# Measuring Regional Inequality for Political Research

Dong-wook Lee*

Claremont Institute for Economic Policy Studies

Melissa Rogers†

Claremont Graduate University

**Abstract**

Political scientists are increasingly interested in the role that geography plays in shaping political preferences and outcomes. At the center of many of these questions is the existence of economic disparities across geographic political units, such as nations, sub-national regions, urban and rural areas, and electoral districts. In this research note, we briefly review recent research examining the relationship between economic and political geography to show its growing impact in the field of political science. Next, we discuss how existing measures of regional inequality from research in economics fit the concepts and units of political geography. We compare measures of regional economic inequality according to their theoretical and empirical properties and show their value in replications of existing scholarly research. Furthermore, we introduce a new measure of regional inequality that is consistent across the scope and scale of political units.

**Preliminary Draft. Comments Welcome.**

*Postdoctoral Research Fellow (`dong-wook.lee@cgu.edu`)

†Assistant Professor, Department of International Studies (`melissa.rogers@cgu.edu`)

# 1 Introduction

Until recently, geography has been "a blind spot for political scientists" (Rodden, 2010, p.322). Conceptual insights, empirical innovations, and breakthroughs in data collection have all brought geography into focus in the field. In particular, recent research has highlighted the importance of variation in economic geography to political phenomena. This research note provides a brief review of the economic geography research agenda in political science and discusses measures of regional economic inequality for use in political science.

Unlike previous research in the fields of economics and geography, political scientists are centrally concerned with what we label the "unit question" of how sub-national economic measures match relevant political jurisdictions. Most political science research offers theory that both political and economic processes are important to the outcome under examination. Thus, scholars are keenly interested in whether regional measures are accurate and stable representations of the political concept of interest.

The most common measures of regional inequality are aggregated, national values of sub-national GDP or income per capita used for cross-national comparison. The process of aggregation introduces complexity related to the comparison of regions across and within units, and other empirical concerns. Research on political geography requires informed choices about the analytical unit (region concept), the capture of spread (dispersion vs. concentration), and the analytic weights (e.g., population, vote share, land area) based on the theoretical construct.

We address these theoretical and empirical concerns of measuring regional inequality in the following sections. We first address the growing interest in regional inequality in political research. We then describe the conceptual and theoretical differences across existing measures and provide guidance on how to locate available data. We discuss the unit of analysis in detail, showing distinctions across unit levels in directly comparable data from the European Union's (EU) member countries. Finally, we offer a new scope and scale-independent measure of regional inequality, and compare existing measures in replications of recent region-focused political science research.

## 2 Politics, Economics, and Regions

Political science is experiencing a surge of interest in electoral and economic geography in theoretical research and empirical applications. Theoretically, geography is highly consequential in electoral and distributive politics, which are nearly always organized spatially. Moreover, economic clustering, both by firms and individuals, may shape political systems organized around geography. While these ideas are not new to the field, political scientists increasingly have empirical tools (data and methods) to explicitly consider geography in their analyses. At the center of this recent research is the measurement of economic concepts at politically-relevant units.

The rising interest in political and economic geography within the discipline of political science comes from a broad recognition of geographically-organized politics and its consequences (Weingast et al., 1981; Rehfeld, 2005; Rogers, 2015). For example, recent scholarship has pointed to economic geography in explanations of why widely accepted institutional hypotheses, such as those regarding the relationship between electoral rules and redistribution, do not always show anticipated results in empirical testing (Rickard, 2012; Jusko, 2015; Jurado & Leon, 2016; Menendez, 2016). In these accounts, it is the variation in voters' economic interests across electoral districts, not simply the institutional design of those districts, that help to explain differences in distribution. Similarly, research emerging primarily from U.S. politics looks to the geographic spread of voters to explain patterns of party success (Bishop, 2009; Rodden, 2010). Each of these authors points to the importance of the economics underlying political geography as their key explanatory factor.

Research on the EU countries has been particularly interested in the intersection of political and economic geography (Hooghe & Marks, 2001). In particular, contributions by Bolton and Roland (1997) and Beramendi (2007, 2012) highlight the inherent distributive conflict in political unions such as the EU with diverse economic geography. It also places economic disparities central to political conflict over secession, whether from the nation state or the EU (Sorens, 2005). Related research focuses on differences in governance quality across EU sub-national regions that is intimately linked to their historical economic development (Charron & Lapuente, 2013).

Growing interest in the relationship between political decentralization and political parties has also emphasized the relevance of regional disparities. Economists have examined whether decentralization (either economic or political) may exacerbate or reduce economic disparities (Lessmann, 2009, 2012; Rodríguez-Pose & Ezcurra, 2009). Political scientists have sometimes reversed this question, asking how regional disparities may shape political outcomes (Beramendi, 2007; Toubeau & Wagner, 2015; Lee & Rogers, 2016). Research into the role of institutions in party system nationalization have also considered the relevance of economic geography (Chhibber & Kollman, 2009; Brancati, 2008; Crisp, Olivella, & Potter, 2013).

Regional economic inequality is also increasingly considered as a relevant parameter predicting civil conflict (Buhaug & Gates, 2002; Lessmann, 2015). Civil conflict tends to be a geographically-bounded phenomenon that requires intense coordination amongst like-minded or similarly-aggrieved combatants (Buhaug et al., 2011; Østby et al., 2009; Raleigh & Hegre, 2009).

Scholars are taking geography more seriously in statistical analyses as well. Spatial concerns are now paramount in many political questions (Franzese & Hays, 2007). Among the many applications of spatial interdependence are fiscal competition amongst nearby nations and states (Tiebout, 1956) and the geographic spread of civil conflict (Buhaug & Lujala, 2005). Scholars are expected to consider the relevance of geographic interdependence, and to model it explicitly in econometric designs. In the next section we detail the advances in data availability on economic geography and regional disparity, and compare existing approaches for their theoretical and empirical strengths.

## 3 Issues in Measuring Regional Inequality

### 3.1 Data Availability

Governments increasingly make data available to measure the extent of the regional divide (Ezcurra & Rodríguez-Pose, 2009). Advanced industrial democracies lead the way by providing very high quality data on economic variables and related concepts at multiple geographic levels. The U.S.

Census Bureau, for example, consistently releases economic indicators for a wide range of geographic aggregations, including the nation, state, county, city, metropolitan statistical area, and electoral district, among others. As part of their coordination within the EU, member countries comply with statistical standards to calculate economic variables at four structural levels: NUTS3 (Parish, Canton, Oblast, City & Regency, County, or Municipality), NUTS2 (Region, State, Province, or Prefecture), NUTS1 (Region, Group of NUTS2), and NUTS0 (Country).[1] These data are widely used in research on the EU, with coverage in most cases since 1970 and in Eastern Europe from 1990 or 1995. OECD nations also provide sub-national data, typically at the equivalent of NUTS2 level in their "Regional Statistics" category.

Outside of the most developed nations, many national statistical agencies also provide data on sub-national regions, typically at the equivalent of the NUTS2 level.[2] These data are most commonly available for federal countries, and those outside of the Middle East and Sub-Saharan Africa. See Online Appendices 1-2 for a summary of available regional GDP data and a global outlook of the intra-country variance in terms of regional GDP disparities at NUTS2-level.

The most common measures of regional inequality use sub-national GDP or GDP per capita values. These are the most widely available data across countries and over time. A smaller set of studies employ sub-national income data, typically from the Luxembourg Income Study (Mahler, 2002). Income data should be employed with particular caution in subnational studies because they may not be directly comparable across countries, regions, or over time. The concepts of economic productivity and income are distinct, yet many scholars will need to rely on GDP data to capture relative incomes given much greater data availability for the latter concept (Gennaioli, La Porta, De Silanes, & Shleifer, 2014).

Data availability can dramatically affect the size of the sample available for examination. This may lead to misleading inferences in comparative contexts. Cross-national analyses using region-

[1] NUTS (Nomenclature of Territorial Units for Statistics) is a geocode standard established by the European Commission's Eurostat (http://ec.europa.eu/eurostat).

[2] Where countries do not provide region level data, or where there is reason to question the quality of provided data, researchers are beginning to look at alternative sources such as the satellite nighttime light data (e.g., Lessmann, 2015; Hodler & Raschky, 2014; Harbers, 2015; Marx & Rogers, 2016).

level economic data are likely to "oversample" affluent nations (Lessmann, 2015). Given that affluent nations have, on average, much lower levels of regional economic disparity than poorer nations, this sample bias may lead scholars to accept or dismiss hypotheses too readily based on the available data.[3]

### 3.2 Reliability

The most important standard for reliable cross-national (or cross-state) regional inequality measures are: they should be independent of the number of regions considered, robust to population size, insensitive to shifts in average GDP levels, and satisfy the Pigou-Dalton transfer principle (Lessmann, 2012). This principle says that an arithmetical transfer from rich to poor regions reduces inequality (see for detail in, Dalton, 1920; Pigou, 1912; Ezcurra & Rodríguez-Pose, 2009). Each of the measures we consider below meet these criteria.

As with all government-produced data, scholars should thoughtfully consider the accuracy of their measures. Regional statistics, when used as evidence of regional leaders' accomplishments or when considered as criteria for centralized inter-regional distribution, might reasonably reflect political bias. The direction of that bias is not entirely obvious. While some leaders, as pointed out by Wallace (2016), have a clear reason to inflate their statistics, other regional actors may have incentives to underestimate their calculations (Kerner, Jerven, & Beatty, 2015).[4] Even where concerns with the politicization of data are minimal, countries surely vary widely in the quality (error) of regional data. Yearly estimates of regional GDP may be extrapolated from the previous economic census, limiting the year to year reliability of the data. Accordingly, the source of these data, their potential problems, and ways to mitigate those problems should be explicitly discussed by scholars.

