shoag 2017), to cities (U.S. Department of Commerce 2023), and down to neighborhoods (Nijman and Wei 2020).

A final inequality concept we need to consider is the role of time. One question is the time unit. Is income measured per person, per year, or is some life-time earnings, or permanent income (Hall 1978) measure employed? A second question pertains to whether the study of inequality is a snapshot at one point in time or focuses on the dynamics of income distribution.

¹It is important to keep in mind that income is measured as a flow whereas wealth is a stock. This distinction matters in terms of the way disparities in the two variables are examined.

Addressing all these issues is beyond the scope of any one study. We raise them here in order to situate the study of spatial inequality in a much broader context. For this chapter, we will hone in on the question of measuring spatial income inequality at the sub-national scale.

Computational Environment

In the following section, we present the packages and computational environment used. The narrative following code cells explains the computational concepts.

Packages

```
import inequality as ineq ①
import numpy as np ②
import pandas as pd
import geopandas as gpd
import libpysal as lps
import seaborn as sns
import matplotlib.pyplot as plt
import watermark
%load_ext watermark
%watermark -a "Sergio Rey" -u -n -t -v \
    -p numpy,pandas,scipy,matplotlib,inequality,seaborn,libpysal
```

① We alias the package `inequality` as `ineq`

② We do the same for `numpy`

The watermark extension is already loaded. To reload it, use:

```
%reload_ext watermark
Author: Sergio Rey
```

Last updated: Wed Aug 21 2024 10:20:08

Python implementation: CPython

Python version : 3.12.4

IPython version : 8.26.0

numpy : 2.0.1

pandas : 2.2.2

scipy : 1.14.0

```
matplotlib: 3.9.1
inequality: 0.1.dev170+gd70b5c0.d20240820
seaborn    : 0.13.2
libpysal   : 4.12.0
```

Our analysis of spatial inequality utilizes packages from the Python Spatial Analysis Library (`pysal`) (Rey et al. 2022), together with `geopandas` (Jordahl et al. 2019) and `seaborn` (Waskom 2021). The main package focusing on measuring spatial disparities is `inequality`, which implements analytics for measuring a-spatial inequality and spatially explicit inequality measures. Also, from `pysal`, we will depend on the `libpysal` package for constructing spatial weights that are central to the analysis of spatial inequality. `geopandas` provides for spatial data processing and producing maps of the spatial distribution of income, while we adopt `seaborn` for constructing a-spatial graphical views of inequality.

For purposes of reproducibility, we include the `watermark` package which reports the version numbers of each of the packages we use in our analysis.

Data

To illustrate the core concepts in spatial inequality measurement, we will rely on a data set for the states in Mexico (Rey and Sastré-Gutiérrez 2010). The variable of interest is state per capita gross domestic product measured in 2000 USD measured for each decade from 1940-2000 for each of 32 areas consisting of the 31 federal states of Mexico plus Mexico City. This data-set is included as an example data-set in `libpysal`.

```
lps.examples.explain("mexico")
```

```
mexico
=====
```

```
Decennial per capita incomes of Mexican states 1940-2000
```

```
-----
* mexico.csv: attribute data. (n=32, k=13)
* mexico.gal: spatial weights in GAL format.
* mexicojoin.shp: Polygon shapefile. (n=32)
```

```
Data used in Rey, S.J. and M.L. Sastre Gutierrez. (2010) "Interregional inequality dynamics : 
```

In addition to the income data contained in the `mexico.csv` file, there are two additional files available in this example: `mexico.gal` which stores information about the contiguity relationships between the states, and `mexicojoin.shp` which is a shapefile.

Figure 1 lists the locations of the 32 Mexican states.²



Figure 1: States of Mexico

```
pth = lps.examples.get_path("mexico.csv")
df = pd.read_csv(pth)
print(f"Shape of dataframe: {df.shape}")
print(f"First 5 rows of dataframe:\n {df.head()}")
df.head()
```

²The python code used to construct this map is hidden.

```
print(f"\nVariables: {df.columns}")
```

Shape of dataframe: (32, 13)

First 5 rows of dataframe:

	State	pcgdp1940	pcgdp1950	pcgdp1960	pcgdp1970	pcgdp1980	\
0	Aguascalientes	10384.0	6234.0	8714.0	16078.0	21022.0	
1	Baja California	22361.0	20977.0	17865.0	25321.0	29283.0	
2	Baja California Sur	9573.0	16013.0	16707.0	24384.0	29038.0	
3	Campeche	3758.0	4929.0	5925.0	10274.0	12166.0	
4	Chiapas	2934.0	4138.0	5280.0	7015.0	16200.0	

	pcgdp1990	pcgdp2000	hanson03	hanson98	esquivel99	inegi	inegi2
0	20787.0	27782.0	2.0	2.0	3.0	4.0	4.0
1	26839.0	29855.0	1.0	1.0	5.0	1.0	1.0
2	25842.0	26103.0	2.0	2.0	6.0	1.0	1.0
3	51123.0	36163.0	6.0	5.0	4.0	5.0	5.0
4	8637.0	8684.0	5.0	5.0	7.0	5.0	5.0

```
Variables: Index(['State', 'pcgdp1940', 'pcgdp1950', 'pcgdp1960', 'pcgdp1970',
                  'pcgdp1980', 'pcgdp1990', 'pcgdp2000', 'hanson03', 'hanson98',
                  'esquivel99', 'inegi', 'inegi2'],
                  dtype='object')
```

We can read the `mexico.csv` file using `pandas` to create a `DataFrame` that will hold the attributes of interest. This dataframe has the shape (32,13) indicating there are 32 observations on 13 variables. The 13 variables include the per capita gross domestic product for each decade, for example `pcgdp1990` for 1990, together with the state name, and six variables that define different regionalization schemes for the country. We will return to these regional variables in a later section.

Measuring Spatial Inequality in Mexico

Visualizing Inequality in Distributions

We begin with different perspectives on the distribution of state incomes in Mexico. From a geographical perspective, Figure 2 shows the spatial distribution of incomes for the last year of the sample (2000).

```

pth = lps.examples.get_path("mexicojoin.shp")
gdf = gpd.read_file(pth)
ax = gdf.plot(column="PCGDP2000", k=5, scheme="Quantiles",
              legend=True,
              edgecolor='grey',
              legend_kwds={"loc": "center left",
                           "bbox_to_anchor": (1, 0.5),
                           "fmt": "{:.0f}"})
ax.set_axis_off()
ax.set_title("PC GDP 2000");

```

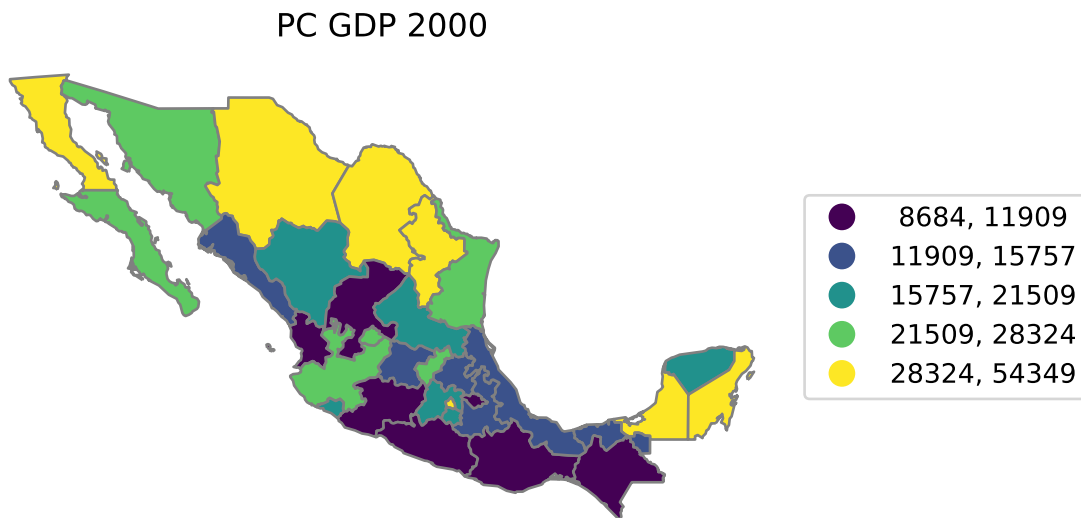


Figure 2: Per Capita Gross Domestic Product by State (Quintiles)

This is a choropleth map using quintiles to classify the incomes. The visual impression is that incomes are not randomly distributed in Mexico, as the states with incomes below the bottom quintile are more concentrated in the south, while in the north, the highest income states dominate. We will be able to make more quantitative evaluation of this spatial pattern later on in this chapter.

