

The European Commission's science and knowledge service

Joint Research Centre



THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Marcos Álvarez-Díaz, Beatrice D'Hombres, Claudia Ghisetti,

Nicola Pontarollo and Laura de Dominicis

COIN 2018 - Composite Indicators & Scoreboards Competence Centre

RSA Winter Conference

London, 16-11-2018

THE DETERMINANTS OF POPULATION CHANGE in EUROPE AT REGIONAL LEVEL

PURPOSE



STUDY OF THE MAIN DRIVERS OF THE POPULATION GROWTH IN THE EU-28 REGIONS (NUTS-3 LEVEL) FROM 2000-2015



FIRST STEP
Literature Review



SECOND STEP
Econometric Analysis



Identify the most important drivers

THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Structure

Literature Review

Empirical Study

Conclusions

1. Literature Review

- ❑ Countries are covered
- ❑ Unit of analysis commonly used
- ❑ Econometric methods employed
- ❑ **MAIN DETERMINANTS** of population growth

2. Empirical application at **NUTS3** level over the period **2000-2015**

- ❑ Model Specification
- ❑ Explanatory Variables
- ❑ Results of the spatial drift via penalized splines PSSD-SAR

3. Conclusions and Future Work

THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Structure

Literature Review

Empirical Study

Conclusions

**MAIN DETERMINANTS
OF POPULATION
GROWTH**

REGULARITIES

Existing evidence: mainly on US data. Evidence on Europe is limited

Most of the literature focuses on cities, counties or Metropolitan Areas

Spatial effects should be accounted for

**MULTIDISCIPLINARY
AREA OF RESEARCH**

**MAIN DETERMINANTS OF
POPULATION GROWTH**

THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Structure

Literature Review

Empirical Study

Conclusions

MAIN DETERMINANTS OF POPULATION GROWTH

DEMOGRAPHIC CHARACTERISTICS

- **Initial Population**
- **Previous Growth**
- Natural pop growth (fertility/death rates)
- Migration
- Age and population distribution
- Ratio women/men
- **Density**
- Blacks or Hispanics

SOCIO ECONOMIC CONDITIONS

SOCIO-ECONOMIC

- **Initial economic conditions**
- **Sectoral composition/ industrial specialization**
- Level of wages
- Housing availability
- real estate values
- Poverty rates
- Public sector investments and taxes
- Crime rate
- Cultural amenities

NATURAL AMENITIES

- **Climate conditions** (Precipitations, temperature, heating and cooling days, humidity, **frozen days**, etc.)
- Forest coverage
- Public land coverage
- Parks
- **Coastal proximity**

TRANSPORTATION ACCESSIBILITY

- **Accessibility**
- Highways, railways, airport
- **Proximity to central cities**
- Proximity to highways, airports etc.
- Public transportation system
- Distance variables
- Regional/province/state dummies

LAND USE AND DEVELOPMENT

- Land use
- Available land
- Topographical characteristics
- Tax exempted lands
- Built-up lands
- Wetland
- Slope

THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Structure

Literature Review

Empirical Study

Conclusions

MODEL SPECIFICATION: spatial drift via penalized splines PSSD-SAR

$$\mathbf{y} = f(s_1, s_2) + \rho \mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{B}\delta + \mathbf{W}\mathbf{Z}\vartheta + \mathbf{u}$$

Y : population growth of the European regions over the period 2000-2015

X : set of demographic and socio-economic variables

B : gathers geographical variables

$f(s_1, s_2)$: spatial interpolation surfaces, for which the nonparametric spatial drift

W·Z : spatially lagged explanatory variables

u : disturbance term

THE DETERMINANTS OF POPULATION CHANGE AT THE REGIONAL LEVEL

Structure

Literature Review

Empirical Study

Conclusions

MODEL SPECIFICATION: spatial drift via penalized splines PSSD-SAR

$$\mathbf{y} = f(s_1, s_2) + \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{B} \boldsymbol{\delta} + \mathbf{W} \mathbf{Z} \boldsymbol{\vartheta} + \mathbf{u}$$