The interpretative value of regional economic statistics may also be challenged by commuting and the cost of living. In the smaller nations of central Europe such as Denmark and the Nether-

[3] See Appendix Figure 1 for a global map of regional inequality of GDP per capita using available data for the RDGINI measure.

[4] See Alt, Lassen, and Wehner (2014) for skepticism about EU economic accounts data.

lands, for example, large percentages of individuals commute outside of their residential location for work.[5] Economic data based on either residential location (such as income surveys) or based on labor-output (such as productivity statistics) can give skewed impressions of region level conditions in some nations and clusters of nations (Spiezia, 2002). Region-level cost of living concerns are also relevant. Income surveys that do not adjust for cost of housing and other basic needs can overestimate regional inequality in standards of living. Scholars should consider whether these biases are likely in their sample and relevant to the theoretical questions, and adjust the calculations of income accordingly (see, e.g. Gennaioli et al., 2014).

### 3.3 Dispersion versus Concentration

Regional inequality can come in many forms, depending on distinct distributions of regional economic output. The two most popular concepts of regional inequality are the dispersion, of overall spread in economic productivity, and the concentration of productivity in particular regions. Dispersion and concentration are related but distinct concepts. Standard measures of regional inequality, such as coefficients of variance or the region adjusted gini coefficient (detailed below), capture economic dispersion. The distributions may take different shapes that can be compared in scales such as skewedness and kurtosis. Most often these comparative statics assume the unimodality of the distribution to examine whether they are left or right skewed (indicating more equal or less equal regions).

Some theoretical questions may be less concerned with dispersion, and more interested in how sub-national regions cluster. In particular, many studies focus on the geographic concentration of economic activity – where relatively small areas of the national territory account for large percentages of national production (Spiezia, 2002). In situations of concentration, certain regions would stand apart from the majority of regions. The relationship between concentration and dispersion is unclear. It may be that one region stands apart, but the remaining regions are very similar in (low) productivity. Measures of dispersion would capture, to some extent, the separation of the highly

[5]The same concern is relevant in multi-state metro areas in the U.S.A., such as Washington D.C. and New York City.

productive region from the cluster of less productive regions but the overall dispersion value would be relatively low. This appears to fit the case of Spain, shown in Figure 1 below. If scholars believe that concentration is an important phenomenon aside from dispersion, they require distinct, appropriate measurement concepts (Spiezia, 2002; Jurado & Leon, 2016; Chen & Rodden, 2013).

### 3.4 Political Units

When measuring regional inequality, scholars need to make choices about the appropriate unit of analysis. In many cases, the units (countries, states / provinces / regions, counties, cities) will emerge clearly from the research question and theory. For example, research on distributive politics and party system nationalization looks to the electoral district to evaluate spatial variation. Research on fiscal federalism focuses on the sub-national administrative levels (state, municipal) as the relevant political locus. However, when it comes to an empirical discussion about the national political outcomes affected by sub-national geography, the units are not always obvious.[6] To take the (extreme) US case, many overlapping jurisdictions could be appropriate units to national outcomes, such as the state (Senate electoral district), the congressional electoral district, the city, the county, and the regional associations of government (in cases of transportation policy). Any or all may be important to a particular political outcome. Researchers need to pay particular attention to what level of analysis best represents their political question.

Even where the unit may be obvious, available data may not match the theoretical unit. A vexing empirical problem for researchers studying comparative or American politics, in particular, is variability between electoral districts and statistical units. Scholars may have theories about distributive politics related to electoral targeting of resources that depend on economic geography (Rickard, 2012; Jurado & Leon, 2016; Menendez, 2016). However, standardized geo-codes for statistical purposes may not always match electoral districts. In the European context, for example, electoral districts may be equivalent to or subsets of NUTS2 or NUTS3 districts in most na-

[6] Underlying the unit question is also a theoretical concern. Geography may matter because of geographic sorting–similar people choose to live in the same area (Bishop, 2009). Or the place itself has a causal effect as people in the same area are exposed to common influences, and as a result exhibit similarities (group-dependent models).

tions, but the economic geography data will not always correspond precisely, particularly for lower house or sub-national electoral districts. In Slovenia, for example, electoral districts span NUTS2 and NUTS3 identifiers, making it nearly impossible to assign economic or socio-demographic indicators to that geographic unit. Moreover, geographic specifications may change over time, such as US House of Representative districts. In other cases, the reporting may be different across the same measure. For instance, within Luxembourg Income Study data, Belgium sometimes reports NUTS2 level aggregation, and sometimes NUTS1 level, eliminating the possibility of direct comparisons of LIS data at the region level over time.

Research on "neighborhood effects" has long recognized inherent challenges in identifying and measuring geographic concepts consistently. For many political questions based on clearly identifiable political units, this issue will be minor. However, research on topics such as dissemination of political ideology or local spatial effects on voting, for example, the unit is not so obvious (Fortunato, Swift, & Williams, 2016). Research of this type must make assumptions about what is an individuals' "neighborhood." These assumptions are likely to vary considerably across individuals, and not always in predictable ways (Kwan, 2012). Work in political geography must directly address this potential for ecological fallacy. It is not always obvious what the "true" causal unit of geography would be. In these cases, an association between a geographically defined variable and an outcome may be spurious or an incorrectly defined geography might lead to an unverifiable false negative.

A related, but distinct, concern in geographic analysis is the Modifiable Areal Unit Problem (MAUP). The simple description of MAUP is that different aggregations of the same data will lead to very different conclusions about that geographic area (Wong, 2009). A clear, politics-relevant example of MAUP is gerrymandering, which is an example of the "zoning problem" of MAUP. Electoral gerrymandering is possible because of spatial grouping of similar individuals, namely voters more inclined to support one party or the other. Accordingly, politicians can design electoral districts that are more favorable to one party or the other simply by reorganizing district aggregation without altering the underlying demographics of the state. Scholars may also face problems when drawing boundaries that create biased or arbitrary units of measure. The zoning problem is most severe when there is severe spatial autocorrelation in the concept of interest (income, partisanship, race, health-factors, etc.).

The scale problem is another concern with MAUP. It is the tendency for data to normalize (reducing the influence of outliers by taking median and mean values) at greater levels of aggregation. For example, the GDP per capita of California is reasonably similar to the GDP per capita of Minnesota. This is a reasonable assessment of the states as a whole, but it obscures tremendous variation at the sub-state level, particularly in California. The GDP per capita numbers appear similar because Sicilon Valley, San Francisco, and Los Angeles, are averaged with low GDP per capita central California areas. Minnesota also has within-state productivity dispersion, but it is not nearly as dramatic as California's. Whether the scale problem is an issue depends fundamentally on the research question. If the unit of analysis is US Senate elections (which are based at the US state level) the scale problem is not a concern because the election matches state-level data. However, if scholars use state-level data to assess House congressional districts, the state level aggregation will be a more accurate predictor across the six districts of Minnesota than it will be across California's fifty-three individual districts. The smaller scale will more accurately capture variation, but the unit should match the theoretical concept.

## 4 Existing Measures of Regional Inequality

Existing studies have adopted several measures of cross-nationally comparable, time-varying indicators of regional inequality. These measures capture different conceptual features of regional inequality. The formulas for calculating regional inequality vary depending on three factors discussed above: 1) the choice of economic indicator (productivity versus income); 2) the use of analytic weights; 3) the applicable concepts of geographic spread in economic properties (dispersion versus concentration). These measures vary according to these three factors.[7]

[7]We present all measures with economic values of GDP calibrated on a per capita basis. We use GDP for consistency across examples. Political scientists are most often concerned with per capita measures of regional GDP rather than aggregate GDP. These measures can be easily adapted to

The most simple, easy to interpret, regional inequality measure is the coefficient of variation (COV). COV is a dispersion measure without analytical weights and is constructed as follows:

$$\text{COV} = \frac{1}{\bar{y}} \left( \frac{1}{n} \sum\nolimits_{i=1}^{n} (\bar{y} - y_i)^2 \right)^{1/2} \quad (1)$$

where $\bar{y}$ denotes the country's average GDP per capita, $y_i$ is per capita GDP of region $i$, and $n$ is the number of regional units. COV is a widely used measure in the literature on regional economic growth and convergence (Barro & Sala-i Martin, 1992; Sala-i Martin, 1996).

Political scientists will often require a regional inequality measure that scales to important political features, such as population or land area. It may be important to know in research on civil conflict, for example, that a highly populated region has a very low GDP per capita or a sparsly populated region has very high per capita GDP. The COV measure would give the same analytical weight to both regions. A population weighted coefficient of variance (WCOV) adjusts for these differences in population according to the following formula:

$$\text{WCOV} = \frac{1}{\bar{y}} \left( \sum\nolimits_{i=1}^{n} p_i (\bar{y} - y_i)^2 \right)^{1/2} \quad (2)$$

where $p_i$ represents the share of the country's total population in the region $i$. Values of WCOV are calculated as the ratio of the standard deviation to the mean ($\bar{y}$) when $p_i$ is equally applied to every region (i.e., 1/n) within a country. However, when assigning a different value of $p_i$ to these regions according to their proportion of population, it is mean-independent. WCOV is thus robust against single extreme observations, in addition to satisfying the Pigou-Dalton transfer principle (Dalton, 1920; Pigou, 1912).