Figure 3 presents a different perspective on income distribution based on the concept of Pen's Parade, introduced by Dutch economist Jan Pen (Pen 1971). The metaphor uses a Parade to illustrate economic inequality, with each person representing a state in the economy, and their height being proportional to the state's per capita income. The Parade starts with the shortest individuals depicting the poorest states, gradually increasing in height as income rises. At the end of the parade, the tallest individuals represent the most affluent states, highlighting

the significant income disparities across states. This visual tool showcases the inequality in income distribution. The parade uses two different scales, with the x-axis showing the ordinal distances between states and the y-axis representing the interval distances in their per capita incomes.

```
from inequality.pen import pen
f = pen(gdf, 'PCGDP2000', 'NAME')
```

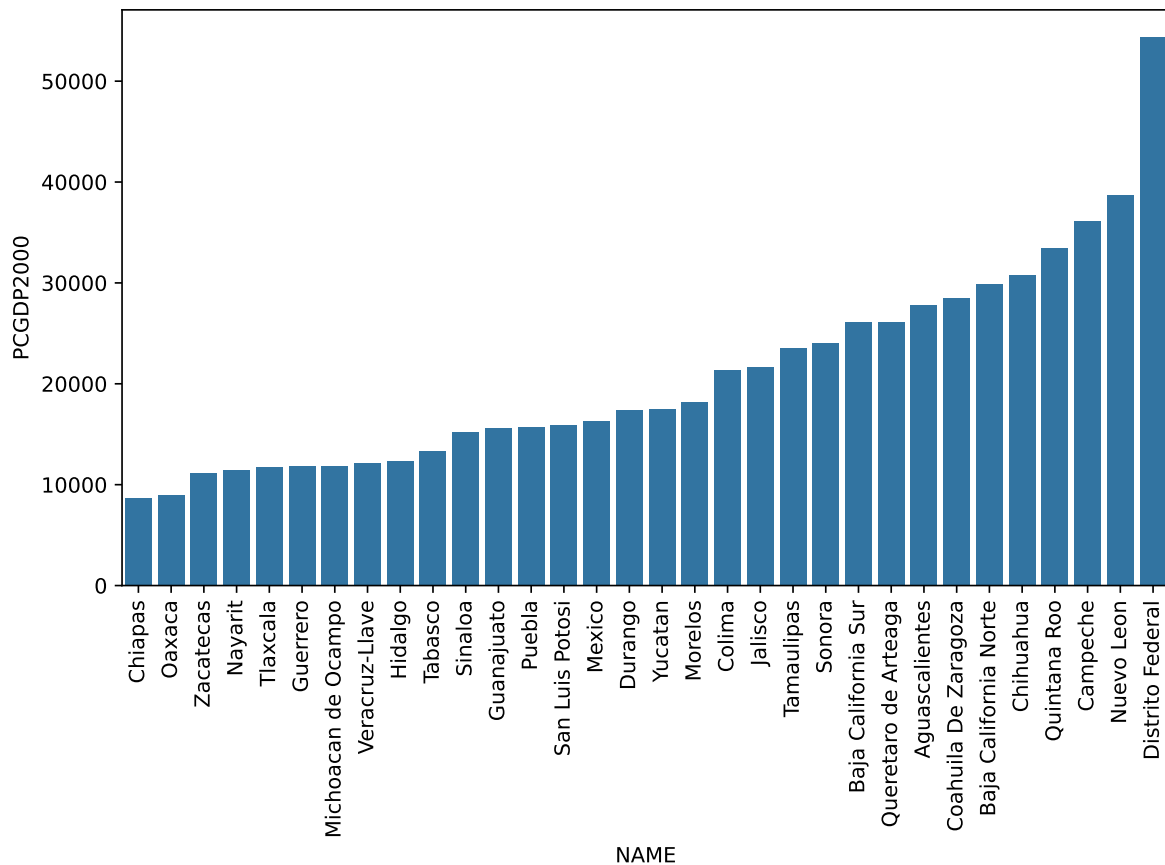


Figure 3: Pen's Parade Per Capita Gross Domestic Product by State 2000

We can start to integrate the spatial and attribute distributions together using a **pengram** from the **pysal-inequality** package:

```
from inequality.pen import pengram
f = pengram(gdf, 'PCGDP2000', 'NAME', xticks=False, leg_pos='lower left',
            fmt="{:.0f}")
```

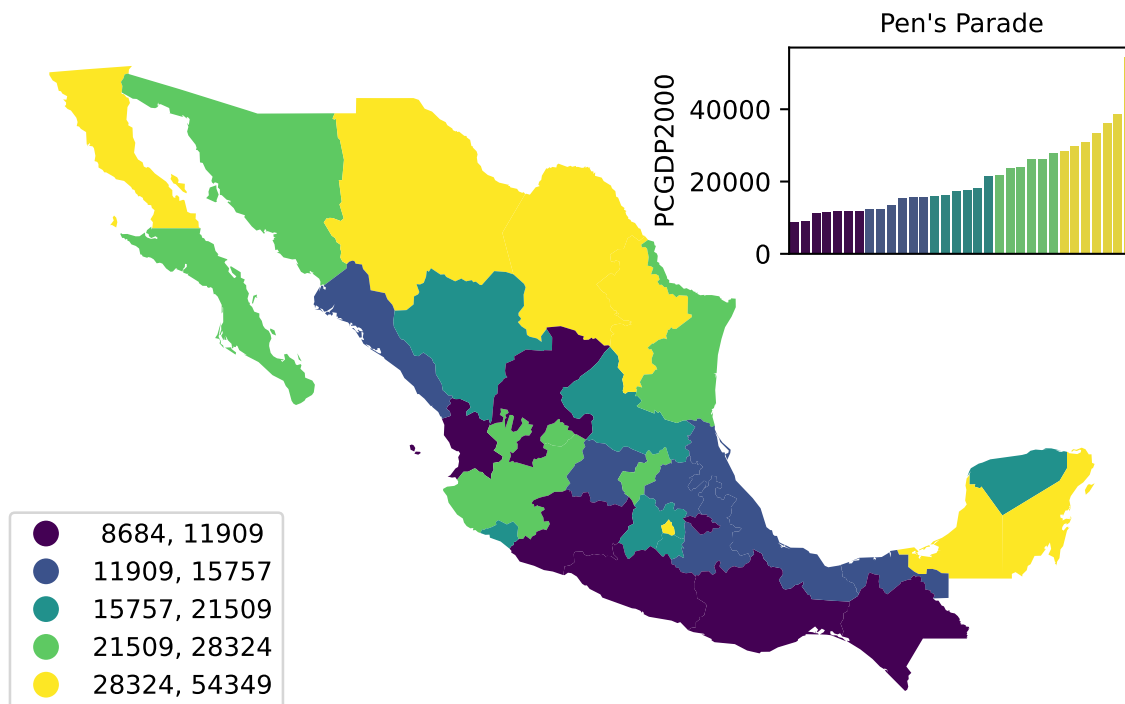


Figure 4: Pengram Per Capita Gross Domestic Product by State 2000

As shown in Figure 4, the **pengram** combines the Pen's Parade alongside the choropleth map. This affords a more granular view of the distribution than those offered by either view in isolation. For example, one of the well-known limitations of a quintile classed map is that the intra-class variation is obscured. In the **pengram**, the intra-class variation now becomes visible through the Pen's Parade, revealing the much larger absolute and relative variance above the upper quintile relative to the other classes.

A second feature of the **pengram** is the ability to query for specific observations. This makes it possible to locate the position of a state in both the Pen's Parade (attribute space) as well as on the map (geographical space). We do this for states occupying the two extremes of the attribute distribution: Chiapas and Distrito Federal in Figure 5. While the two reside in the extremes of the attribute distribution, the high income Distrito Federal is in the center of the geographic distribution while Chiapas is on the southern border of the country. Moreover, although Distrito Federal stands out in the Pen's Parade, its small geographic area makes it difficult to identify on the map without the query functionality of the **pengram**.

```
from inequality.pen import pengram
f = pengram(gdf, 'PCGDP2000', 'NAME', leg_pos='lower left',
            fmt="{:.0f}"),
```

```
xticks=False,
query=['Chiapas', 'Distrito Federal'])
```

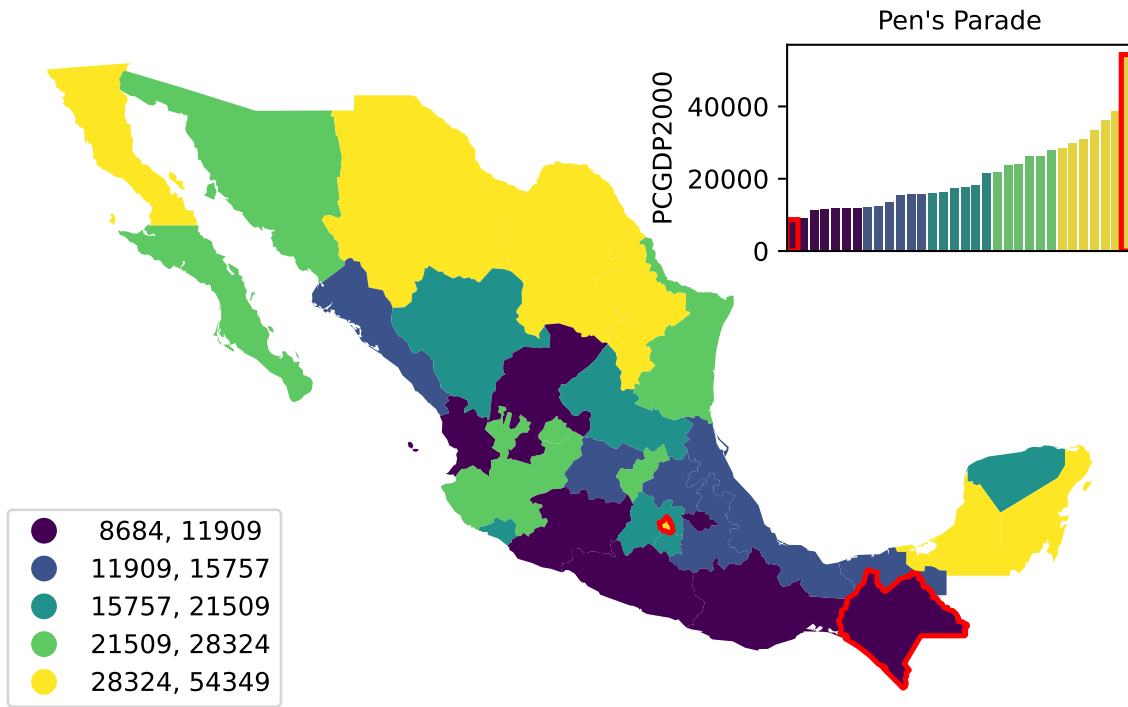


Figure 5: Querying the Pengram of Per Capita Gross Domestic Product by State 2000