The PSSD-SLM specifications (Montero et al., 2018) take both spatial autocorrelation and spatial heterogeneity into account by combining a nonparametric spatial drift with a standard SLM

The marginal effects are computed as:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta}$$

Where $(\mathbf{I} - \rho \mathbf{W})^{-1}$ is the so-called spatial multiplier, that takes into account the spatial spillover effects

ONGOING RESEARCH ON THE MAIN DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

DEMOGRAPHIC AND SOCIO-ECONOMIC VARIABLES

	Explanatory Variable	Definition	Expected Effect
X Matrix	GDP per capita	Log of GDP per capita at 2005 constant prices	+
	Industrial sector	Pct. of workers in industrial sector in 2000	+
	Agricultural sector	Pct. of workers in agricultural sector in 2000	-
	Population	Population in 2000	+
	Population density	Population over regional area in 2000	+
	Accessibility	Log of potential accessibility index	+
	Remoteness	Share of people living within NN minutes from cities with at least 250,000 people	-

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

GEOGRAPHICAL VARIABLES

	Explanatory Variable	Definition	Expected Effect
B matrix	Costal regions	Dummy 1 if the regions is costal, 0 otherwise	+
	Mountain regions	Dummy 1 if the regions is mountain, o otherwise	+
	Intermediate regions	Dummy 1 if people living in rural areas are beteewn 15% and 50%, 0 otherwise	+
	Urban regions	Dummy 1 if people living in rural areas are less than 15%, 0 otherwise	+
	Frozen days	Share of days below zero degree (five years average)	-
	Rain intensity	Average millimetres of rain (five years average)	-

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

SPATIAL WEIGHTING MATRIX W

The elements of the Weighting Matrix W is based on a k nearest neighbors, where we account for the 4 nearest neighbours.

This avoid us to have isolated islands and an asymmetric W .

The matrix W is row-standardized.

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

SPATIALLY LAGGED VARIABLES

	Explanatory Variable	Definition	Expected Effect
WX Matrix	W GDP per capita	Log of GDP per capita at 2005 constant prices in neighbour regions	+
	W Population density	Population over regional area in 2000 in neighbour regions	+
	W Accessibility	Log of potential accessibility index in neighbour regions	+

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

SOME NOTES ON THE DATA

- ❑ The data come from **2 Sources of Information: Cambridge Econometric's Regional Database and EUROSTAT.**
- ❑ The sample includes **1095 NUTS-3 regions** out of 1391.
 - All regions belonging to the same metro region were **merged** into a single NUTS3 region.
- ❑ The **explanatory variables were measured at the beginning of the sample period** to mitigate the possible problems of endogeneity or double directionality.

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

SPATIAL MODEL SELECTION

	2000-2015			2000-2007			2008-2015		
Total	statistic	p.value		statistic	p.value		statistic	p.value	
LMerr	623.02	< 2.2e-16 ***		LMerr	615.0457 < 2.2e-16 ***		524.523	< 2.2e-16 ***	
LMlag	643.668	< 2.2e-16 ***		LMlag	665.0619 < 2.2e-16 ***		505.098	< 2.2e-16 ***	
RLMerr	9.306	0.002284**		RLMerr	1.4884	0.2225	23.469	1.27E-06***	
RLMlag	29.953	4.43E-08***		RLMlag	51.5046	7.14E-13***	4.045	0.0443**	

- ❑ The model selection is based on standard spatial models
- ❑ LM tests for spatial drift via penalized splines not yet implemented
- ❑ Spatial drift via penalized splines is a refinement of linear spatial models

Thus, in a first approximation, we can rely on standard LM tests

Spatial lag is the chosen model

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Full Sample: 2000-2015

	2000-2015		
	Coeff	Std. err.	P value
logPOP	0.00074	0.00000	0.00002***
Share agric. Employment	0.00029	0.00000	0.87437
Share ind. Employment	-0.00287	0.00000	0.11754
Log(GDP per capita)	0.00416	0.00000	0.00000***
Logo(pop. Density)	-0.00062	0.00000	0.00000***
Accessibility	-0.00061	0.00000	0.04026**
Remoteness	-0.00227	0.00000	0.00000***
Urban region	0.00033	0.00000	0.77624
Intermediate region	0.00103	0.00000	0.00002***
Frozen days	0.00008	0.00000	0.50078
Rain days	-0.00001	0.00000	0.43089
Mountain region	-0.00057	0.00000	0.11279
Costal region	0.00079	0.00000	0.04427**
W log(GDP per capita)	-0.00155	0.00000	0.03000***
W accessibility	-0.00017	0.00000	0.41264
W log(pop. Density)	-0.00039	0.00000	0.37051
rho	0.43696	0.00360***	