Similar to COV or WCOV, the measure of region-adjusted Gini Coefficient (RDGINI) is popularly used as an alternative to the dispersion of income spread across sub-national regions. Unlike COV or WCOV, RDGINI retains meaningful information about the extent of relative deprivation,

other economic concepts and values.

not merely economic spread. In RDGINI, additional weight is given to a region's per capita productivity as it veers father away from the mean of the inter-regional wealth distribution. This weighted value makes the inequality measure more sensitive to changes in the upper or lower tail of this distribution. RDGINI is calculated as follows:

$$\text{RDGINI} = \frac{2\sum_{i=1}^{n} iy_i}{n\sum_{i=1}^{n} y_i} - \frac{n}{n-1} \tag{3}$$

Existing research in economics and political geography that employs the COV, WCOV, and RDGINI typically relies on the most widely available regional unit, NUTS2, and assumes that the variance in the number of units has no effect on the measures. Nations vary considerably in how many regional units exist at this level (see Table 1 for variation in our EU sample across countries and across NUTS2 and NUT3 levels.) This assumption may be misleading: it would require that there is no within-unit variation (Bochsler, 2010). We build upon existing research on the measurement of party system nationalization (which has similar unit concerns) to construct a measure that is standardized across the number of units (see Bochsler, 2010, p.163). We calculate a new standardized gini indicator based on the formula developed by Bochsler (2010) to measure regional inequality (a modification of RDGINI), independent of the number political units considered. Thus, at least in theory, this measure should not be sensitive to the size or number of regions involved. We demonstrate in the replications below that this indicator is the most stable available across different political units and nations. The standardized adjusted Gini coefficient of regional GDP, labeled INEQ_SPNS, is constructed as follows:

$$\text{INEQ\_SPNS} = 1 - \left( \frac{2\sum_{i=1}^{n}(p_i(\sum_{j=1}^{i} y_j - \frac{y_i}{2}))}{\sum_{i=1}^{n} p_i \sum_{i=1}^{n} y_i} \right)^{1/\log\left(\frac{(\sum_{i=1}^{n} p_i)^2}{\sum_{i=1}^{n} p_i^2}\right)} \tag{4}$$

where $y_i$ is for a region's GDP and $y_j$ is thus the accumulated proportion of regional GDP. In the equation, the value subtracted from 1 represents portion of regional equality, equivalent to the measure of ‘party nationalization’ developed by Bochsler (2010). We do this subtraction to obtain the residual commensurate with the measure of regional inequality. This modification creates an cross-nationally comparable regional inequality index ranging from 0 to 1, where the larger the value, the higher the level of regional inequality.

The measures described above–COV, WCOV, RDGINI, and INEQ_SPNS–are dispersion measures that weight elements differently. The adjusted geographic concentration (AGC) is a measure of economic concentration that captures whether economic productivity is disproportionately held in one or a small number of regions (Spiezia, 2002). This AGC is constructed as follows:[8]

$$\text{AGC} = \frac{\sum_{i=1}^{n} |y_i - a_i|}{2(1 - a_{min})} \tag{5}$$

where $a_i$ is the area of the region i as a share of the country area. AGC is an index ranging from 0 to 1, with the higher value being more concentration of national GDP in a certain administrative sub-national region. It uses both the economic and geographic weight over all regions within a country, thus accounting for within and between country differences in the size of regions.

Figure 1 provides a general idea of country distributions according to WCOV, INEQ_SPNS, and AGC with data from the NUTS2 level. All data for this analysis are from Cambridge Econometrics and Eurostat. In this figure we use the GDP concept (GDP per capita, population-weighted GDP per capita, share of GDP) relevant to their respective measures. For each measure, we have chosen countries with low values (top row) and high values (bottom row) to illustrate the distinct distributions the measures capture. In the first column, we include the Netherlands and the UK. The Netherlands has a relatively narrower distribution of GDP per capita (all regions between 25000€ and 45000€) in comparison with the UK (20000€ to over 80000€). The frequencies have been weighted by population to demonstrate the WCOV concept. The highly populated Dutch regions are near the middle of the country distribution. In comparison, the more of the UK’s population is located in regions closer to the bottom of the highly dispersed distribution. In column 2 we compare the Netherlands and Belgium, representing variation in INEQ_SPNS. Again, the Netherlands

[8]Chen and Rodden (2013) and Jurado and Leon (2016), for example, use this measure.

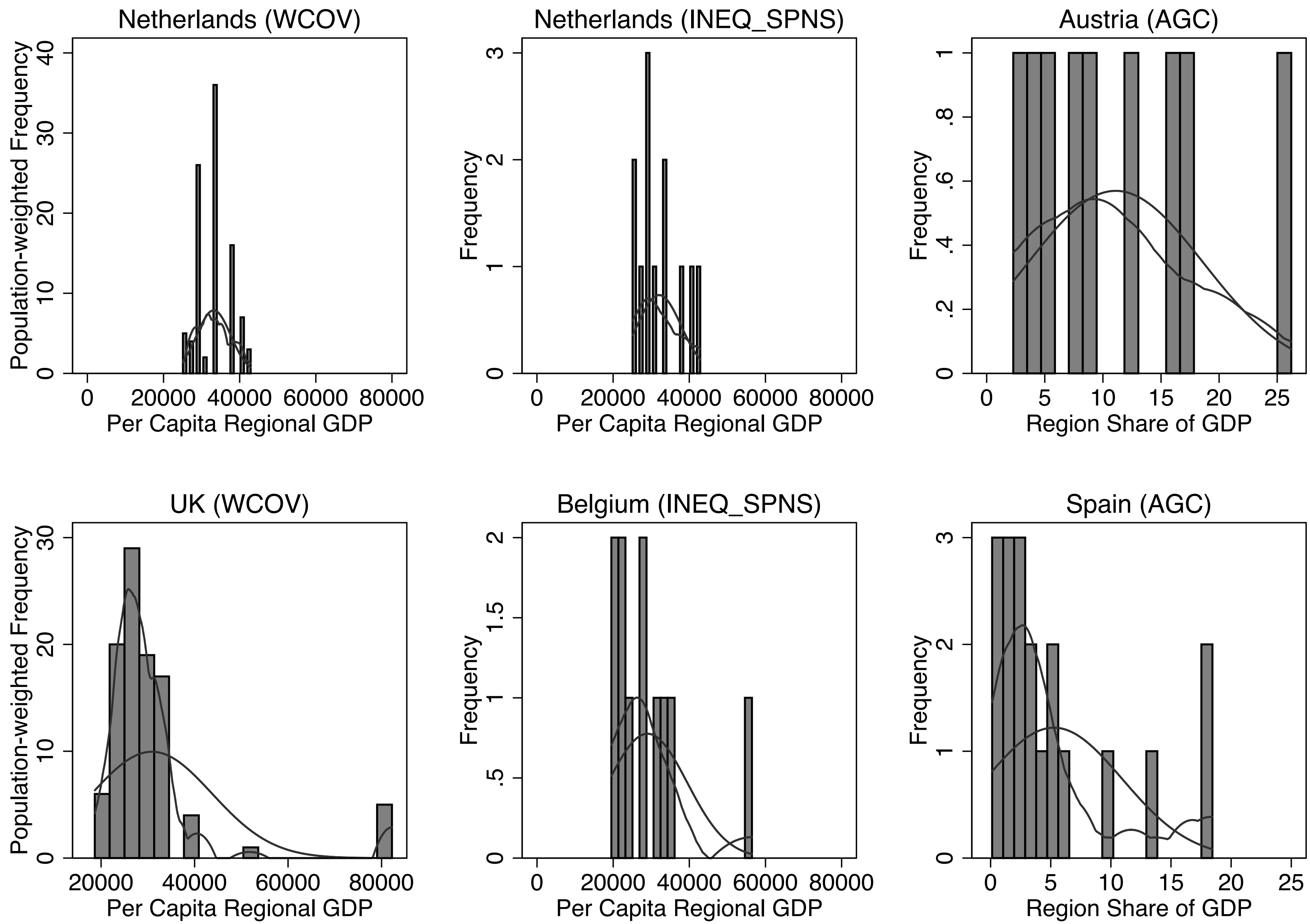



Figure 1: Comparison of Dispersion and Concentration Measures

has a very compact distribution (this time the frequency is unweighted). Belgium's regions are more dispersed, and one region (Brussels) is much more productive than the other regions. Because INEQ_SPNS weights more heavily at the ends of the distribution, Belgium has a high value on this measure. Column 3 shows variation on AGC, which compares the concentration in regions' share of GDP. Both Austria and Spain have 1 or 2 regions that produce a substantial percentage of GDP (18-25%). However, the majority of Spain's regions produce very little (>5%) of GDP while most of Austria's regions produce over 5% of the nation's GDP.

## 5 Country Variation in Regional Units

Each of the summarized regional inequality measures has analytic value, depending on the theoretical question at hand. In this section we point out some challenges to the stability and interpretation of these measures based on the "unit question."

Table 1 summarizes cross-national differences in the number of regions at the NUTS2 and NUTS3 levels in our sample.[9] The observed variation both across countries and across unit levels is substantial. The number of NUTS2 regions in Germany is more than 15 times larger than that in Ireland and Slovenia. The size of unit differences across Germany's NUTS2 and NUTS3 is also very large, as shown in column 3's calculation of the difference across units. This is noteworthy, in part, because NUTS2 data is commonly substituted for NUTS3 data in years in which NUTS3 are not available (and vice versa) in datasets in which one of the levels is missing.