Returning to a more granular view of the attribute distribution, Figure 6 combines a histogram of the distribution together with a kernel density estimate, and a rug plot. The latter signifies the positions of each state as short ticks on the x-axis. The outlier nature of the Distrito Federal that we saw in the `pengram` is responsible for the positive (right) skew of the density function. There is some evidence of polarization in the distribution with the mode being at the poorest group of states and other, lower, peaks in the middle of the distribution.

```
years = range(1940, 2010, 10)
yvars = [f'pcgdp{year}' for year in years]
sns.histplot(df[yvars[-1]], kde=True, element="step",
             bins=10, edgecolor="black")
sns.rugplot(df[yvars[-1]], color='blue')
plt.show()
```

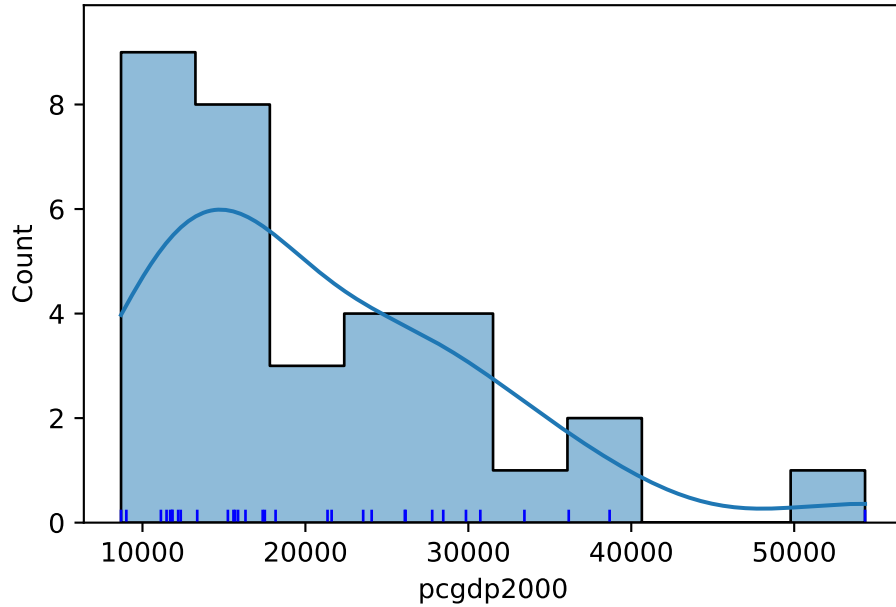


Figure 6: Distribution Per Capita Gross Domestic Product by State 2000, histogram, kernel density, rug plot

Inequality in a distribution is often considered by an inspection of the shares of overall income belonging to units at different locations within the distribution. Here, we must keep in mind, the distinction between the different units under study in the analysis of spatial versus personal income inequality. In spatial inequality analysis, we essentially treat each state as an “individual” and set that individual’s level of income to the state’s per capita income. The share for the state is then derived as the ratio of its per capita income to the sum of the per capita income of all states.

These shares can be portrayed in a Lorenz curve, shown in Figure 7, which orders the states by their per capita incomes from lowest to highest. Then, against the cumulative proportion of states (x-axis) we plot the cumulative income share on the y-axis. Both scales have limits of $[0, 1]$. In the case of perfect equality, where all state per capita incomes are equal, this plot would be a 45-degree line, the so called line of perfect equality. Any departure from perfect equality will result in states with above average per capita income receiving more than $1/n$ share of $n\bar{y}$, while states with below average per capita incomes will have shares below $1/n$.³

```
from inequality.schutz import Schutz
s = Schutz(gdf, 'PCGDP2000')
s.plot(xlabel='Share of States', ylabel='Share of Per Capita Income')
```

³States with shares below $1/n$ would also have location quotients of less than one for their relative per capita incomes, where their per capita income was expressed relative to national per capita income.

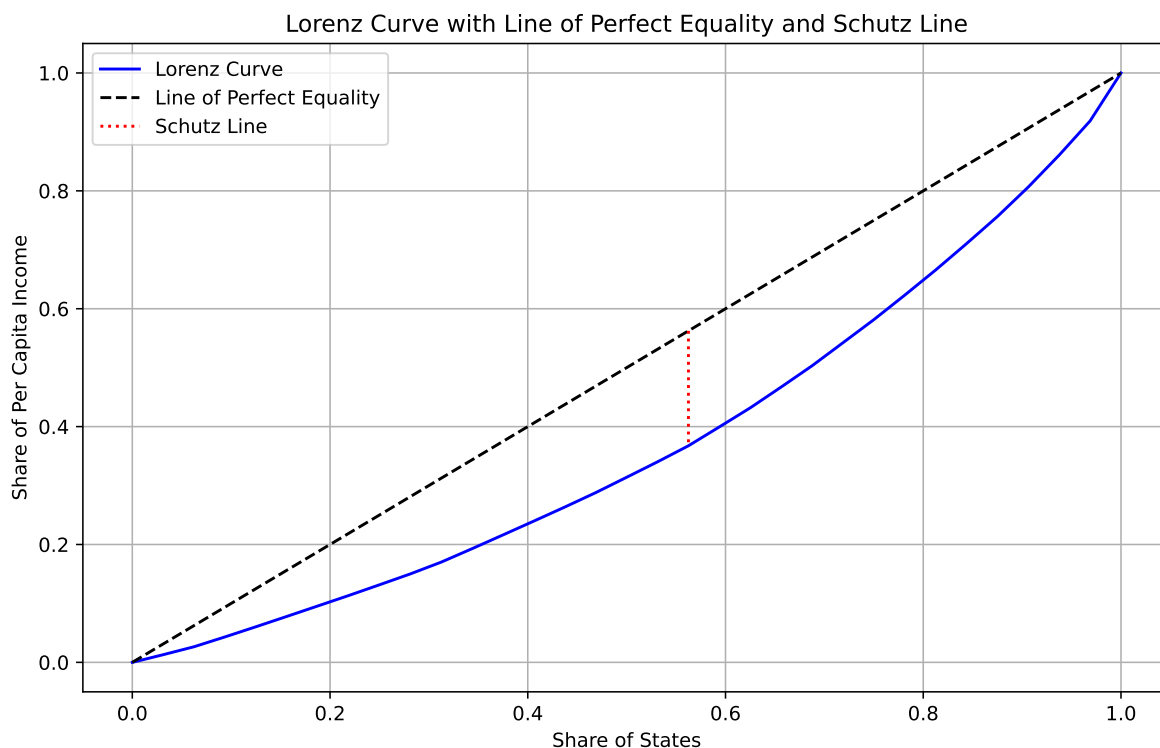


Figure 7: Lorenz Curve and Schutz Line Per Capita Gross Domestic Product by State 2000

Measures of Inequality

We can derive two indices from the Lorenz curve: the Gini coefficient and the Schutz coefficient. The Gini coefficient is an area measurement expressed as a ratio of the area of the lens above the Lorenz curve but below the diagonal to area under the line of perfect equality. The Schutz coefficient is a distance measure defined as the maximum vertical distance between the equality diagonal and the Lorenz curve. Both coefficients measure the extent to which the Lorenz curve departs from perfect equality. Moreover, both indices run from 0, perfect equality, to 1, maximal inequality.

For 2000, the Gini coefficient is 0.258, while the Schutz coefficient is 0.195. As they are both scaled to be within the unit interval, the temptation is to compare the two values. However, this would be misleading as we recall one is a measure of area, the second a measure of distance. Instead, each coefficient is often used to compare across distributions, either over time or space. We can do this for Mexico by asking what has happened to inequality over time. Along the way, we add a third commonly used indicator of inequality, the coefficient of variation (CV). The coefficient of variation is a relative measure of variation as it is defined as the ratio of the standard deviation to the sample mean.

```

yvars = [f'PCGDP{year}' for year in years]
ginis = [ineq.gini.Gini(gdf[yvar]).g for yvar in yvars]
res_df = pd.DataFrame(data=ginis, columns=['Gini'], index=years)
cv = gdf[yvars].std() / gdf[yvars].mean()
res_df['CV'] = cv.values
s = [ineq.schutz.Schutz(gdf, yvar).distance for yvar in yvars]
res_df['Schutz'] = s
res_df['Gini_rank'] = res_df['Gini'].rank()
res_df['CV_rank'] = res_df['CV'].rank()
res_df['Schutz_rank'] = res_df['Schutz'].rank()
res_df

```

Table 1: Inequality Index Rankings

	Gini	CV	Schutz	Gini_rank	CV_rank	Schutz_rank
1940	0.353724	0.719858	0.260037	7.0	7.0	7.0
1950	0.296446	0.624611	0.213920	6.0	6.0	6.0
1960	0.253718	0.492447	0.181549	3.0	3.0	3.0
1970	0.255134	0.472039	0.185266	4.0	2.0	4.0
1980	0.245053	0.462657	0.179702	1.0	1.0	1.0
1990	0.251818	0.497729	0.181363	2.0	5.0	2.0
2000	0.258113	0.492565	0.195043	5.0	4.0	5.0

Table 1 shows that the three indices agree that inequality was lowest in 1980, and the highest in 1940. However, while the Gini and Schutz coefficients agree in their rankings, the inclusion of the CV creates discordance. Part of the discordance reflects the sensitivity of the measures to different parts of the income distribution. The Gini coefficient puts more weight on the middle of the distribution, while the CV is more affected by the right tail of the distribution. The discordance also reflects the property that when the Lorenz curves do not intersect, the CV and Gini would agree on the rankings of inequality. However, Figure 8 shows that there are cases where the Lorenz curves intersect.

