

Convergence Clubs and Regression Trees

0686 - Spatial Economics

Nikolas, Philipp, Lukas & Daniel

17 January, 2019

Data Recap

European Regional Database by Cambridge Econometrics

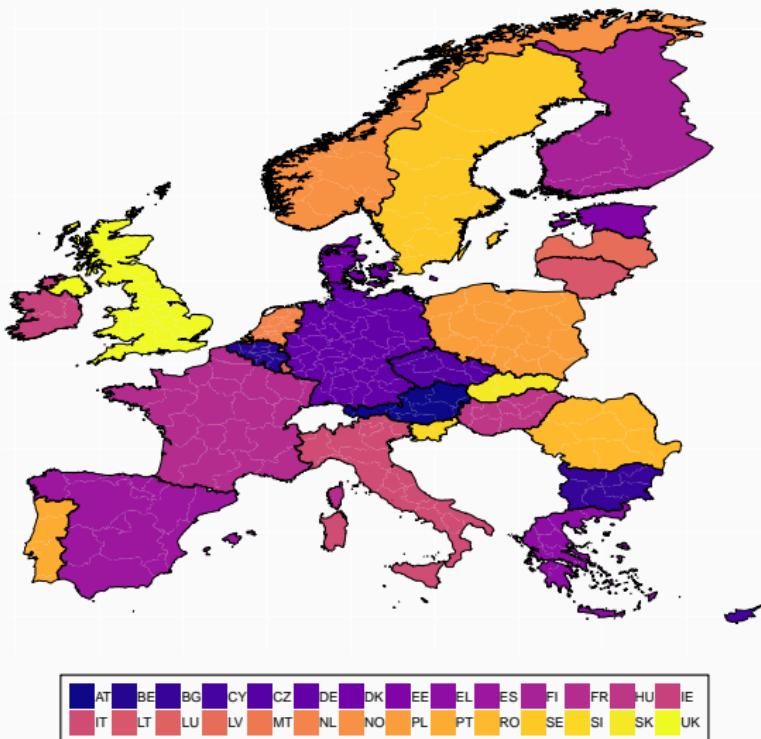
- Comprehensive dataset for NUTS2-regions over long period of time

We limit the dataset to:

- Focus on period after 2000 (possible breakpoint financial crisis 2008/9)
- Exclude Croatia (missing observations for employment indicators)

Results in a dataset of 273 regions for 28 countries

Data Recap



Oh what a merry regression tree

Split observations into clubs:

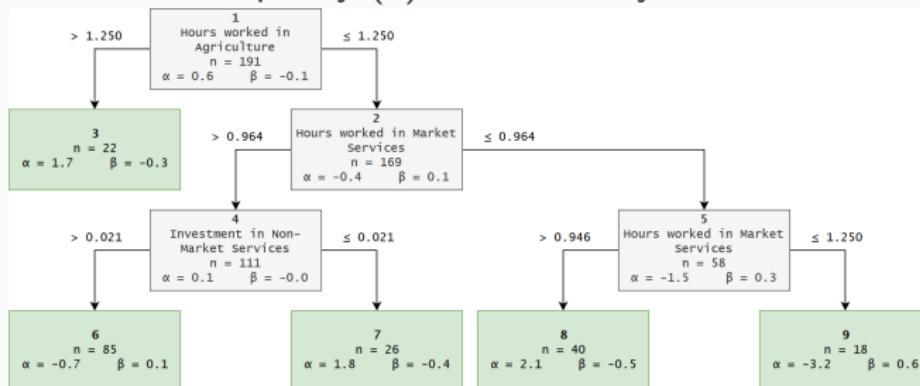
```
tree <- function(data, split_vars, end_criteria) {  
  split <- find_best_split(...)  
  if (!end_criteria) {  
    return(list(tree(split$data1, ...),  
               tree(split$data2, ...)))  
  } else { # if(end_criteria)  
    return(data)  
  }  
}
```

Regression Tree

We receive a recursive, tree-like data structure that is:

- hard to deal with (**a lot** of helper functions are necessary)
- nice

<!-- Add pretty (?) tree, dummy ahead: -->

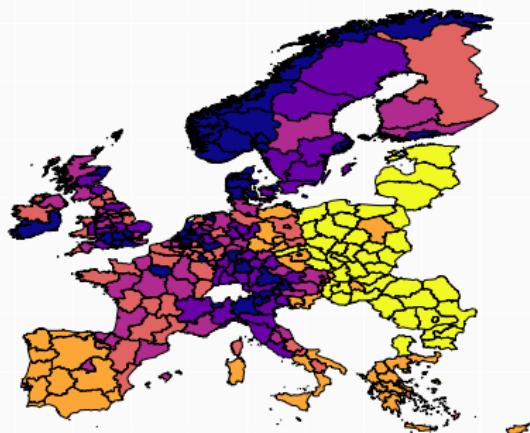


Regression Tree

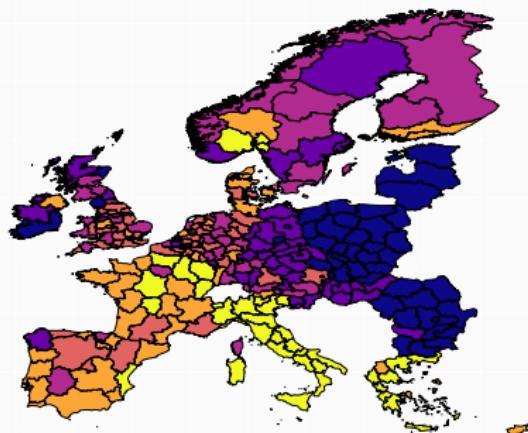
Our results are comparable to partykit (Hothorn and Zeileis 2015).