The concern with variation across the number of units may be minor if it does not influence the aggregation measures. We examine this question in Figure 1, which compares regional inequality measures (dispersion and concentration), depending on the level of geographic aggregation for Germany, Belgium, and Sweden. These countries were selected to provide variation in difference between the number of NUTS2 and NUTS3 regions (column 3 of Table 1). Conceptually, the regional inequality trend of NUTS2 should be directly comparable with that of its subset data at

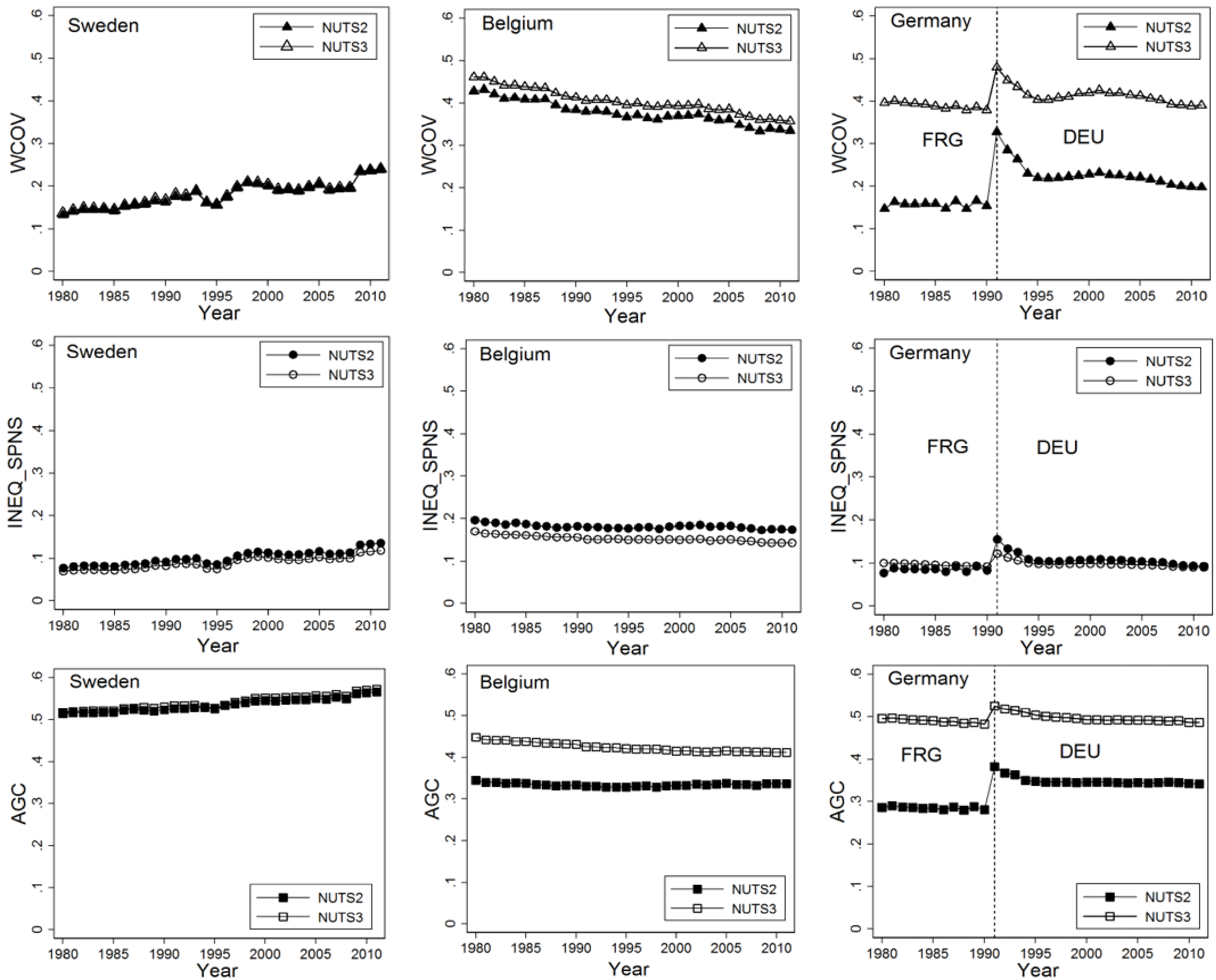
[9] We also report the comparative measures of regional inequality by separating the NUTS2 level from the NUTS3 level. This comparison report can be found in Online Appendix 3.

Table 1: The Number of Political Units by NUTS, 2011

| Countries | NUTS2 | NUTS3 | Ranked by Difference |
|---|---|---|---|
| Germany | 31 | 326 | 295 |
| United Kingdom | 41 | 133 | 92 |
| Italy | 21 | 107 | 86 |
| France | 26 | 100 | 74 |
| Poland | 16 | 66 | 50 |
| Spain | 19 | 59 | 40 |
| Greece | 13 | 51 | 38 |
| Romania | 8 | 42 | 34 |
| Belgium | 11 | 44 | 33 |
| Netherlands | 12 | 40 | 28 |
| Austria | 9 | 35 | 26 |
| Portugal | 7 | 30 | 23 |
| Bulgaria | 6 | 28 | 22 |
| Finland | 5 | 20 | 15 |
| Hungary | 7 | 20 | 13 |
| Sweden | 8 | 21 | 13 |
| Norway | 7 | 19 | 12 |
| Slovenia | 2 | 12 | 10 |
| Czech Republic | 8 | 14 | 6 |
| Denmark | 5 | 11 | 6 |
| Ireland | 2 | 8 | 6 |
| Slovenia | 4 | 8 | 4 |

Notes: The NUTS classification (Nomenclature of Territorial Units for Statistics) is defined on the basis of administrative divisions applied to the Eurostat member countries.

Figure 2: Measures of Regional Inequality by the Size of Political Units



the NUTS3 level.[10] So long as these values are calculated using the same measure of regional inequality, their trends and values should be similar.

Figure 2 shows that unit level choices in aggregation measures may influence how we evaluate cross-country inequality. We demonstrate these concerns with three cases, Belgium (33), Germany (295), and Sweden (13), that vary considerably in difference in the regional units across NUTS levels (see Table 1). Figure 1 shows that, at least for the three indicators shown (WCOV, INEQ_SPNS, AGC), Sweden's measures are very stable across NUTS2 and NUTS3 levels. NUTS2 and NUTS3 in the Swedish case are relatively similar in size and trends. The unit challenge becomes more apparent in the Belgian case. While the trends across the two units in each of the three measures are similar, their levels vary. In particular, the interpretation of the Belgium's regional concentration level would be potentially very different if measured at the NUTS3 level (high) or NUTS2 level (moderate). This may reflect theoretical differences across the levels, or more likely may reflect differences in the statistical aggregation of the levels and the number of units (MAUP). Germany shows the biggest difference across NUTS2 and NUTS3, with big diverence most apparent in WCOV and AGC. The choice of unit can have a significant impact on the measures in question.

As shown in the center row of Figure 2, the INEQ_SPNS measure is more consistent in all three nations across the NUTS2 and NUTS3 levels. INEQ-SPNS is designed to provide a measure for regional inequality that is not influenced by the number or the size of the units. When scholars do not have a specific regional concept in their theory, or where data are not available at for the unit concept of interest, INEQ-SPNS is likely to provide a more reliable measure of regional dispersion.

## 6 Regional Inequality in Relevant Studies

To demonstrate the properties of the common regional inequality measures, and to show their value in emerging research, we replicate results from three relatively recent published, high impact articles focused on regionalism and economic and political geography.[11] We chose these three

[10]NUTS3 regions are always subsets of NUTS2 regions.
[11]Replication data are available at melissazrogers.com

papers for their variation in the dependent variables, variation in their samples, and their variation in methodology. We intend to demonstrate that regional inequality is relevant to these studies and show how the different measures perform, based on theoretical expectations and the unit of analysis. For each study, we show results with dispersion and concentration measures calculated at two regional units (NUTS2 and NUTS3). Using NUTS2 and NUTS3, which are produced from the same data, allows direct comparison of the effect of the unit. Data constraints in our regional inequality measures mean the original samples could not be fully preserved, but the general findings are universally preserved in the replicated studies.

First, we replicated Toubeau and Wagner (2015), which studies how national parties' left-right ideological position on the economic or cultural dimensions may affect their preferences for decentralized government policy decision making. They argue that national parties placing a higher value on market efficiency (economically right wing parties) are more likely to advocate decentralized decision making. Parties emphasizing national unity and territorial integrity (culturally conservative parties), on the other hand, tend to prefer centralization.

Table 2 presents our replication results for (Toubeau & Wagner, 2015). Their dependent variable is party positions on decentralization and their econometric approach is multi-level (party and country level) modeling. M1 is our replication of their basic finding, with the sample reduced to match our data on regional inequality. M1 and all additional replications (M2-M7) show strong support for their basic findings.

In Toubeau and Wagner's basic model (M1), they include the COV in regional GDP per capita, with data from the NUTS2 level. Their results show a significant, positive effect of dispersion on party positions in favor of decentralization. We add to their model three other regional inequality measures (two dispersion, one concentration), measured at both the NUTS2 and NUTS3 levels. The idea is to examine how the different indicators and different units hold across the same specification.