The discordance in rankings complicates whether income inequality between states in Mexico has increased or decreased. The answer now depends upon which temporal interval one chooses. Both the Gini and CV agree that inequality has declined since 1940, irrespective of the terminal year selected. However, they disagree on the answer when the question is whether inequality decreased from, say, 1960 to 1970 or between 1990 and 2000 (see Table 1).

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

```

```

from scipy.stats import cumfreq

def lorenz_curve(incomes):
    sorted_incomes = np.sort(incomes)
    cumulative_incomes = np.cumsum(sorted_incomes)
    normalized = cumulative_incomes / cumulative_incomes[-1]
    lorenz_curve = np.insert(normalized, 0, 0)
    n = len(incomes)
    x = np.linspace(0.0, 1.0, n + 1)
    return x, lorenz_curve

plt.figure(figsize=(10, 8))

for yvar in yvars:
    incomes = gdf[yvar].values
    x, y = lorenz_curve(incomes)
    plt.plot(x, y, label=yvar)

# Plotting the line of equality
plt.plot([0, 1], [0, 1], color='black', linestyle='--')

# Adding titles and labels
plt.title('Lorenz Curves for PCGDP by Year')
plt.xlabel('Cumulative Share of States')
plt.ylabel('Cumulative Share of PCGDP')
plt.legend()
plt.grid(True)
plt.show()

```

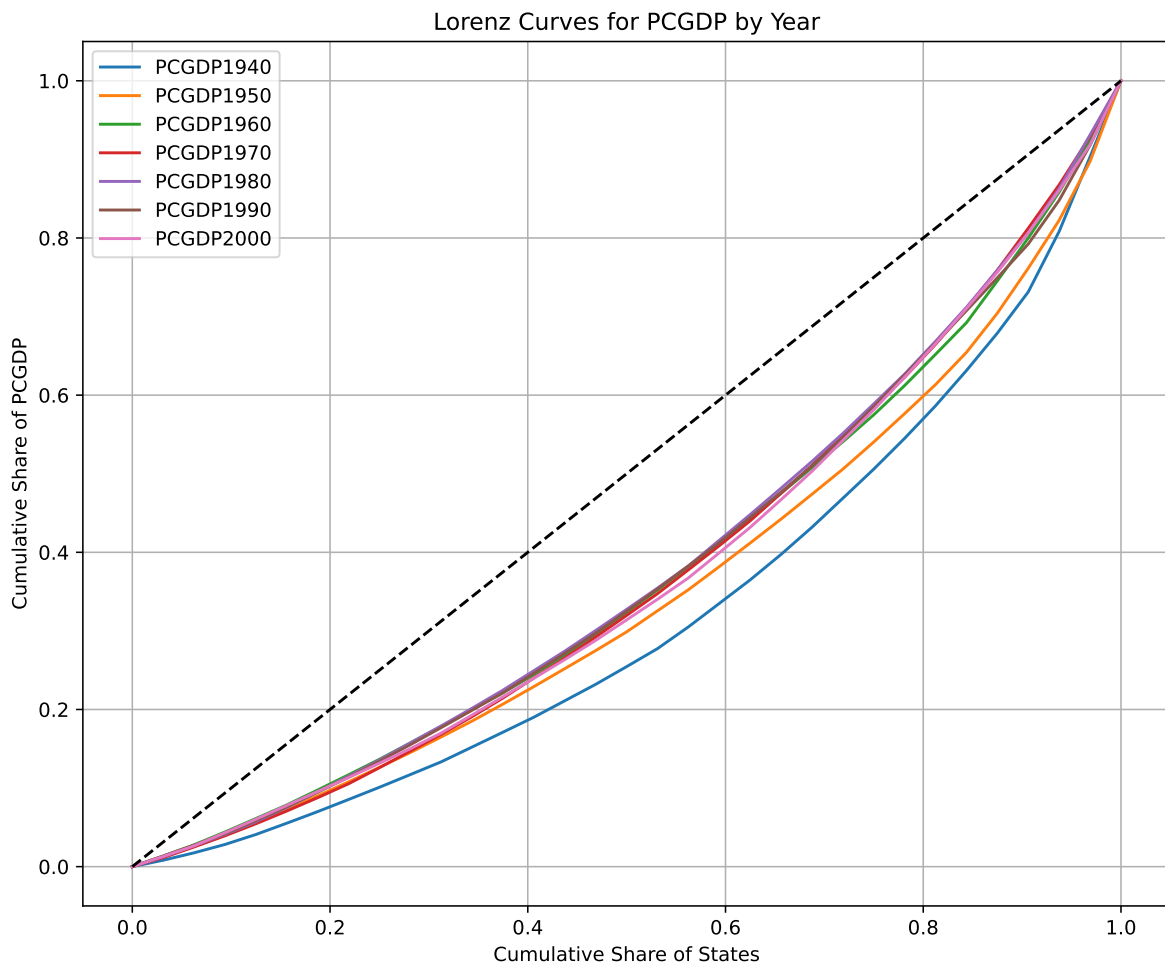


Figure 8: Lorenz Curves by Decade

The CV and Gini are but two of a collection of scalar measures of inequality we could apply. Other commonly employed inequality measures include the variance, the relative mean deviation, logarithmic variance, variance of the logarithms, Atkinson, and Theil.⁴ Choosing among the rich diversity of inequality measures has been the subject of a vast literature. To help guide that selection, there are five desirable properties of an inequality measure:

1. Symmetry or anonymity
2. Principle of transfers
3. Scale invariance
4. Replication invariance
5. Zero normalization

⁴See Cowell (2011) for an overview of inequality measures.

Symmetry implies that the names of the income receiving unit should be immaterial. That is, if we swap the income of one geographical unit with that of another, the overall inequality measure should not change. The principle of transfers implies that the measure should reflect a reduction in inequality if income is transferred from a richer unit to a poorer unit, as long as the transfer does not reverse their income ranking. Scale invariance means that if all incomes are multiplied by the same constant, inequality remains unchanged. Replication invariance implies that an inequality measure should be unaffected if the population is replicated, meaning that duplicating the entire income distribution does not alter the measure of inequality. Zero normalization means that an inequality measure assigns a value of zero to a perfectly equal income distribution, serving as a baseline where all units have the same income.

Not all of the inequality measures satisfy all of these five properties. For example, the variance is not scale invariant, and both the logarithmic variance and variance of logarithms fail the principle of transfers property. Even for measures that respect these five properties, their use in comparing the levels of inequality between countries at one point in time, or the same country at different points of time, requires the researcher to specify a social welfare function (Atkinson 1970).

However, in the study of spatial inequality, it is important to note that all the measures mentioned above share a sixth property: *spatial invariance*. This means they are insensitive to the *geographical* distribution of the income values. The spatial invariance is demonstrated in Figure 9, where we compare two spatial distributions, one that is the actual distribution for 2000 and one where we artificially permute the incomes randomly. Despite markedly different spatial patterns, the income distributions summarized in the histograms in the bottom row are identical. Since all of the inequality measures mentioned only consider information about the statistical distribution, each measure will take on the same value whether applied to the spatial distribution on the left or right of the figure.

From a geographical perspective, spatial invariance is not a desirable property in an inequality measure. Let's now turn our attention to spatially explicit measures of inequality.

```
fig, axs = plt.subplots(2, 2)

gdf.plot(column='PCGDP2000', ax=axs[0, 0], scheme='quantiles',
         cmap='viridis')
axs[0, 0].set_title('PCGDP2000')
axs[0, 0].axis('off')

gdf['PCGDP2000r'] = np.random.permutation(gdf.PCGDP2000)

gdf.plot(column='PCGDP2000r', ax=axs[0, 1], scheme='quantiles',
         cmap='viridis')
axs[0, 1].set_title('PCGDP2000 Random')
```

```

axs[0, 1].axis('off')

axs[1, 0].hist(gdf['PCGDP2000'], bins=30, color='skyblue',
              edgecolor='black')
axs[1, 0].set_title('PCGDP2000 Histogram')
axs[1, 0].set_xlabel('PCGDP2000')
axs[1, 0].set_ylabel('Frequency')

axs[1, 1].hist(gdf['PCGDP2000r'], bins=30, color='skyblue',
              edgecolor='black')
axs[1, 1].set_title('PCGDP2000 Random Histogram')
axs[1, 1].set_xlabel('PCGDP2000r')
axs[1, 1].set_ylabel('Frequency')

plt.tight_layout()

