

Assignment 4:

Regional GDP Inequality in 4 Selected European Economies - Synthesised Results.

Kristoffer Tufta and Harald Vika

Monday 1 Dec, 2025

1 Assignment 4 Consolidation of Key findings

1.1 Abstract

1.2 Introduction

1.2.1 Background

1.2.2 Objectives

1.2.3 Sinificance

1.3 Literature Review - Kristoffer

1.3.1 Previous Work

A central reference point for our work is Lessmann & Seidel (2017), who shows that the link between economic development and regional inequality is not linear. Their results suggest something closer to an inverted-U (and in some cases N-shaped) pattern, where inequality tends to rise in the middle stages of development before falling again in high-income settings. This gives a natural expectation that fast-growing regions may show lower inequality if they are already relatively advanced, while poorer or developing regions might move in the opposite direction. Their work is also relevant because they rely on both traditional data and satellite-based luminosity data, which shows that inequality can be meaningfully captured even when official data is patchy — a recurring issue even in Eurostat datasets, as we experienced directly.

Eurostat's own methodological documentation is also part of the relevant background. Their notes on missing values help explain why regional GDP and population numbers are incomplete in certain years. These gaps arise from reclassification of regions, confidentiality rules, or delayed reporting. Understanding why these “quirks” happen is important because it directly affects our results.

For our segmentation work, we also build on Eurostat’s definitions of population-density classes used for identifying rural areas, urban clusters, and high-density urban centres. While Eurostat’s approach operates at the 1 km² continuous grid-cell level, their breakpoints (300 and 1500 people per square kilometer) give a practical structure for grouping NUTS2 regions by density. This matters because urban density is often correlated with both productivity and access to public-services, and therefore a likely influence when it comes to regional GDP inequality.

1.3.2 Research Gap

While there is existing literature on national GDP inequality trends and on broader development patterns, there is less focus on how short-term growth shocks play out inside individual NUTS2 regions, and even less on whether the effect of growth differs across structural categories like urban density, unemployment rates, or workforce size. Our assignments aim to examine exactly that: to see whether any of these structural categories influence GDP inequality in NUTS2-regions, and if so, to what degree.

1.4 Data and Methodology

1.4.1 Data Sources - Harald

1.4.2 Methodological approach - Kristoffer

1.4.2.1 Cross-Sectional Estimation

We first combined the Eurostat datasets covering GDP (in million euros) and population, then calculated GDP per capita and Gini coefficients using a formula weighted by population, like Lessmann & Seidel (2017). We then visualized the calculated Gini coefficients for the different regions using different plots from ggplot2. In assignment 2, we made a cross-sectional analysis of the year 2017 using a regression model to test the effect of change in regional GDP per capita on regional inequality. We specified the following linear regression model:

$$\text{Gini}_{i,2017} = \alpha + \beta \Delta \text{GDPpc}_{i,2016 \rightarrow 2017} + u_i$$

Where:

- $\text{Gini}_{i,2017}$ — Calculated Gini coefficient of region i in 2017 (regional GDP per capita inequality)
- $\Delta \text{GDPpc}_{i,2016 \rightarrow 2017}$ — percent change in GDP per capita from 2016 to 2017 in region i
- α — intercept term (baseline inequality when GDP-per-capita change is zero)

- β — slope coefficient showing how inequality changes with a one-percentage-point increase in GDP-per-capita growth
- u_i — error term capturing unobserved regional factors

We then experimented with other determinants of inequality by making a new OLS model with a set of new variables: population density (persons per square kilometer), unemployment rate and total workforce, also fetched from Eurostat. We specified the MLR as follows:

$$\text{Gini}_{i,2017} = \alpha + \beta_1 \text{Workforce}_i + \beta_2 \text{PopDensity}_i + \beta_3 \text{Unemployment}_i + u_i$$

where α is the intercept, β_1 , β_2 , and β_3 are the slope coefficients for each explanatory variable, and u_i is the error term capturing unobserved factors affecting inequality.

1.4.2.2 Segmentation

We also segmented regions into categories to examine if the development–inequality relationship differed across different structural environments. Population density was split into “Rural”, “Medium Density”, and “High Density” using Eurostat-inspired breakpoints (0–300, 300–1500, 1500+). For workforce size, we used quantile-based grouping (Low, Medium, High). Unemployment rate was also split into Low-High segments, but here we set the breakpoints manually at 0%-5%, 5%-10% and 10+%. These subsets were used to run separate regressions for each group.

1.4.2.3 Panel Estimation

Finally, we estimated fixed-effects models using the full panel, applying `plm()` with region, country, year, and two-way effects. Before estimation, we converted the dataset into a `pdata.frame` and cleaned remaining NA/Inf values to avoid dropped regions. The panel models allowed us to control for all time-invariant regional characteristics, giving a “less noisy” picture of how regional GDP inequality changes inside regions over time.

1.4.2.4 Handling of NA-values and Heteroskedasticity

After joining the datasets onto each other, including the new variables used for segmenting, we ended up with a few obvious gaps in our data and some NA/Inf rows. These were removed to keep the panel consistent, but a side effect of this is that we only have panel data from 2013 onwards.

We also ran Breusch–Pagan tests to check for heteroskedasticity across all regressions. Where alternative functional forms were required, we introduced quadratic and logarithmic transformations of the key variables and evaluated their residual behaviour the same way.

1.5 Empirical Findings - Harald

1.5.1 Cross- sectional Estimates

1.5.2 Alternative functional forms and panel estimates

1.6 Discussion - Begynner kvar for oss, samskriver etterpå

1.6.1 Key insights

1.6.2 Policy Implications

1.7 Limitations and Future Research - Harald

1.7.1 Reasearch limitations

1.8 Conclusion

1.8.1 Summary

1.8.2 Final Reflection

1.9 References

Appendix

Lessmann, C., & Seidel, A. (2017). Regional inequality, convergence, and its determinants – a view from outer space. *European Economic Review*, 92, 110–132. <https://doi.org/10.1016/j.eurocorev.2016.11.009>