In theory, regional inequality may influence party preferences for decentralization through its effects on fiscal structures or on the party system (Beramendi, 2012). Centralization is a subject

Table 2: Multilevel Linear Regression Model Predicting the Effect of Party Positions on Decentralization (Toubeau & Wagner, 2015)

| | M1 Toubeau & Wagner (2015) | M2 Extension | M3 Extension | M4 Extension | M5 Extension | M6 Extension | M7 Extension |
|---|---|---|---|---|---|---|---|
| Ideological scales (0-20) | | | | | | | |
| Economically right wing party | 0.215** | 0.215** | 0.216** | 0.207** | 0.203** | 0.202** | 0.207** |
| | (0.094) | (0.094) | (0.095) | (0.094) | (0.094) | (0.097) | (0.097) |
| Culturally conservative party | -0.330*** | -0.331*** | -0.319*** | -0.327*** | -0.325*** | -0.301*** | -0.310*** |
| | (0.061) | (0.061) | (0.061) | (0.061) | (0.061) | (0.062) | (0.063) |
| *Party-level Controls* | | | | | | | |
| Regionalist party distance | | | | | | | |
| Economic dimension | -0.375*** | -0.374*** | -0.373** | -0.382*** | -0.385*** | -0.387*** | -0.382** |
| | (0.144) | (0.143) | (0.145) | (0.144) | (0.143) | (0.149) | (0.149) |
| Cultural dimension | -0.002 | -0.004 | 0.007 | -0.001 | -0.001 | 0.026 | 0.022 |
| | (0.082) | (0.082) | (0.083) | (0.082) | (0.082) | (0.084) | (0.084) |
| Vote share | -3.354 | -3.398 | -4.187 | -3.379 | -3.433 | -4.628 | -3.984 |
| | (3.670) | (3.653) | (3.674) | (3.666) | (3.660) | (3.808) | (3.829) |
| Participation in the national government | -0.700 | -0.719 | -0.623 | -0.699 | -0.701 | -0.444 | -0.467 |
| | (0.837) | (0.834) | (0.841) | (0.836) | (0.835) | (0.857) | (0.859) |
| Mean government / opposition position | -0.159 | -0.185 | -0.119 | -0.200 | -0.225 | 0.014 | 0.066 |
| | (0.211) | (0.213) | (0.209) | (0.219) | (0.225) | (0.206) | (0.195) |
| *Country-level Controls* | | | | | | | |
| Level of self-rule | 0.055 | 0.069 | 0.141 | 0.049 | 0.088 | -0.019 | 0.027 |
| | (0.062) | (0.063) | (0.090) | (0.061) | (0.068) | (0.083) | (0.066) |
| Regionally based ethnic group | -0.082 | -0.556 | -0.925 | 0.808 | 0.614 | -0.077 | 0.946 |
| | (1.314) | (1.368) | (1.513) | (1.281) | (1.276) | (1.786) | (1.489) |
| Area (logged) | 1.673*** | 2.194*** | 2.366*** | 1.324*** | 1.890*** | 0.613 | 1.047** |
| | (0.551) | (0.700) | (0.866) | (0.472) | (0.610) | (0.758) | (0.493) |
| Population (logged) | -0.555 | -1.401* | -1.914* | 0.366 | 0.237 | 0.030 | -0.098 |
| | (0.669) | (0.848) | (1.140) | (0.676) | (0.658) | (0.715) | (0.663) |
| Regional Economic Disparity | | | | | | | |
| COV (NUTS2) | 0.078** | | | | | | |
| | (0.036) | | | | | | |
| WCOV (NUTS2) | | 0.139** | | | | | |
| | | (0.061) | | | | | |
| WCOV (NUTS3) | | | 0.154* | | | | |
| | | | (0.081) | | | | |
| INEQ SPNS (NUTS2) | | | | 0.233** | | | |
| | | | | (0.106) | | | |
| INEQ SPNS (NUTS3) | | | | | 0.491** | | |
| | | | | | (0.221) | | |
| AGC (NUTS2) | | | | | | 0.035 | |
| | | | | | | (0.063) | |
| AGC (NUTS3) | | | | | | | -0.025 |
| | | | | | | | (0.047) |
| Intercept | -5.616 | -10.516 | -13.810 | -5.544 | -14.601 | 5.866 | 1.114 |
| | (7.458) | (8.555) | (10.642) | (7.428) | (9.960) | (9.700) | (7.040) |
| -2log likelihood | -135.100 | -134.886 | -135.630 | -135.069 | -135.015 | -137.242 | -137.257 |
| AIC | 300.201 | 299.772 | 301.259 | 300.138 | 300.031 | 304.484 | 304.515 |
| N (parties) | 62 | 62 | 62 | 62 | 62 | 62 | 62 |
| N (countries) | 8 | 8 | 8 | 8 | 8 | 8 | 8 |

*Notes:*
Outcome variable: party support for decentralization, scaled from 1 to 20, with 20 the most positive stance.
Multilevel mixed-effects linear regression.
Standard errors are in parentheses and significant at ***p<0.01, **p<0.05, *p<0.1.
Data for 8 OECD countries in 2003
Country coverage: Belgium, Bulgaria, Finland, Italy, Romania, Slovakia, Spain, the United Kingdom.

of contention as regional inequality grows because rich regions increasingly bear the fiscal burden of the country. Furthermore, regional inequality increases heterogeneity in the national party system, putting competing pressures to represent local constituencies with different interests. In both theoretical accounts, dispersion of regional inequality is expected to positively correlate with preferences for decentralization.

Table 2 shows a consistently significant, positive relationship between dispersion of regional economic productivity (COV, WCOV, INEQ_SPNS) and decentralization, irrespective of political units considered at NUTS2 or NUTS3. The coefficient estimates vary somewhat, depending on the unit choice, but the results are not sensitive to the choice of measure or unit.

The measure of concentration of regional GDP per capita, AGC, on the other hand, is not significantly related to party positions on decentralization at the NUTS2 or NUTS3 level. We attribute this null finding to a theoretical difference between dispersion and concentration that may impact either fiscal structures or the party system. Concentration implies a small number of very productive regions that may prefer decentralization for either economic reasons or concerns with the party system. However, it also implies a lot of less productive regions with shared interest in centralization for fiscal reasons. Thus the ultimate effect on the party positions in the aggregate is not clear. Importantly, this result shows that the theoretical choice of dispersion or concentration can be very important, depending on the question at hand and the sample of countries.

In Table 3, we replicate Rickard's (2012) results linking electoral rules and the size of government spending, dependent on the geographic concentration or spread of economic interests. Rickard argues that whether proportional representation electoral systems or plurality electoral rules target more resources depends on whether economic interests are dispersed or concentrated. Specifically, governments with proportional representation rules are expected to distribute more resources when economic interests (here measured with manufacturing employment) are more broadly dispersed. However, when manufacturing employment is geographically concentrated, governments with plurality electoral rules are expected to distribute more in response to constituency interests.

Table 3 summarizes replication results from Rickard's (2012) with our reduced sample. Her dependent variable is subsidies to manufacturing (% of total spending). Throughout M1-M7, Rickard's basic results hold, showing that PR governments tend to spend more, but plurality electoral systems spend more when manufacturing interests are geographically concentrated. We again introduce our disperion and concentration measures at the two unit levels to this analysis. Across all measures, the relationship between regional inequality and manufacturing subsidies is positive, and nearly always significant. The results suggest that both dispersed and concentrated regional inequality are relevant to government spending of this kind, and in similar ways.[12] However, the WCOV dispersion measure does show some sensitivity to the unit choice. The NUTS2 unit is positive and significant, but the NUTS3 unit is positive and insignificant. This null finding may reflect a theoretical issue, that the NUTS2 level more closely captures the electoral unit of interest in these cases, or the finding may be spurious. In such cases, validating across different dispersion measures and different units increases the validity of the findings. Moreover, the results shown in M4 and M5 reveal the greater consistency of the INEQ_SPNS measure across the different unit measures.

Finally, we replicate Brancati's (2008) study showing political decentralization increases the strength of regional parties. Brancati argues that decentralization creates greater opportunities for regional parties in the region and the nation, thus increasing the likelihood that regional parties emerge as political decentralization increases.

Table 4 replicates Brancati's (2008) data analysis with added regional inequality measures. Brancati's basic result, that political decentralization increases regional party strength, is preserved across models. Controlling for regional inequality in these models substantially increases the model fit (shown in an R-squared increase from 0.350 in M1 to 0.619 in M2). The results have a smaller sample than Brancati's original models, however, so they cannot be directly compared.

The regional inequality measures each have significant correlations to regional party strength

[12]Our regional inequality concentration measure of GDP per capita may be somewhat collinear with Rickard's geographic concentration of manufacturing employment measures adopted by Rickard (2012). Rickard's blends data from NUTS2 and NUTS3 regions to produce her measures. This is theoretically reasonable when it matches the electoral unit of interest.

Table 3: OLS Regression Model Predicting the Effect of PR on Subsidy Budget Shares (Rickard, 2012)

| | M1<br>Rickard (2012) | M2<br>Extension | M3<br>Extension | M4<br>Extension | M5<br>Extension | M6<br>Extension | M7<br>Extension |
|---|---|---|---|---|---|---|---|
| L. PR | 2.126*** | 1.327*** | 1.931*** | 1.776*** | 1.769*** | 1.258*** | 1.339** |
| | (0.451) | (0.489) | (0.479) | (0.459) | (0.457) | (0.438) | (0.514) |
| L. Concentration (of Manufacturing Employment) | 29.277** | -34.831* | 13.045 | 21.280* | 19.504 | 3.538 | -0.600 |
| | (12.034) | (20.638) | (18.335) | (12.092) | (11.825) | (12.471) | (15.861) |
| L.PR * L. Concentration | -53.126*** | -0.195 | -38.032** | -44.893*** | -42.719*** | -24.605* | -25.203 |
| | (13.006) | (19.724) | (18.607) | (12.858) | (13.031) | (13.395) | (16.443) |
| L.Trade | 0.017*** | 0.011*** | 0.017*** | 0.010** | 0.017*** | 0.010** | 0.015*** |
| | (0.004) | (0.004) | (0.004) | (0.005) | (0.004) | (0.005) | (0.004) |
| L.GDP per capita (log) | -1.989*** | -2.445*** | -2.075*** | -2.155*** | -1.842*** | -1.797*** | -2.042*** |
| | (0.190) | (0.235) | (0.212) | (0.194) | (0.180) | (0.190) | (0.190) |
| L. Area (log) | 0.308*** | 0.241*** | 0.321*** | 0.142 | 0.437*** | 0.021 | 0.137 |
| | (0.078) | (0.081) | (0.077) | (0.097) | (0.072) | (0.105) | (0.107) |
| *Regional Economic Disparity* | | | | | | | |
| L. WCOV (NUTS2) | | 0.043*** | | | | | |
| | | (0.012) | | | | | |
| L. WCOV (NUTS3) | | | 0.011 | | | | |
| | | | (0.010) | | | | |
| L. INEQ_SPNS (NUTS2) | | | | 0.039*** | | | |
| | | | | (0.008) | | | |
| L. INEQ_SPNS (NUTS3) | | | | | 0.102*** | | |
| | | | | | (0.021) | | |
| L. AGC (NUTS2) | | | | | | 0.032*** | |
| | | | | | | (0.010) | |
| L. AGC (NUTS3) | | | | | | | 0.030** |
| | | | | | | | (0.014) |
| Intercept | 14.595*** | 20.576*** | 15.207*** | 18.492*** | 10.755*** | 16.281*** | 16.867*** |
| | (2.461) | (2.975) | (2.517) | (2.806) | (2.105) | (2.377) | (2.489) |
| N (Country Years) | 167 | 167 | 167 | 167 | 167 | 167 | 167 |
| $R^2$ | 0.533 | 0.580 | 0.539 | 0.585 | 0.607 | 0.567 | 0.547 |

*Notes:*
Outcome variable: the national government spending on subsidies to manufacturing industries (% of total government spending).
OLS regression with year fixed-effects.
Standard errors in parentheses and significant at ***$p<0.01$, **$p<0.05$, *$p<0.1$.
Data for 13 OECD countries from 1981 to 1997.
Country coverage: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Netherlands, Norway, Spain, Sweden, the United Kingdom.