plt.show()

```

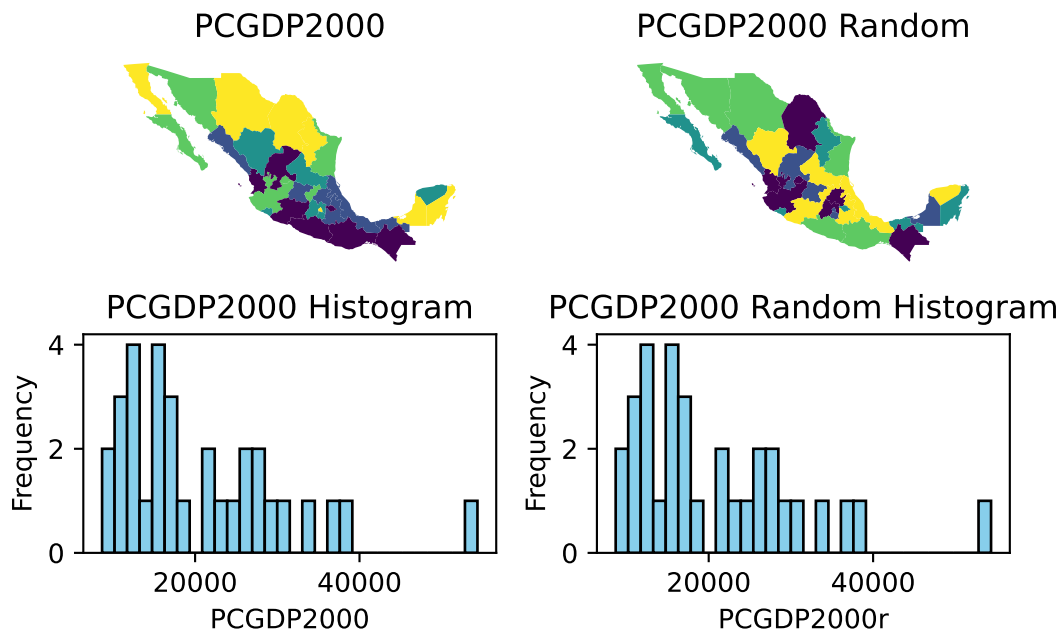


Figure 9: Spatial Invariance of Distributions

Putting Space into the Measurement of Inequality

The inability of the inequality measures to capture any of the geographical dimensions of inequality stems from the treatment of the geographical units of measurement as individuals, and the desire to respect the principle of symmetry or anonymity that is desired in a classic inequality measure. This enables researchers to draw upon the wealth of knowledge about the properties of classic inequality measures, but at the cost of ignoring geography. In other words, these measures say a lot about inequality in the statistical distribution of incomes but they are silent on the spatial distribution of incomes.

In this section, we discuss the approaches used to integrate the spatial and statistical distributions in the study of spatial inequality. Given that these methods take the geographical distribution into account, they can be said to be *spatially explicit measures of inequality*.

These approaches all can be seen as special cases of inequality decomposition.

Regional Inequality Decomposition

A common approach to introducing geography in the measurement of inequality leverages the fact that certain inequality measures can be *decomposed* if the individual income receiving units are placed into a set of mutually exclusive and exhaustive groups. The decomposition then identifies the overall inequality that is due to inequality within groups and between the groups.

To see how this works, we return to the Theil inequality measure that we mentioned briefly above. Let y_i be the income of unit i , and s_i represent the income share of unit i such that $s_i = \frac{y_i}{\sum_i y_i}$ and $\sum_i s_i = 1$. Then, consider the distribution of the shares. When all units have the same income $s_i = s_j = 1/n$. Then the entropy of the shares given as

$$H(y) = \sum_{i=1}^n s_i \ln \frac{1}{s_i}$$

will be maximized at $\ln n$. In the case of extreme inequality, all but one unit have $s_i = 0$ and a single unit has all income $s_j = 1$, and $H(y) = 0$. Thus, the entropy function can be viewed as an indicator of *income equality*.

To generate an indicator of *income inequality*, we can contrast an observed distribution's equality against the maximum:

$$T(y) = \ln n - H(y) = \ln n - \sum_{i=1}^n s_i \ln \frac{1}{s_i} = \sum_{i=1}^n s_i \ln n s_i.$$

This can be viewed as a weighted average of the logarithmic deviations of the shares, with the weights defined as the shares. The logarithmic deviation of share i from perfect equality is $\ln \frac{s_i}{1/n} = \ln s_i n$.

Alternatively, the Theil index can be defined using relative incomes:

$$T = \frac{1}{n} \sum_i \frac{y_i}{\mu} \ln \frac{y_i}{\mu}$$

where y_i is the income of unit i and $\mu = \frac{1}{n} \sum_i y_i$.

Decomposition of the overall T measure requires assigning each unit to exactly one of G sets S_1, S_2, \dots, S_G , with the size of each set given as n_g so that:

$$\sum_{g=1}^G n_g = n.$$

Given this, we have:

$$Y_g = \sum_{i \in S_g} y_i \quad g = 1, \dots, G$$

Defining $\omega_g = \frac{n_g}{n} \frac{\bar{Y}_g}{\mu}$, the Theil index can be rewritten as:

$$T = \sum_{g=1}^G \omega_g T_g + \sum_{g=1}^G \omega_g \ln \frac{\bar{Y}_g}{\mu}. \quad (1)$$

The first term is the within group inequality defined as a weighted average of the inequality within each group with the weights equal to the group's share of overall income, with:

$$T_g = \frac{1}{n_g} \sum_{i \in S_g} \frac{y_i}{\bar{Y}_g} \ln \frac{y_i}{\bar{Y}_g}.$$

The second term in Equation 1 is the between group inequality component, measuring the inequality that would exist if within each group there was no inequality (i.e., all members of the same set have the same income).

Decomposition of inequality had been widely applied in economics to study inequality between occupational groups, sexes, and races, where individuals would be placed into the mutually exclusive groups and overall individual inequality decomposed into that due to the differences between and within the groups. It was a short jump to adopt this to spatial inequality by

using regions to define the groups, with individual units (in our case, states) being assigned to one and only one region.⁵

We will illustrate this for the case of Mexican states using a regional partition due to Hanson (1996) as shown in Figure 10. This regionalization scheme consists of 5 regions, with the size of the regions ranging from 2 states to 10 states.

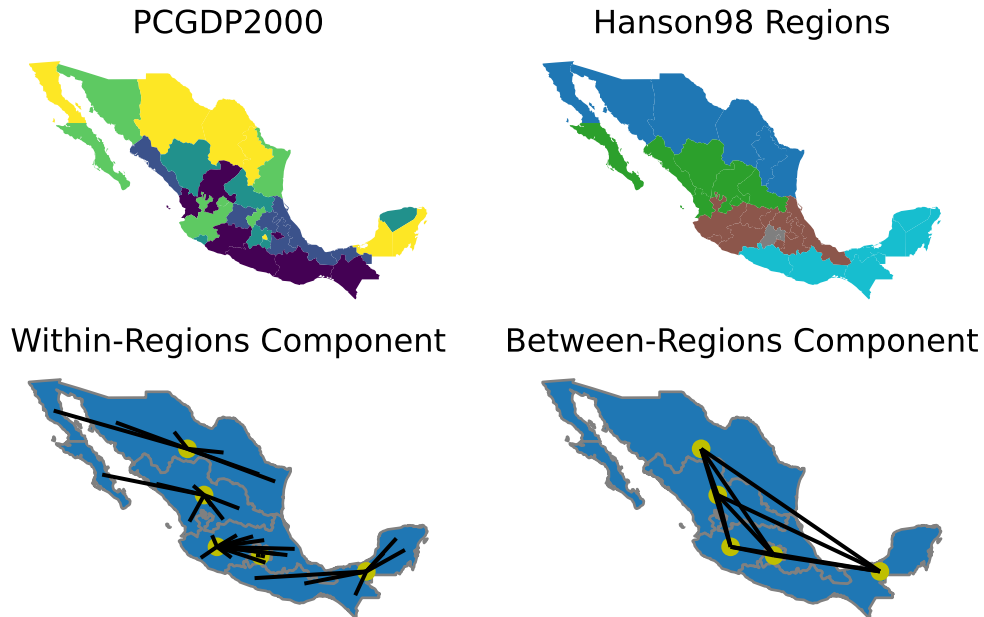


Figure 10: Regional Inequality Decomposition

We can apply the `theil` module from `pysal-inequality` to calculate the value of the overall level of inequality as measured by the global Theil index, and its decomposition into the between region inequality and within region inequality components:

```
fig, axs = plt.subplots(2, 2)
from inequality.theil import TheilDSim
np.random.seed(12345)

gdf.plot('PCGDP2000', ax=axs[0, 0], scheme='quantiles',
         cmap='viridis')
axs[0, 0].set_title('PCGDP2000')
axs[0, 0].axis('off')

# Extract the per capita GDP and regimes
```

⁵One of the earliest applications of decomposition for regional inequality analysis is Theil (1967).

```

income = gdf['PCGDP2000']
regimes = gdf['HANSON98']

res = TheilDSim(income, regimes, 999)

gdf['PCGDP2000r'] = np.random.permutation(gdf.PCGDP2000)
gdf.plot(column='PCGDP2000r', ax=axes[0, 1], scheme='quantiles',
         cmap='viridis')
axes[0, 1].set_title('PCGDP2000 Random')
axes[0, 1].axis('off')

gdf['PCGDP2000r'] = np.random.permutation(gdf.PCGDP2000)
gdf.plot(column='PCGDP2000r', ax=axes[1, 0], scheme='quantiles',
         cmap='viridis')
axes[1, 0].set_title('PCGDP2000 Random')
axes[1, 0].axis('off')

msg=f'Spatial polarization: {res.bg[0][0]/res.T:.3f}'
msg=f'{msg}, pseudo p-value: {res.bg_pvalue[0]}'
print(msg)
realizations = np.array([t.bg/t.T for t in res.results])
print(f'Ho mean: {realizations.mean():.3f}')

kde = sns.kdeplot(realizations, fill=False, color='blue', ax=axes[1,1])
x, y = kde.get_lines()[0].get_data()
plt.legend([], [], frameon=False)
# Fill the area to the right of the specified value
plt.fill_between(x, y, where=(x >= realizations[0]),
                interpolate=True, color='red', alpha=0.5)