Still there's the caveat of spatially filtering the data. ## Motivation

GDP p.c. in 2000
Quantile map

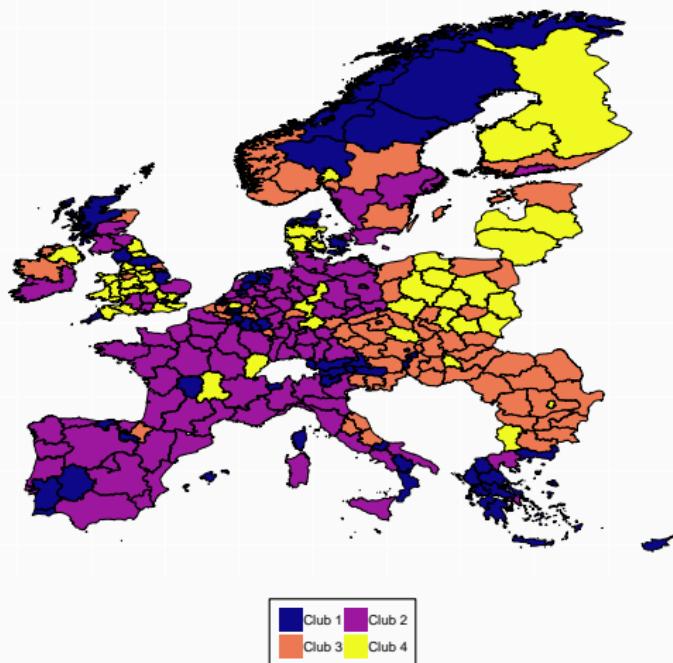


GDP p.c. growth 2000–15
Quantile map



Results

Convergence clubs NUTS 2
Unfiltered data



Results

Table 1: Regression results using unfiltered data

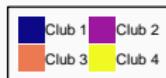
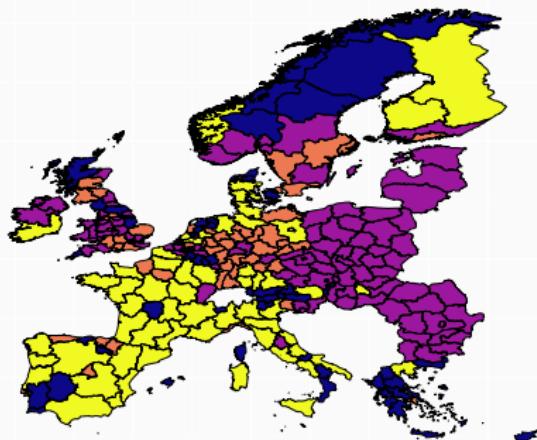
<i>Dependent variable:</i>				
	GDP p.c. growth rate 2000-15			
	(1)	(2)	(3)	(4)
Constant	-1.139*** (0.323)	-0.265 (0.360)	1.769*** (0.146)	2.922*** (0.147)
Initial GDP p.c.	0.120*** (0.032)	0.035 (0.036)	-0.159*** (0.016)	-0.275*** (0.015)
Observations	63	92	67	51
Residual Std. Error	0.118 (df = 61)	0.105 (df = 90)	0.129 (df = 65)	0.086 (df = 49)

Note:

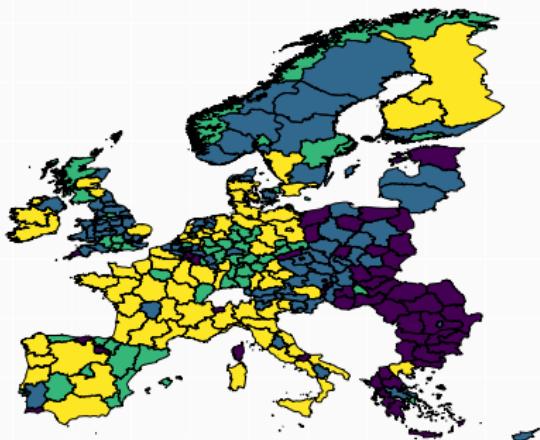
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Results

Convergence clubs NUTS 2
SAR-filtered data



Convergence clubs NUTS 2
SEM-filtered data



Results

Table 2: Regression results using SAR-filtered data

<i>Dependent variable:</i>				
	GDP p.c. growth rate 2000-15			
	(1)	(2)	(3)	(4)
Constant	-1.174*** (0.343)	1.445*** (0.122)	1.296*** (0.383)	-0.037 (0.470)
Initial GDP p.c.	0.109*** (0.034)	-0.142*** (0.013)	-0.128*** (0.037)	-0.003 (0.047)
Observations	63	97	55	58
Residual Std. Error	0.125 (df = 61)	0.124 (df = 95)	0.073 (df = 53)	0.110 (df = 56)

Note:

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Results

Table 3: Regression results using SEM-filtered data

<i>Dependent variable:</i>				
	GDP p.c. growth rate 2000-15			
	(1)	(2)	(3)	(4)
Constant	-0.039** (0.018)	0.088*** (0.014)	0.016 (0.020)	-0.021 (0.022)
Initial GDP p.c.	-0.277*** (0.022)	-0.265*** (0.026)	-0.061** (0.028)	-0.132*** (0.047)
Observations	55	89	59	70
Residual Std. Error	0.117 (df = 53)	0.120 (df = 87)	0.086 (df = 57)	0.106 (df = 68)

Note:

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

- club-plots
- some first LM vs. SAR vs. SEM comparisons

Literatur

Hothorn, Torsten, and Achim Zeileis. 2015. “partykit: A Modular Toolkit for Recursive Partytioning in R.” *Journal of Machine Learning Research* 16: 3905–9.
<http://jmlr.org/papers/v16/hothorn15a.html>.