Table 4: OLS Regression Model Predicting the Effect of Decentralization on Regional Party Strengths(Brancati, 2008)

| | M1 Brancati (2008) | M2 Extension | M3 Extension | M4 Extension | M5 Extension | M6 Extension | M7 Extension |
|---|---|---|---|---|---|---|---|
| Political Decentralization | 16.305*** | 22.195*** | 20.427*** | 12.997*** | 16.251*** | 21.891*** | 16.756*** |
| | (2.903) | (2.410) | (2.325) | (3.166) | (2.511) | (3.120) | (2.839) |
| Presidentialism | 3.293 | -0.629 | 4.295 | 0.488 | 2.181 | 19.275*** | 10.832** |
| | (3.596) | (2.822) | (2.767) | (3.689) | (3.120) | (5.721) | (5.188) |
| Total Number of Regions | 0.107 | 0.101 | 0.192*** | -0.020 | 0.135* | 0.184** | 0.103 |
| | (0.082) | (0.063) | (0.064) | (0.097) | (0.071) | (0.078) | (0.080) |
| First Elections | -2.658 | -0.629 | -6.879 | -1.935 | -6.883 | -3.495 | 1.543 |
| | (6.621) | (5.079) | (5.131) | (6.399) | (5.806) | (6.065) | (6.797) |
| Mixed Electoral Systems | -5.547 | -9.186* | 0.759 | -7.005 | -7.650 | -2.837 | -7.483 |
| | (6.154) | (4.745) | (4.833) | (5.975) | (5.343) | (5.687) | (6.078) |
| Majority / Plurality Systems | -3.124 | 7.744* | 5.887 | 1.169 | -6.414 | -7.095 | -6.405 |
| | (5.451) | (4.514) | (4.428) | (5.598) | (4.773) | (5.122) | (5.568) |
| *Regional Economic Disparity* | | | | | | | |
| WCOV (NUTS2) | | -0.622*** | | | | | |
| | | (0.099) | | | | | |
| WCOV (NUTS3) | | | -0.553*** | | | | |
| | | | (0.088) | | | | |
| INEQ_SPNS (NUTS2) | | | | -0.455** | | | |
| | | | | (0.203) | | | |
| INEQ_SPNS (NUTS3) | | | | | -1.218*** | | |
| | | | | | (0.276) | | |
| AGC (NUTS2) | | | | | | 0.668*** | |
| | | | | | | (0.196) | |
| AGC (NUTS3) | | | | | | | 0.465* |
| | | | | | | | (0.236) |
| Intercept | -3.662 | 8.762*** | 8.455** | 6.462 | 12.241** | -35.123*** | -27.050** |
| | (3.411) | (3.269) | (3.727) | (6.327) | (4.876) | (10.751) | (11.716) |
| N (Country Years) | 64 | 64 | 64 | 64 | 64 | 64 | 64 |
| Adjusted R2 | 0.350 | 0.619 | 0.616 | 0.394 | 0.514 | 0.455 | 0.382 |

*Notes:*
Outcome variable: the percentage of votes received by regional parties in an election.
OLS regression with decade fixed-effects.
Standard errors in parentheses and significant at ***p<0.01, **p<0.05, *p<0.1.
Data for 12 OECD countries from 1980 to 2002.
Country coverage: Belgium, Finland, Germany, Greece, Ireland, Norway, Poland, Romania, Slovenia, Spain, Sweden, the United Kingdom.

across all models in Table 4. The results are not sensitive to the unit, but are quite different across the dispersion and concentration measures. Regional party strength is negatively correlated with regional economic dispersion, but positively correlated with regional economic concentration. This distinction appears to be theoretical, reflective of different constellations of regional interests depending on whether regions have divergent or shared interests. These different inequality measures thus capture distinct economic distributions within nations that impact the formation of national policy coalitions.

Overall, the replications demonstrate the importance of regional inequality to related studies. They also show that theoretical consideration of these measures, both in the nature of distribution, and the political unit from which economic data are drawn, can impact the results and our interpretations of the importance of regional inequality. The replications also reveal the consistency of our INEQ_SPNS measure, showing it may be an optimal choice where the unit level is ambigious or data from the most appropriate unit are unavailable.

## 7 Conclusions and Research Implications

Existing research in economics and geography has made considerable advances in measuring regional economic variation within and across countries. These indicators include calculations of dispersion and concentration of economic productivity. However, these literatures have not been centrally concerned with the unit question, and the implications for data comparability and availability across units. We offer advice on common research challenges in measuring regional inequality, and delve deeply into the unit question.

We also show that our measure, INEQ_SPNS, is more reliable across political units than other measures. If the appropriate data are unavailable and available data are used instead, as is common practice in cross-national studies, researchers should consider using INEQ_SPNS for more stable estimates. Furthermore, we provide direction on where to find data on regional economies, and link to our datasets, which include both national aggregates and regional level data.

We have also demonstrated important theoretical and empirical differences across the two most common regional inequality concepts, dispersion and concentration. We show that these distributions do not always show the same relationship to important political phenomena. The difference across these configurations of regional inequality may be an important area for future research.

## References


Alt, J., Lassen, D. D., & Wehner, J. (2014). It isn't just about greece: domestic politics, transparency and fiscal gimmickry in europe. *British Journal of Political Science*, *44*(04), 707–716.

Barro, R. J., & Sala-i Martin, X. (1992). Convergence. *Journal of political Economy*, 223–251.

Beramendi, P. (2007). Inequality and the territorial fragmentation of solidarity. *International Organization*, *61*(4), 783.

Beramendi, P. (2012). *The political geography of inequality: regions and redistribution*. Cambridge University Press.

Bishop, B. (2009). *The big sort: Why the clustering of like-minded america is tearing us apart*. Houghton Mifflin Harcourt.

Bochsler, D. (2010). Measuring party nationalisation: A new gini-based indicator that corrects for the number of units. *Electoral Studies*, *29*(1), 155–168.

Bolton, P., & Roland, G. (1997). The breakup of nations: a political economy analysis. *The Quarterly Journal of Economics*, 1057–1090.

Brancati, D. (2008). The origins and strengths of regional parties. *British Journal of Political Science*, *38*(01), 135–159.

Buhaug, H., & Gates, S. (2002). The geography of civil war. *Journal of Peace Research*, *39*(4), 417–433.

Buhaug, H., Gleditsch, K. S., Holtermann, H., Østby, G., & Tollefsen, A. F. (2011). It's the local economy, stupid! geographic wealth dispersion and conflict outbreak location. *Journal of Conflict Resolution*, *55*(5), 814–840.

Buhaug, H., & Lujala, P. (2005). Accounting for scale: Measuring geography in quantitative studies of civil war. *Political Geography*, *24*(4), 399–418.

Charron, N., & Lapuente, V. (2013). Why do some regions in europe have a higher quality of government? *The Journal of Politics*, *75*(03), 567–582.

Chen, J., & Rodden, J. (2013). Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science*, *8*(3), 239–269.

Chhibber, P., & Kollman, K. (2009). *The formation of national party systems: Federalism and party competition in canada, great britain, india, and the united states*. Princeton University Press.

Crisp, B. F., Olivella, S., & Potter, J. D. (2013). Party-system nationalization and the scope of public policy the importance of cross-district constituency similarity. *Comparative Political Studies*, *46*(4), 431–456.

Dalton, H. (1920). The measurement of the inequality of incomes. *The Economic Journal*, *30*(119), 348–361.

Ezcurra, R., & Rodríguez-Pose, A. (2009). Measuring the regional divide. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories* (p. 329-353). Edward Elgar Publishing.

Fortunato, D., Swift, C. S., & Williams, L. K. (2016). All economics is local: Spatial aggregations of economic information. *Political Science Research and Methods*, 1–21.

Franzese, R. J., & Hays, J. C. (2007). Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis*, *15*(2), 140–164.

Gennaioli, N., La Porta, R., De Silanes, F. L., & Shleifer, A. (2014). Growth in regions. *Journal of Economic growth*, *19*(3), 259–309.

Harbers, I. (2015). Taxation and the unequal reach of the state: Mapping state capacity in ecuador. *Governance*, *28*(3), 373–391.

Hodler, R., & Raschky, P. (2014). Regional favoritism. *Quarterly Journal Of Economics [P]*, *129*(2), 995–1033.

Hooghe, L., & Marks, G. (2001). *Multi-level governance and european integration*. Rowman & Littlefield.

Jurado, I., & Leon, S. (2016). Geography matters: The conditional effect of electoral systems on social spending. *British Journal of Political Science*.

Jusko, K. L. (2015). Electoral geography and redistributive politics. *Journal of Theoretical Politics*, *27*(2), 269–287.