# Add vertical line at the specified value
plt.axvline(x=realizations[0], color='red', linestyle='--')
plt.xlabel("Spatial Polarization")
print(f'Theil: {res.T}')
print(f'KB p-value: {(realizations >= realizations[0]).sum()/1000}')

```

Spatial polarization: 0.341, pseudo p-value: 0.036
 Ho mean: 0.139
 Theil: 0.10660832349588023
 KB p-value: 0.036

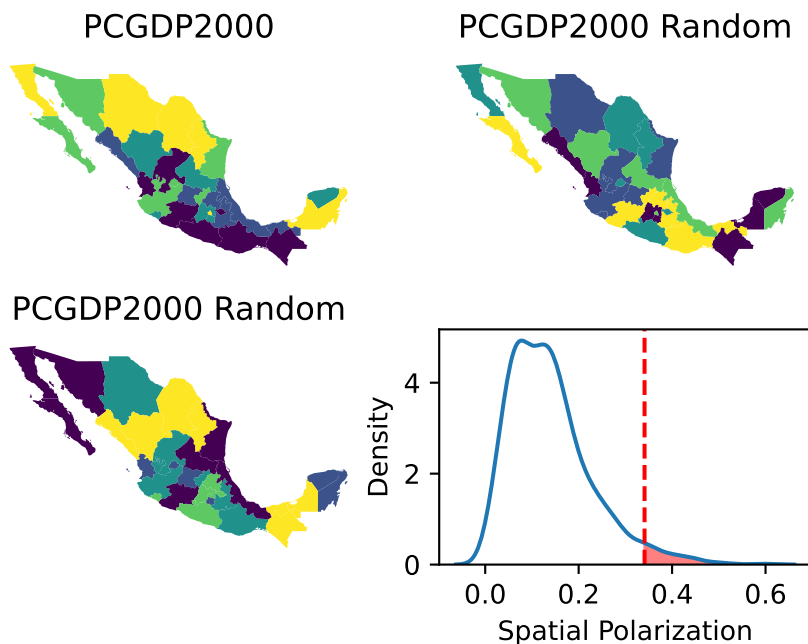


Figure 11: Spatial Polarization

The overall level of regional inequality is 0.106. The between region component stands at 0.036, while the within region element is 0.070. In relative terms, inequality between the regions in Mexico accounts for 34 percent of state income inequality, while inequality between states from the same region is the larger share presenting 72 percent of spatial inequality.

The ratio of between to within region inequality has been suggested as a measure of *spatial polarization* (Zhang and Kanbur 2001). Since the two components sum to overall inequality, we can re-express the measure of spatial polarization as the ratio of between region to overall inequality. Thus, the level of spatial polarization of state incomes stood at 34 percent in 2000 in Mexico.

A relevant question here is whether this level of spatial polarization in 2000 should be considered high or not? Another way to express this, is to ask if the polarization is higher than we would expect if state incomes were randomly distributed in space in Mexico in 2000.

We can answer this question by developing counter-factual spatial distributions that reflect a null hypothesis of spatially random income distributions. Given the n observations, in our case states, there are $n!$ permutations of incomes that are equally likely. In Figure 11, the maps in the top-right and lower-left are two such realizations, where the observed incomes from 2000 (top-left) have been randomly reassigned to states.

Since it is not feasible to generate all $32!$ maps in our case,⁶ we take a sample of 999 such maps from the distribution of permutations. For each of these maps we calculate the spatial polarization measure, and collect all measures to develop a reference distribution for our index under the null of spatially random income distributions. We evaluate our observed spatial polarization index against this distribution and derive a pseudo p-value as the number of counterfactual distributions that generate polarization levels as large as the observed value over the number of permutations plus one.

The reference distribution for the polarization index is reported in the bottom right of Figure 11. The area to the right of our observed polarization index of 0.34 is 0.036 of the distribution. By convention, this p-value points to a significant departure from the null, and we would conclude that the level of spatial polarization of incomes in Mexico is significantly different from that expected were incomes generated by a spatially random process.

The regional decomposition is a spatially explicit measure of spatial inequality as it is indeed sensitive to the geographical distribution of the income values. It is important to note, however, that it is the spatial polarization index, and not the overall T that is spatially explicit. Each of the counterfactual spatial distributions used to construct the reference distribution for the spatial polarization index would still have the same attribute distribution, with similar means and variances, and levels of overall inequality.

While the regional inequality decomposition is spatially explicit, it treats geography in an aggregated fashion. There are, in essence, two scales of spatial inequality implied in this decomposition. As displayed in Figure 10, the within-regions component might be considered the “local” measure as it compares incomes belonging to the same region to the mean income of the group. The between-regions inequality component, by contrast, views spatial inequality in a more aggregate fashion considering only the differences in regional means.

A close inspection of the decomposition Equation 1 reveals both the within and between inequality components are functions of a T index. In the former case this is applied to states from the same region, and in the latter case, the T is applied to the means of regional incomes. For the within region component, the actual spatial distribution of the member states within each of their regions is ignored. By the same token, the geographical location of the regions is immaterial to the calculation of the between region component. In other words, once the states have been assigned to regions, geography no longer matters. Thus, we draw a distinction between *regional* inequality decomposition and *spatial* inequality decomposition.

Spatial Inequality Decomposition

To capture these ignored dimensions of the spatial distribution, Rey and Smith (2013) suggested a spatial decomposition of the Gini coefficient. Starting from the Gini coefficient ex-

⁶ $32! = 2.6313084e + 35$.

pressed as half the relative mean absolute deviation:⁷

$$Gini = \sum_{i=1}^n \sum_{j=1}^n \frac{|y_i - y_j|}{2n^2 \bar{y}} \quad (2)$$

the sum of the absolute deviations can be decomposed as:

$$\sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} |y_i - y_j| + (1 - w_{i,j}) \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| \quad (3)$$

where $w_{i,j} = 1$ if states i and j are geographical neighbors, $w_{i,j} = 0$ otherwise. Here we define neighbors as states that share a border.

Substituting Equation 3 into Equation 2 gives:

$$Gini = \sum_{i=1}^n \sum_{j=1}^n \frac{w_{i,j} |y_i - y_j|}{2n^2 \bar{y}} + \sum_{i=1}^n \sum_{j=1}^n \frac{(1 - w_{i,j}) |y_i - y_j|}{2n^2 \bar{y}} \quad (4)$$

The first term represents a measure of inequality between neighboring states, while the second term captures inequality between “distant” pairs of states. For most spatial configurations, the number of neighboring pairs will be dwarfed by the number of pairs of states that are distant. So while, in the case of spatially clustered income values, the expectation would be for the average difference in incomes to be smaller for neighboring rather than distant states, our measure of spatial clustering here has to take into account the different cardinality of the two sets of pairs. As such, the relevant comparison is if the first term (second term) is smaller (larger) than what could be expected if state incomes were randomly distributed in space.

We apply the Spatial Gini Decomposition to Mexican State incomes in 2000 using `pysal-inequality` in Figure 12. The adjacency graph based on the criterion of Queen neighbors is shown in the top-right figure. An edge defines a pair of neighboring states.

```
fig, axs = plt.subplots(2, 2)
from inequality.gini import Gini_Spatial
import libpysal
np.random.seed(12345)

gdf.plot('PCGDP2000', ax=axs[0, 0], scheme='quantiles', cmap='viridis')
axs[0, 0].set_title('PCGDP2000')
axs[0, 0].axis('off')
```

⁷It is sometimes stated that the maximum value of the Gini in this form is 1 (e.g., Wang et al. 2024). This is technically incorrect, as in this form, the Gini has a range of $[0, (n-1)/n]$. As n grows larger, the upper bound approaches 1. Moreover, because the upper bound is a function of n , care should be taken when using the Gini in this form when comparing distributions of different sizes.

```

wq = libpysal.weights.Queen.from_dataframe(gdf)
wq.transform = 'B'
gs2000 = Gini_Spatial(gdf["PCGDP2000"], wq)

income = gdf['PCGDP1940']

gdf.plot(ax=axes[0, 1])
axes[0, 1].set_title('Neighbor Graph')
axes[0, 1].axis('off')

wq.plot(gdf, ax=axes[0,1])

axes[1, 0].set_title('Counterfactual')
axes[1, 0].axis('off')
gdf['PCGDP2000r'] = np.random.permutation(gdf.PCGDP2000)
gdf.plot(column='PCGDP2000r', ax=axes[1,0], scheme='quantiles',
         cmap='viridis')

adsum = gs2000.dtotal
realizations = gs2000.wcgp / adsum
kde = sns.kdeplot(realizations, fill=False, color='blue', ax=axes[1,1])
x, y = kde.get_lines()[0].get_data()
plt.legend([], [], frameon=False)
# Fill the area to the right of the specified value
plt.fill_between(x, y, where=(x >= gs2000.wcgp/adsum),
                interpolate=True, color='red', alpha=0.5)

# Add vertical line at the specified value
plt.axvline(x=gs2000.wcgp/adsum, color='red', linestyle='--')
plt.xlabel("Spatial Gini")