GDP per capita and labor market conditions drivers of regional population growth

Intermediate regions have higher growth rate in comparison with the rural regions.

Population density has a negative relationship with population growth.

Population has a positive relationship with population growth.

Accessibility and being a neighbor of a metropolitan area have a negative effect on population growth

Costal regions have a sustained population growth.

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Full Sample: 2000-2015

	2000-2015		
	Coeff	Std. err.	P.value
logPOP	0.00074	0.00000	0.00002***
Share agric. Employment	0.00029	0.00000	0.87437
Share ind. Employment	-0.00287	0.00000	0.11754
Log(GDP per capita)	0.00416	0.00000	0.00000***
Logo(pop. Density)	-0.00062	0.00000	0.00000***
Accessibility	-0.00061	0.00000	0.04026**
Remoteness	-0.00227	0.00000	0.00000***
Urban region	0.00033	0.00000	0.77624
Intermediate region	0.00103	0.00000	0.00002***
Frozen days	0.00008	0.00000	0.50078
Rain days	-0.00001	0.00000	0.43089
Mountain region	-0.00057	0.00000	0.11279
Costal region	0.00079	0.00000	0.04427**
W log(GDP per capita)	-0.00155	0.00000	0.03000***
W accessibility	-0.00017	0.00000	0.41264
W log(pop. Density)	-0.00039	0.00000	0.37051
rho	0.43696	0.00360***	

Spatial Effects - W-Z - weakly statistically significant

Population growth in one region not only affected by its own characteristics, but also by those of the neighboring regions

Population growth of one specific region increases if its neighboring regions **display favorable economic conditions**

Spatial lag highly significant and equal to 0.44
It means that **78% of the effect on growth in a region passes through the neighbor regions**

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Full Sample: 2000-2015

		Direct	Indirect	Total
X matrix	logPOP	0.00078***	0.00054***	0.00132***
	Share agric. Employment	0.00031	0.00021	0.00052
	Share ind. Employment	-0.00301	-0.00208	-0.00510
	Log(GDP per capita)	0.00437***	0.00303***	0.00740***
	Logo(pop. Density)	-0.00065***	-0.00045***	-0.00110***
	Accessibility	-0.00064**	-0.00045**	-0.00109**
	Remoteness	-0.00238***	-0.00165***	-0.00403***
	B matrix	Urban region	0.00034	0.00024
Intermediate region		0.00108***	0.00075***	0.00183***
Frozen days		0.00008	0.00005	0.00013
Rain days		-0.00001	-0.00001	-0.00001
Mountain region		-0.00060**	-0.00042**	-0.00102**
Costal region		0.00083**	0.00057**	0.00140**

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Pre crisis: 2000-2007

	2000-2007		
	Coeff	Std. err.	P.value
logPOP	0.00082	0.00000	0.00013***
Share agric. Employment	0.00354	0.00001	0.11994
Share ind. Employment	0.00055	0.00001	0.80880
Log(GDP per capita)	0.00421	0.00000	0.00000***
Logo(pop. Density)	-0.00029	0.00000	0.02113***
Accessibility	-0.00087	0.00000	0.01962***
Remoteness	-0.00201	0.00000	0.00001***
Urban region	0.00077	0.00000	0.58602
Intermediate region	0.00125	0.00000	0.00003***
Frozen days	0.00016	0.00000	0.26647
Rain days	-0.00002	0.00000	0.18464
Mountain region	-0.00017	0.00000	0.70710
Costal region	0.00130	0.00000	0.00734***
W log(GDP per capita)	-0.00007	0.00000	0.93454
W accessibility	-0.00024	0.00000	0.34344
W log(pop. Density)	-0.00007	0.00000	0.90232
rho	0.35178	0.00441***	

GDP per capita and labor market conditions drivers of regional population growth

Intermediate regions have higher growth rate in comparison with the rural regions.