Kerner, A., Jerven, M., & Beatty, A. (2015). Does it pay to be poor? testing for systematically underreported gni estimates. *The Review of International Organizations*, 1–38.

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, *102*(5), 958–968.

Lee, D., & Rogers, M. (2016). *Regional inequality and the dynamics of government spending* (Tech. Rep.).

Lessmann, C. (2009). Fiscal decentralization and regional disparity: evidence from cross-section and panel data. *Environment and Planning A*, *41*(10), 2455–2473.

Lessmann, C. (2012). Regional inequality and decentralization: an empirical analysis. *Environment and Planning A*, *44*(6), 1363–1388.

Lessmann, C. (2015). Regional inequality and internal conflict. *German Economic Review*.

Mahler, V. A. (2002). Exploring the subnational dimension of income inequality: An analysis of the relationship between inequality and electoral turnout in the developed countries. *International Studies Quarterly*, *46*(1), 117–142.

Marx, A., & Rogers, M. (2016). Analysis of panamanian dmsp/ols nightlights corroborates suspicions of inaccurate fiscal data. *International Journal of Remote Sensing*.

Menendez, I. (2016). Globalization and welfare spending: How geography and electoral institutions condition compensation. *International Studies Quarterly*, sqw028.

Østby, G., Nordås, R., & Rød, J. K. (2009). Regional inequalities and civil conflict in sub-saharan africa1. *International Studies Quarterly*, *53*(2), 301–324.

Pigou, A. C. (1912). *Wealth and welfare*. Macmillan and Company, limited.

Raleigh, C., & Hegre, H. (2009). Population size, concentration, and civil war. a geographically disaggregated analysis. *Political geography*, *28*(4), 224–238.

Rehfeld, A. (2005). *The concept of constituency: Political representation, democratic legitimacy,* *and institutional design*. Cambridge University Press.

Rickard, S. J. (2012). Electoral systems, voters' interests and geographic dispersion. *British Journal of Political Science*, *42*(4), 855–77.

Rodden, J. (2010). The geographic distribution of political preferences. *Annual Review of Political Science*, *13*, 321–340.

Rodríguez-Pose, A., & Ezcurra, R. (2009). Does decentralization matter for regional disparities? a cross-country analysis. *Journal of Economic Geography*, lbp049.

Rogers, M. Z. (2015). *The politics of place and the limits of redistribution*. Routledge.

Sala-i Martin, X. X. (1996). Regional cohesion: evidence and theories of regional growth and convergence. *European Economic Review*, *40*(6), 1325–1352.

Sorens, J. (2005). The cross-sectional determinants of secessionism in advanced democracies. *Comparative political studies*, *38*(3), 304–326.

Spiezia, V. (2002). Geographic concentration of production and unemployment in oecd countries. *Cities and Regions*.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The journal of political economy*, 416–424.

Toubeau, S., & Wagner, M. (2015). Explaining party positions on decentralization. *British journal of political science*, *45*(1), 97–119.

Wallace, J. L. (2016). Juking the stats? authoritarian information problems in china. *British Journal of Political Science*, *46*(01), 11–29.

Weingast, B. R., Shepsle, K. A., & Johnsen, C. (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. *The Journal of Political Economy*, 642–664.

Wong, D. (2009). *The modifiable areal unit problem*. SAGE Publications: London, UK.

# Measuring Regional Inequality for Political Research

## Online Appendix

Appendix 1. Data Availability for Regional Level GDP around the World

| Country | Regional Unit | # of Units | Time Coverage | Source |
|---|---|---|---|---|
| Albania | Region | 12 | 1990,2001,2009 | Gennaioli et al. 2014 |
| Argentina | Province, Capital | 24 | 1958-2006 | Porto 1990, National Accounts |
| Australia | State, Capital | 7 | 1953-2010 | National Accounts, Gennaioli et al. 2014 |
| Austria | State (NUTS2) | 9 | 1961-2010 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Bangladesh | Region | 21 | 1982-2005 | Gennaioli et al. 2014 |
| Belgium | Province (NUTS2), Capital | 11 | 1970-2014 | Cambridge Econometrics |
| Benin | Province | 6 | 1992, 1998, 2004 | Gennaioli et al. 2014 |
| Bolivia | Department | 9 | 1980-2010 | National Accounts |
| Bosnia and Herzegovina | Canton, Entity | 12 | 1963, 2010 | Gennaioli et al. 2014 |
| Brazil | State, Capital | 27 | 1970-2014 | National Accounts |
| Bulgaria | Oblast | 28 | 1990-2010 | Gennaioli et al. 2014 |
| Canada | Province, Territory | 13 | 1970-2010 | National Accounts |
| Chile | Region | 15 | 1960-2014 | National Accounts, Gennaioli et al. 2014 |
| China | Province, Autonomous Region, Municipality | 31 | 1980-2010 | National Accounts |
| Colombia | Department, Capital | 33 | 1990-2012 | National Accounts |
| Croatia | County, Capital | 21 | 1995-2010 | Gennaioli et al. 2014 |
| Czech Republic | Region (NUTS2) | 14 | 1990-2014 | Cambridge Econometrics |
| Denmark | Region (NUTS2) | 5 | 1970-2014 | Cambridge Econometrics |
| Ecuador | Province | 24 | 1996-2010 | National Accounts, Gennaioli et al. 2014 |
| Egypt, Arab Rep. | Governorates | 21 | 1992, 1998, 2007 | Gennaioli et al. 2014 |
| El Salvador | Department | 14 | 1996, 1999, 2010 | Gennaioli et al. 2014 |
| Estonia | County | 15 | 1996-2010 | Gennaioli et al. 2014 |
| Finland | NUTS2 | 5 | 1960-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| France | Region (NUTS2) | 22 | 1950-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Germany | Lander | 16 | 1980-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Greece | Peripheries (NUTS 2) | 13 | 1970-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Guatemala | Department | 22 | 1995, 2000, 2008 | Gennaioli et al. 2014 |
| Honduras | Department | 18 | 1995, 2000, 2003 | Gennaioli et al. 2014 |
| Hungary | NUTS 2 | 7 | 1990-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| India | State, Union Territory | 36 | 1980-2014 | National Accounts |
| Indonesia | Province, Special Division | 34 | 1971-2010 | National Accounts, Gennaioli et al. 2014 |

Notes: We have accumulated the largest available global sample of regional level GDP measures at the equivalent of the NUTS2 level. These measures are also available at melissazrogers.com. The table below includes descriptions of the region unit, the number of region units, the year coverage of the data, and the data source.

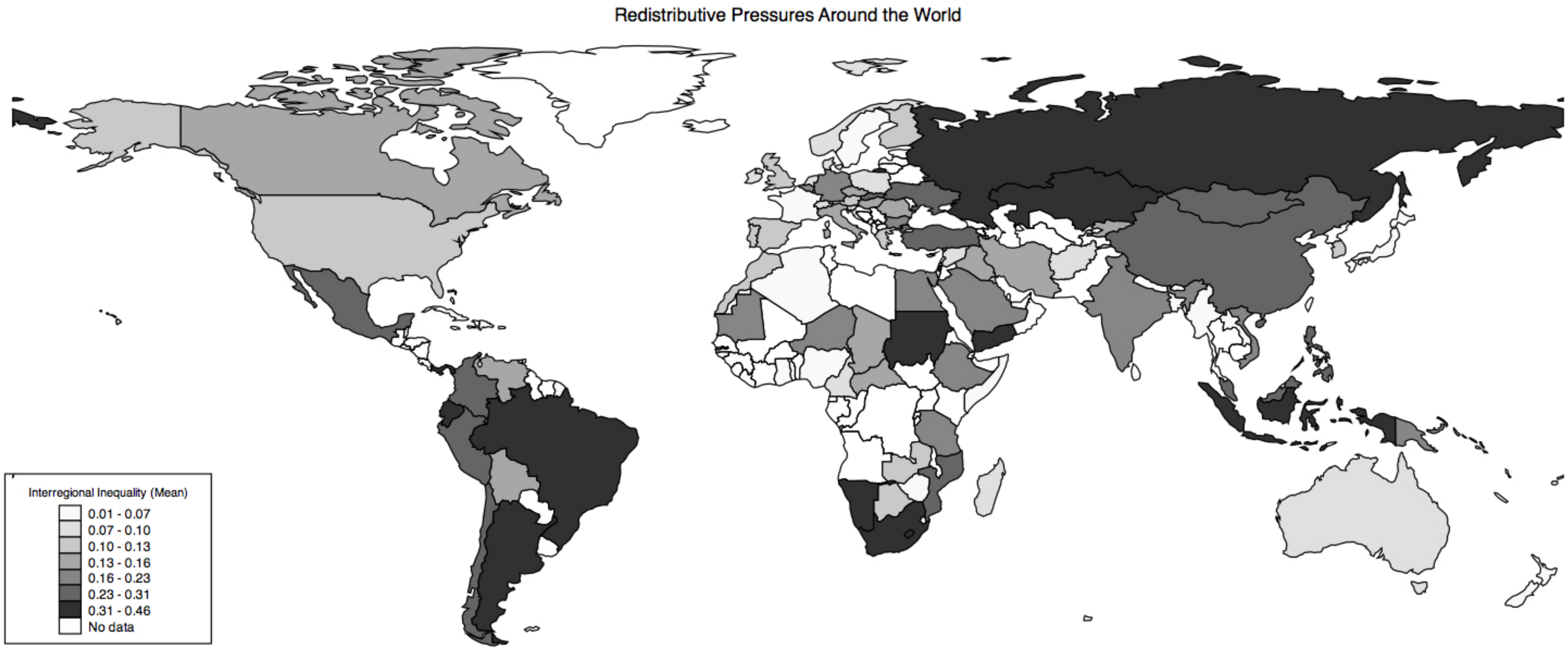
Appendix 1. Regional GDP Data Availability (Continued)