plt.tight_layout()
plt.show()
G = gs2000
msg = f'Expected Distant SADS/Total SADS: {G.e_wcgp/adsum:.2f}'
msg = f'{msg}, Observed: {G.wcgp/adsum:.2f}\n p-value: {G.p_sim}'
print(msg)

```

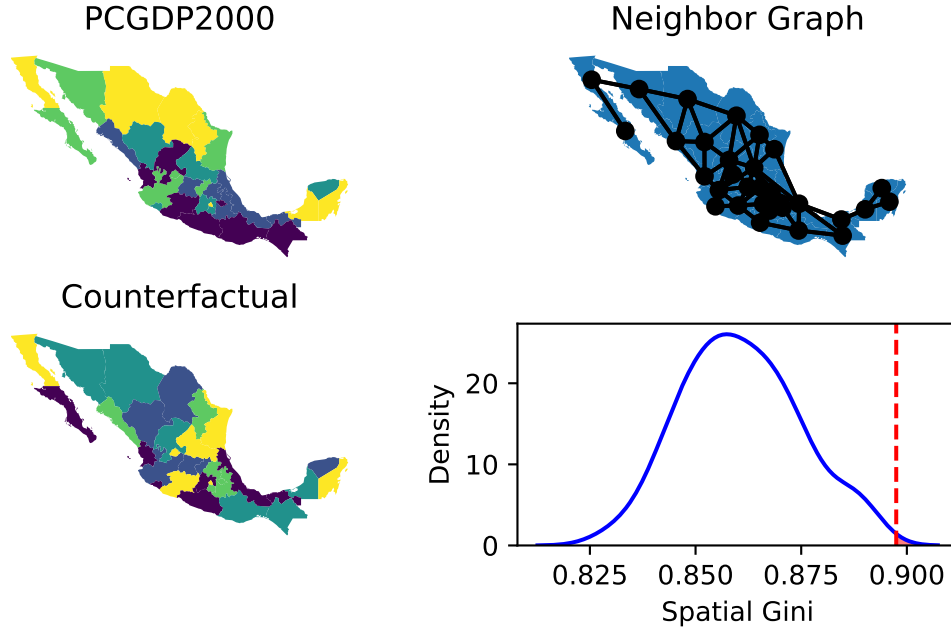


Figure 12: Spatial Inequality Decomposition

Expected Distant SADS/Total SADS: 0.86, Observed: 0.90
p-value: 0.01

For inference on the spatial Gini, the same computationally based approach that we used in the Theil decomposition is employed. One of the counterfactuals representing a random permutation of the incomes is shown in the lower-left figure. The reference distribution for the spatial Gini index is shown on the bottom right. Here the spatial Gini is expressed as the share of the overall absolute pair-wise differences due to inequality between distant (non-neighbor) pairs of states. The pseudo p-value (0.01) for the observed index is calculated as the area to its right under the distribution.

Under the null, the distant pairs should account for 86 percent of the absolute differences, however, the observed share is much higher at 90 percent. In fact, the p-value indicates that none of the counterfactual spatial distributions of income generated a spatial Gini as large as the one observed. In other words, the inequality between distant pairs of states is larger than the inequality between neighboring states. This pairwise orientation of the spatial Gini decomposition offers a useful complement over the regional decomposition of the Theil approach, as it introduces a more spatially explicit view of the income distribution that demonstrates how spatial autocorrelation affects overall inequality across the states.⁸

⁸For a recent extension of the spatial Gini decomposition see Panzera and Postiglione (2020).

It should be noted that the spatial Gini decomposition does not allow for the exact additivity of within-group and between-group inequality components. This is because the Gini index is affected by the degree of overlap in incomes of states belonging to different regions. Any overlap complicates the decomposition because part of the total inequality arises from the overlap itself, which cannot be easily attributed to either within-group or between-group inequality. So instead of a two-way additive decomposition, a residual term is required due to the overlap. Because this residual term is difficult to interpret, the Gini is seldom used for regional inequality decomposition.

Weighted or Unweighted Inequality: Places versus People

A final issue we explore is the question of whether the measure of spatial inequality should take into account the population sizes of the enumeration units. This was briefly mentioned earlier in the context of measuring international inequality. In the regional literature, a debate rages as to whether population weighted or unweighted approaches should be used to measure spatial inequality (Gluschenko 2018) .

To frame the debate, it is helpful to consider three different concepts of inequality at the international scale suggested by Milanović (2007). Here we adapt them to the question of measuring spatial inequality at the sub-national scale. Concept 1 is unweighted spatial inequality, where each state is one unit of measurement, and we use its per capita income irrespective of the state's population. In addition to being the dominant approach in regional inequality analysis, this concept is at the core of the literature on regional convergence (Rey and Montouri 1999) where the focus is on whether the incomes of poor and rich states in a system are coming together or growing apart over time.

Concept 2 takes into account the population of the individual states, recognizing that a state like Nuevo Leon with a population of 8.6 million in 2000 having its average income increase by 10,000 USD is likely to have a larger impact than is seeing the per capita income of Colima, with a population of 500,000 change by the same amount. So here, the per capita incomes of the states are now weighted by the population of the states. This is weighted spatial inequality.

Finally, Concept 3, measures inequality between all the individuals in the country. Here we would require information on the individuals both in terms of their incomes and their state of residence. Were it available, such data would allow us to measure personal income inequality. However, in the literature on regional inequality, such data is scarce, and so this concept is not operationalized.

This means the debate in the regional literature is between proponents of Concept 1 and Concept 2.

Thus far in this paper, we have adopted the Concept 1 definition of spatial inequality, that is, unweighted spatial inequality. We can contrast this with the perspective offered by Concept

2, by developing the Weighted Pengram under concept 2 shown in Figure 13 and comparing it with the Unweighted Pengram from Figure 3. The number of bars for each state in the Weighted Pengram is proportional to the population of that state. The logic behind the weighted spatial inequality is that rather than having one representative observation (individual) from the state with their level of income equal to the state per capita income, now there are a proportional number of individuals from each state having their level of income equal to the state's per capita income.

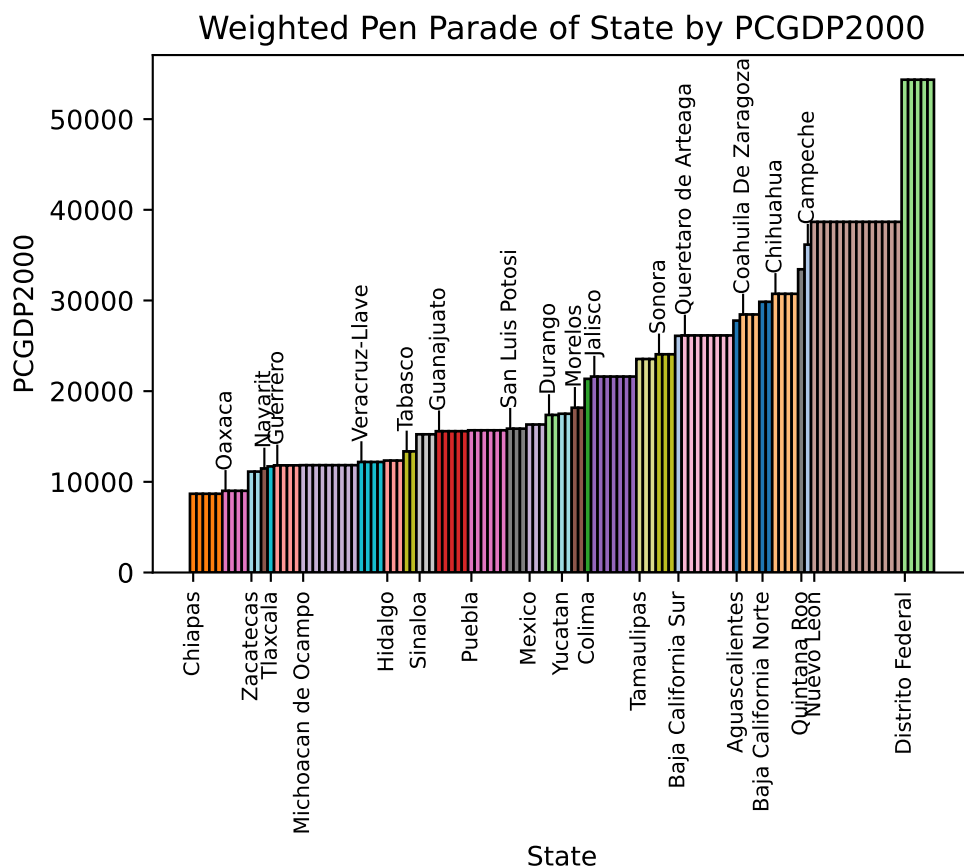


Figure 13: Populated Weighted Pengram

A fundamental issue with this approach is that it assumes that inequality among individuals within the state is zero, since all members have the same level of income. This also implies that for two states that have different levels of per capita income, the poorest members of the richer state will be richer than the richest member of the poorer state. This is at odds with empirical reality.

Conclusion

Regional inequality continues to attract the attention of researchers and policy makers alike. This chapter has presented the key methodological approaches available to researchers interested in analyzing spatial disparities. By highlighting the challenges posed in adapting classical inequality measures to the question of spatial inequality, the chapter draws important distinctions between different types of inequality decomposition approaches based on their treatment of space. These methods reveal the spatial dimensions of inequality, particularly the extent to which income disparities are geographically clustered.

While the focus has been on regional inequality at the sub-national scale, these methods have a broad scope that researchers can apply at different spatial scales, from analyzing geographical disparities at the international level (Redding and Venables 2004), inter-regional (Bathelt, Buchholz, and Storper 2024), as well as intra-urban scale (OECD 2018). This scope is vital, as the mechanisms of inequality can differ at each scale. For example, trade policy likely plays a more significant role on an international scale. At the same time, industrial restructuring is more influential on an inter-regional scale, and residential sorting is operative on an intra-urban scale. Furthermore, the resulting patterns of spatial inequality may also vary across these scales. The work by Ganong and Shoag (2017) shows a substantial divergence of incomes across states but reports a more mixed pattern when measuring convergence using labor market areas.

Linking the spatial inequality measures to possible policy interventions offers some possibilities. For example, the distinction between inter-regional (between-region) and intra-regional (within-region) inequality in the decomposition may help researchers and policymakers design such interventions in a tailored fashion. If between-region inequality dominates, there may be a strong case for place-based policies. Conversely, if the within-region inequality component accounts for the majority share, more nationally oriented policies, such as federal tax laws, may be more appropriate.

In future research, it would be beneficial to investigate the relationship between spatial inequality and place mobility. Place mobility refers to the economic trajectory of a location within the national context, similar to the concept of inter-generational income mobility (Rey and Casimiro Vieyra 2023). Exploring the link between place mobility and spatial polarization is crucial, especially how the movement of states within the income distribution impacts broader regional inequalities. Moreover, gaining insight into the interaction between place mobility and “place-based policies” could help in designing more effective regional development strategies that reduce inequality while fostering economic mobility.