Population density has a negative relationship with population growth.

Population has a positive relationship with population growth.

Accessibility and being a neighbor of a metropolitan area have a negative effect on population growth

Costal regions have a sustained population growth.

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

	2000-2007			
	Coeff	Std. err.	P.value	
logPOP	0.00082	0.00000	0.00013***	
Share agric. Emp.	0.00354	0.00001	0.11994	
Share ind. Employment	0.00055	0.00001	0.80880	
Log(GDP per capita)	0.00421	0.00000	0.00000***	
Logo(pop. Density)	-0.00029	0.00000	0.02113***	
Accessibility	-0.00087	0.00000	0.01962***	
Remoteness	-0.00201	0.00000	0.00001***	
Urban region	0.00077	0.00000	0.58602	
Intermediate region	0.00125	0.00000	0.00003***	
Frozen days	0.00016	0.00000	0.26647	
Rain days	-0.00002	0.00000	0.18464	
Mountain region	-0.00017	0.00000	0.70710	
Costal region	0.00130	0.00000	0.00734***	
W log(GDP per capita)	-0.00007	0.00000	0.93454	
W accessibility	-0.00024	0.00000	0.34344	
W log(pop. Density)	-0.00007	0.00000	0.90232	
rho	0.35178	0.00441***		

Pre crisis: 2000-2007

Spatial Effects - **W-Z** - are not statistically significant

Anyway,
Spatial lag highly significant and equal to 0.35.
 It means that **54% of the effect on growth in a region passes through the neighbor regions**

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Pre crisis: 2000-2007

		Direct	Indirect	Total
W matrix	logPOP	0.00085***	0.00042***	0.00127***
	Share agric. Employment	0.00364*	0.00181*	0.00546*
	Share ind. Employment	0.00056	0.00028	0.00085
	Log(GDP per capita)	0.00408***	0.00203***	0.00611***
	Logo(pop. Density)	-0.00037***	-0.00018***	-0.00056***
	Accessibility	-0.00114***	-0.00057***	-0.00171***
B matrix	Remoteness	-0.00207***	-0.00103***	-0.00311***
	Urban region	0.00016**	0.00008**	0.00024**
	Intermediate region	-0.00002**	-0.00001**	-0.00002**
	Frozen days	0.00079	0.00040	0.00119
	Rain days	0.00128***	0.00064***	0.00192***
	Mountain region	-0.00017	-0.00009	-0.00026
	Costal region	0.00134***	0.00067***	0.00200***

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

2008-2015

Coeff Std. err. P.value

logPOP 0.00079 0.00000 0.00001***

Share agric. Employment -0.00309 0.00001 0.20501

Share ind. Employment -0.00654 0.00000 0.00110***

Log(GDP per capita) 0.00485 0.00000 0.00000***

Logo(pop. Density) -0.00015 0.00000 0.14875

Accessibility -0.00022 0.00000 0.47919

Remoteness -0.00242 0.00000 0.00000***

Urban region -0.00011 0.00000 0.92652

Intermediate region 0.00069 0.00000 0.00614***

Frozen days 0.00002 0.00000 0.83193

Rain days 0.00000 0.00000 0.80992

Mountain region -0.00097 0.00000 0.00956***

Costal region 0.00026 0.00000 0.52123

W log(GDP per capita) -0.00268 0.00000 0.00062***

W accessibility -0.00009 0.00000 0.65918

W log(pop. Density) -0.00052 0.00000 0.25173

rho 0.41259 0.00371***

Post crisis: 2008-2015

GDP per capita and labor market conditions drivers of regional population growth

Intermediate regions have higher growth rate in comparison with the rural regions.

Population density has a negative relationship with population growth.

Population has a positive relationship with population growth.

Accessibility and being a neighbor of a metropolitan area have a negative effect on population growth

Costal regions have a sustained population growth.