| Country | Regional Unit | # of Units | Time Coverage | Source |
|---|---|---|---|---|
| Iran, Islamic Republic | Province | 25 | 2000, 2005, 2010 | Gennaioli et al. 2014 |
| Ireland | NUTS 2 | 2 | 1980-2014 | Cambridge Econometrics |
| Italy | Region (NUTS 2) | 20 | 1950-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Japan | Region | 10 | 1955-2014 | National Accounts, Gennaiolli et al. 2014 |
| Jordan | Province | 12 | 1997, 2002, 2010 | Gennaioli et al. 2014 |
| Kazakhstan | Oblast, City | 16 | 1990-2010 | National Accounts |
| Kenya | Province | 5 | 1962, 2005 | Gennaioli et al. 2014 |
| Korea, Rep. | Province, Capital | 10 | 1985-2011 | National Accounts |
| Kyrgyz Republic | Oblast | 7 | 1996, 2000, 2005 | Gennaioli et al. 2014 |
| Latvia | District | 26 | 1995, 2000, 2005 | Gennaioli et al. 2014 |
| Lesotho | District | 6 | 1986, 1996, 2000 | Gennaioli et al. 2014 |
| Lithuania | County | 10 | 1995, 2000, 2005, 2010 | Gennaioli et al. 2014 |
| Macedonia, FYR | Statistical Region | 8 | 1963, 1990-2010 | Gennaioli et al. 2014 |
| Malaysia | State | 12 | 1970-2010 | Gennaioli et al. 2014 |
| Mexico | State, Capital | 32 | 1950-2014 | National Accounts, Gennaioli et al. 2014 |
| Mongolia | Province, Municipality | 20 | 1990-2010 | National Accounts, Gennaioli et al. 2014 |
| Morocco | Region | 7 | 1990-2010 | Gennaioli et al. 2014 |
| Mozambique | Province | 10 | 1996-2009 | Gennaioli et al. 2014 |
| Nepal | Development Region | 5 | 1999, 2006 | Gennaioli et al. 2014 |
| Netherlands | NUTS 2 | 11 | 1960-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| New Zealand | Island (NUTS 2) | 2 | 1980-2014 | National Accounts |
| Nicaragua | Region | 7 | 1974, 2000, 2005 | Gennaioli et al. 2014 |
| Nigeria | Region | 4 | 1992, 2008 | Gennaioli et al. 2014 |
| Norway | County (NUTS 2) | 7 | 1970-2014 | Cambridge Econometrics |
| Pakistan | Province | 4 | 1970-2004 | Gennaioli et al. 2014 |
| Panama | Province | 9 | 1996-2012 | National Accounts |
| Paraguay | Department, Capital | 18 | 1992, 2002, 2008 | Gennaioli et al. 2014 |
| Peru | Region | 24 | 1970-2012 | National Accounts, Gennaioli et al. 2014 |
| Philippines | Region | 17 | 1975-2014 | National Accounts, Gennaioli et al. 2014 |
| Poland | Voivodship (NUTS 2) | 16 | 1990-2014 | Cambridge Econometrics |
| Portugal | Region (NUTS 2) | 5 | 1977-2014 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Romania | Region (NUTS 2) | 8 | 1995-2010 | Cambridge Econometrics |
| Russian Federation | Subject | 83 | 1995-2010 | National Accounts |
| Serbia | District, City | 25 | 1963, 2002 | Gennaioli et al. 2014 |

Appendix 1. Regional GDP Data Availability (Continued)

| **Country** | **Regional Unit** | **# of Units** | **Time Coverage** | **Source** |
|---|---|---|---|---|
| Slovak Republic | Region (NUTS 2) | 4 | 1990-2014 | Cambridge Econometrics |
| Slovenia | Region (NUTS 2) | 2 | 1990-2014 | Cambridge Econometrics |
| South Africa | Province | 9 | 1995-2011 | National Accounts |
| Spain | Community (NUTS 2) | 19 | 1970-2014 | Cambridge Econometrics |
| Sri Lanka | Province | 9 | 1990-2010 | Gennaioli et al. 2014 |
| Sweden | NUTS 2 | 8 | 1980-2010 | Cambridge Econometrics |
| Switzerland | NUTS 2 | 7 | 1965-2010 | Cambridge Econometrics, Gennaiolli et al. 2014 |
| Tanzania | Region | 20 | 1980-2010 | National Accounts, Gennaioli et al. 2014 |
| Thailand | Province | 76 | 1982-2011 | National Accounts, Gennaioli et al. 2014 |
| Turkey | NUTS 2 | 26 | 1995-2014 | Cambridge Econometrics |
| Ukraine | Region, City | 26 | 1990-2013 | National Accounts |
| United Arab Emirates | Emirate | 7 | 1981-2009 | Gennaioli et al. 2014 |
| United Kingdom | County (NUTS 2) | 40 | 1980-2014 | Cambridge Econometrics |
| United States | States, Capital | 51 | 1950-2014 | National Accounts |
| Uruguay | Department | 19 | 1961, 1991, 1995, 2000 | Gennaioli et al. 2014 |
| Uzbekistan | Regions | 12 | 1995, 2000, 2005 | Gennaioli et al. 2014 |
| Venezuela | State | 23 | 1961, 1971, 1981, 1990 | Gennaioli et al. 2014 |
| Vietnam | Province | 39 | 1990-2008 | Gennaioli et al. 2014 |

Appendix 2. Global Map of Regional Inequality




Notes: Interregional inequality is a measure of country mean values based on RDGINI (region-adjusted gini coefficeint) at NUTS2 level. It ranges from 0 to 1, with 1 being the highest level of regional inequality.

Appendix 3. Summary Statistics by Country, NUTS level, and Regional Inequality Indicator

| Countries | No. of Political Units NUTS2 (NUTS3) | Values Circa 2011: NUTS2 (NUTS3) | | | | |
|---|---|---|---|---|---|---|
| | | COV | WCOV | RDGINI | INEQ_SPNS | AGC |
| Germany | 31 (326) | 0.21 (0.35) | 0.20 (0.39) | 0.11 (0.18) | 0.09 (0.09) | 0.34 (0.49) |
| United Kingdom | 41 (133) | 0.36 (0.43) | 0.42 (0.58) | 0.08 (0.17) | 0.12 (0.12) | 0.46 (0.55) |
| Italy | 21 (107) | 0.23 (0.23) | 0.24 (0.23) | 0.13 (0.14) | 0.12 (0.10) | 0.26 (0.33) |
| France | 26 (100) | 0.24 (0.33) | 0.28 (0.46) | 0.10 (0.13) | 0.12 (0.10) | 0.38 (0.46) |
| Poland | 16 (66) | 0.25 (0.41) | 0.29 (0.54) | 0.12 (0.20) | 0.14 (0.14) | 0.26 (0.41) |
| Spain | 19 (59) | 0.18 (0.19) | 0.20 (0.21) | 0.10 (0.11) | 0.12 (0.09) | 0.43 (0.49) |
| Greece | 13 (51) | 0.23 (0.29) | 0.27 (0.30) | 0.11 (0.14) | 0.19 (0.20) | 0.47 (0.53) |
| Romania | 8 (42) | 0.56 (0.41) | 0.53 (0.57) | 0.23 (0.21) | 0.25 (0.19) | 0.34 (0.33) |
| Belgium | 11 (44) | 0.33 (0.32) | 0.34 (0.36) | 0.17 (0.16) | 0.17 (0.14) | 0.34 (0.41) |
| Netherlands | 12 (40) | 0.17 (0.22) | 0.14 (0.23) | 0.10 (0.11) | 0.09 (0.09) | 0.32 (0.36) |
| Austria | 9 (35) | 0.17 (0.25) | 0.18 (0.25) | 0.10 (0.14) | 0.12 (0.12) | 0.30 (0.47) |
| Portugal | 7 (30) | 0.22 (0.27) | 0.25 (0.35) | 0.11 (0.14) | 0.20 (0.17) | 0.42 (0.57) |
| Bulgaria | 6 (28) | 0.43 (0.47) | 0.49 (0.68) | 0.21 (0.20) | 0.31 (0.27) | 0.36 (0.46) |
| Finland | 5 (20) | 0.27 (0.23) | 0.17 (0.26) | 0.14 (0.11) | 0.18 (0.14) | 0.50 (0.43) |
| Hungary | 7 (20) | 0.39 (0.43) | 0.46 (0.61) | 0.19 (0.19) | 0.30 (0.25) | 0.46 (0.44) |
| Sweden | 8 (21) | 0.20 (0.18) | 0.24 (0.57) | 0.08 (0.06) | 0.14 (0.12) | 0.57 (0.57) |
| Norway | 7 (19) | 0.23 (0.25) | 0.25 (0.32) | 0.12 (0.11) | 0.30 (0.13) | 0.44 (0.50) |
| Slovenia | 2 (12) | 0.19 (0.24) | 0.19 (0.27) | 0.09 (0.12) | 0.28 (0.16) | 0.28 (0.28) |
| Czech Republic | 8 (14) | 0.45 (0.36) | 0.44 (0.44) | 0.16 (0.13) | 0.18 (0.16) | 0.28 (0.30) |
| Denmark | 5 (11) | 0.16 (0.22) | 0.17 (0.21) | 0.08 (0.12) | 0.13 (0.12) | 0.34 (0.34) |
| Ireland | 2 (8) | 0.24 (0.32) | 0.19 (0.36) | 0.12 (0.17) | 0.34 (0.23) | 0.56 (0.40) |
| Slovakia | 4 (8) | 0.77 (0.57) | 0.55 (0.55) | 0.29 (0.24) | 0.36 (0.26) | 0.27 (0.29) |

Notes: The NUTS classification (Nomenclature of Territorial Units for Statistics) is defined on the basis of administrative divisions applied to the Eurostat member countries.