Agnieszka, Widuto. 2019. “Regional Inequalities in the EU.” *European Parliamentary Research Service*.

Atkinson, Anthony B. 1970. “On the Measurement of Inequality.” *Journal of Economic Theory* 2 (3): 244–63. [https://doi.org/10.1016/0022-0531\(70\)90039-6](https://doi.org/10.1016/0022-0531(70)90039-6).

- Barber, Michael, and John B. Holbein. 2022. “400 Million Voting Records Show Profound Racial and Geographic Disparities in Voter Turnout in the United States.” *PLoS ONE* 17 (6): e0268134. <https://doi.org/10.1371/journal.pone.0268134>.
- Bathelt, Harald, Maximilian Buchholz, and Michael Storper. 2024. “The Nature, Causes, and Consequences of Inter-Regional Inequality.” *Journal of Economic Geography* 24 (3): 353–74. <https://doi.org/10.1093/jeg/lbae005>.
- Brandolini, Andrea, and Timothy M. Smeeding. 2011. *Income Inequality in Richer and OECD Countries*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199606061.013.0004>.
- Cowell, Frank A. 2011. *Measuring Inequality*. 3rd ed. LSE Perspectives in Economic Analysis. Oxford: Oxford University Press.
- Deb Nath, Nirmalendu, and Agricola Odoi. 2024. “Geographic Disparities and Temporal Changes of Diabetes-Related Mortality Risks in Florida: A Retrospective Study.” *PeerJ* 12 (June): e17408. <https://doi.org/10.7717/peerj.17408>.
- Frank, Mark W. 2009. “Inequality and Growth in the United States: Evidence from a New State-Level Panel of Income Inequality Measures.” *Economic Inquiry* 47 (1): 55–68.
- Ganong, Peter, and Daniel Shoag. 2017. “Why Has Regional Income Convergence in the U.S. Declined?” *Journal of Urban Economics* 102 (November): 76–90. <https://doi.org/10.1016/j.jue.2017.07.002>.
- Gaubert, Cecile, Patrick Kline, Damián Vergara, and Danny Yagan. 2021. “Trends in US Spatial Inequality: Concentrating Affluence and a Democratization of Poverty.” *AEA Papers and Proceedings* 111 (May): 520–25. <https://doi.org/10.1257/pandp.20211075>.
- Gluschenko, Konstantin. 2018. “Measuring Regional Inequality: To Weight or Not to Weight?” *Spatial Economic Analysis* 13 (1): 36–59. <https://doi.org/10.1080/17421772.2017.1343491>.
- Graetz, Nicholas, Lauren Woyczynski, Katherine F. Wilson, Jason B. Hall, Kalkidan Hassen Abate, Foad Abd-Allah, Oladimeji M. Adebayo, et al. 2020. “Mapping Disparities in Education Across Low- and Middle-Income Countries.” *Nature* 577 (7789): 235–38. <https://doi.org/10.1038/s41586-019-1872-1>.
- Hall, Robert E. 1978. “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence.” *Journal of Political Economy* 86 (6): 971–87. <https://doi.org/10.1086/260724>.
- Hanson, Gordon H. 1996. *U.S.-Mexico Integration and Regional Economies*. Cambridge: National Bureau of Economic Research.
- Jordahl, Kelsey, Joris Van Den Bossche, Jacob Wasserman, James McBride, Jeffrey Gerard, Jeff Tratner, Matthew Perry, et al. 2019. “Geopandas/Geopandas: V0.4.1.” Zenodo. <https://doi.org/10.5281/ZENODO.2585849>.
- Kanbur, S. M. R., and A. Venables. 2005. *Spatial Inequality and Development*. Oxford University Press.
- Khedmati Morasae, Esmaeil, Daniel W. Derbyshire, Payam Amini, and Tahera Ebrahimi. 2024. “Social Determinants of Spatial Inequalities in COVID-19 Outcomes Across England: A Multiscale Geographically Weighted Regression Analysis.” *SSM - Population Health* 25 (February): 101621. <https://doi.org/10.1016/j.ssmph.2024.101621>.
- Knaap, Elijah. 2017. “The Cartography of Opportunity: Spatial Data Science for Equitable

- Urban Policy.” *Housing Policy Debate* 27 (6): 913–40. <https://doi.org/10.1080/10511482.2017.1331930>.
- Ledić, Marko, Ivica Rubil, and Ivica Urban. 2023. “Tax Progressivity and Social Welfare with a Continuum of Inequality Views.” *International Tax and Public Finance* 30 (5): 1266–96. <https://doi.org/10.1007/s10797-022-09752-y>.
- Milanović, Branko. 2007. *Worlds Apart: Measuring International and Global Inequality*. Princeton (N.J.): Princeton university press.
- . 2018. *Global Inequality: A New Approach for the Age of Globalization*. First Harvard University Press paperback edition. Cambridge, Massachusetts London, England: The Belknap Press of Harvard University Press.
- Nijman, Jan, and Yehua Dennis Wei. 2020. “Urban Inequalities in the 21st Century Economy.” *Applied Geography* 117 (April): 102188. <https://doi.org/10.1016/j.apgeog.2020.102188>.
- OECD. 2018. *Divided Cities: Understanding Intra-urban Inequalities*. OECD. <https://doi.org/10.1787/9789264300385-en>.
- Overman, Henry G, and Xiaowei Xu. 2022. “Spatial Disparities Across Labour Markets.” Institute for Fiscal Studies.
- Panzer, Domenico, and Paolo Postiglione. 2020. “Measuring the Spatial Dimension of Regional Inequality: An Approach Based on the Gini Correlation Measure.” *Social Indicators Research* 148 (2): 379–94. <https://doi.org/10.1007/s11205-019-02208-7>.
- Partridge, Jamie S, Mark D Partridge, and Dan S Rickman. 1998. “State Patterns in Family Income Inequality.” *Contemporary Economic Policy* 16 (3): 277–94.
- Pen, J. 1971. *Income distribution*. London: Allen Lane.
- Piketty, Thomas, and Emmanuel Saez. 2003. “Income Inequality in the United States, 1913–1998.” *The Quarterly Journal of Economics* 118 (1): 1–39.
- Redding, Stephen, and Anthony J. Venables. 2004. “Economic Geography and International Inequality.” *Journal of International Economics* 62 (1): 53–82. <https://doi.org/10.1016/j.jinteco.2003.07.001>.
- Rey, Sergio J., Luc Anselin, Pedro Amaral, Dani Arribas-Bel, Renan Xavier Cortes, James David Gaboardi, Wei Kang, et al. 2022. “The PySAL Ecosystem: Philosophy and Implementation.” *Geographical Analysis* 54 (3): 467–87. <https://doi.org/10.1111/gean.12276>.
- Rey, Sergio J., Dani Arribas-Bel, and Levi John Wolf. 2023. *Geographic Data Science with Python*. Chapman & Hall/CRC Texts in Statistical Science. Boca Raton: CRC Press.
- Rey, Sergio J., and Esau Casimiro Vieyra. 2023. “Spatial Inequality and Place Mobility in Mexico: 2000–2015.” *Applied Geography* 152 (March): 102871. <https://doi.org/10.1016/j.apgeog.2023.102871>.
- Rey, Sergio J., and Brett D. Montouri. 1999. “US Regional Income Convergence: A Spatial Econometric Perspective.” *Regional Studies* 33 (2): 143–56.
- Rey, Sergio J., and Myrna L Sastré-Gutiérrez. 2010. “Interregional Inequality Dynamics in Mexico.” *Spatial Economic Analysis* 5 (3): 277–98.
- Rey, Sergio J., and Richard J Smith. 2013. “A Spatial Decomposition of the Gini Coefficient.” *Letters in Spatial and Resource Sciences* 6: 55–70.
- Roemer, John E. 1998. *Equality of Opportunity*. Cambridge, Mass.: Harvard Univ. Press.
- Sala-i-Martin, Xavier. 2006. “The World Distribution of Income: Falling Poverty and...

- Convergence, Period.” *The Quarterly Journal of Economics* 121 (2): 351–97. <https://www.jstor.org/stable/25098796>.
- Sen, Amartya. 2004. *Inequality Reexamined*. Reprint. New York: Oxford Univ. Press.
- Suss, Joel, Tom Kemeny, and Dylan S. Connor. 2024. “GEOWEALTH-US: Spatial Wealth Inequality Data for the United States, 1960–2020.” *Scientific Data* 11 (1): 253. <https://doi.org/10.1038/s41597-024-03059-9>.
- Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam: North Holland.
- Tine, Michele. 2017. “Growing up in Rural Vs. Urban Poverty: Contextual, Academic, and Cognitive Differences.” In *Poverty, Inequality and Policy*. IntechOpen. <https://doi.org/10.5772/intechopen.68581>.
- U.S. Department of Commerce. 2023. “Geographic Inequality on the Rise in the U.S.” *Regional Economic Research Initiative Blog*.
- Venter, Zander S., Helene Figari, Olve Krange, and Vegard Gundersen. 2023. “Environmental Justice in a Very Green City: Spatial Inequality in Exposure to Urban Nature, Air Pollution and Heat in Oslo, Norway.” *Science of The Total Environment* 858 (February): 160193. <https://doi.org/10.1016/j.scitotenv.2022.160193>.
- Wang, Jianying, Mei-Po Kwan, Gezhi Xiu, and Fangxu Deng. 2024. “A Robust Method for Evaluating the Potentials of 15-Minute Cities: Implications for Sustainable Urban Futures.” *Geography and Sustainability*, August, S2666683924000646. <https://doi.org/10.1016/j.geosus.2024.07.004>.
- Waskom, Michael. 2021. “Seaborn: Statistical Data Visualization.” *Journal of Open Source Software* 6 (60): 3021. <https://doi.org/10.21105/joss.03021>.
- Zhang, Xiaobo, and Ravi Kanbur. 2001. “What Difference Do Polarisation Measures Make?: An Application to China.” *Journal of Development Studies* 37: 85–98.