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Post crisis: 2008-2015

	2008-2015		
	Coeff	Std. err.	P.value
logPOP	0.00079	0.00000	0.00001 ***
Share agric. Employment	-0.00309	0.00001	0.20501
Share ind. Employment	-0.00654	0.00000	0.00110 ***
Log(GDP per capita)	0.00485	0.00000	0.00000 ***
Logo(pop. Density)	-0.00015	0.00000	0.14875
Accessibility	-0.00022	0.00000	0.47919
Remoteness	-0.00242	0.00000	0.00000 ***
Urban region	-0.00011	0.00000	0.92652
Intermediate region	0.00069	0.00000	0.00614 ***
Frozen days	0.00002	0.00000	0.83193
Rain days	0.00000	0.00000	0.80992
Mountain region	-0.00097	0.00000	0.00956 ***
Costal region	0.00026	0.00000	0.52123
W log(GDP per capita)	-0.00268	0.00000	0.00062 ***
W accessibility	-0.00009	0.00000	0.65918
W log(pop. Density)	-0.00052	0.00000	0.25173
rho	0.41259	0.00371 ***	

Spatial Effects - W-Z – are weakly statistically significant

Population growth of one specific region increases if its neighboring regions **display favorable economic conditions**

Anyway,

Spatial lag highly significant and equal to 0.35.

It means that **69% of the effect on growth in a region passes through the neighbor regions**

DETERMINANTS OF POPULATION GROWTH

Structure

Literature Review

Empirical Study

Conclusions

Post crisis: 2008-2015

		Direct	Indirect	Total
W matrix	logPOP	0.00083 ***	0.00052 ***	0.00135 ***
	Share agric. Employment	-0.00322	-0.00204	-0.00526
	Share ind. Employment	-0.00682 ***	-0.00431 ***	-0.01113 ***
	Log(GDP per capita)	0.00496 ***	0.00314 ***	0.00810 ***
	Logo(pop. Density)	-0.00070 ***	-0.00044 ***	-0.00115 ***
	Accessibility	-0.00033	-0.00021	-0.00053
	Remoteness	-0.00253 ***	-0.00160 ***	-0.00412 ***
B matrix	Urban region	0.00002	0.00001	0.00004
	Intermediate region	0.00000	0.00000	0.00000
	Frozen days	-0.00011	-0.00007	-0.00019
	Rain days	0.00072 ***	0.00045 ***	0.00117 ***
	Mountain region	-0.00101 ***	-0.00064 ***	-0.00165 ***
	Costal region	0.00027	0.00017	0.00044

CONCLUSIONS

Structure

Literature Review

Empirical Study

Conclusions

- **Economic variables tend to be the main determinants of population growth**
 - GDP per capita and total population have positive impact
 - Population density negative impact
 - Accessibility and remoteness have negative affects
 - **Geographical variables have important effect**
 - Mountain regions tend to have lower population growth
 - Costal regions have higher population growth
 - Natural factors are weakly significant
 - The degree of urbanization matters -> intermediate regions grow faster than rural ones
- Possibly related to overcrowded areas?

CONCLUSIONS

Structure

Literature Review

Empirical Study

Conclusions

- **The period before the crisis have different dynamics from the post-crisis period**
 - **Geographical variables have a stronger effects before the crisis**
 - **After the crisis economic variables matter more**
- **Spatial dependence is stronger before the crisis**
- **spatial drift via penalized splines PSSD-SAR specification, rather than requiring an a priori specification of the trend model form, lets the data suggest such a form. The smoothed spatial drift allows to account for spatial heterogeneity, and increasing the precision in the estimation of the models.**

CONCLUSIONS

Structure

Literature Review

Empirical Study

Conclusions

- **Possible steps ahead**

- **Considering separately EU13 and EU15**
- **Using the variability of the variables related to weather conditions**
- **Including variables related to the age**
- ...

...suggestions?



Thank you!

Any questions?

Welcome to email us at: jrc-coin@ec.europa.eu

COIN in the EU Science Hub

<https://ec.europa.eu/jrc/en/coin>

COIN tools are available at:

<https://composite-indicators.jrc.ec.europa.eu/>

