Università degli Studi Roma Tre
Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Magistrale in Matematica

Sintesi della Tesi di Laurea Magistrale in Matematica

# Taylor's law applied to the spatial distribution of human population density

Candidata:
Cecilia Bricci

Relatore:
Prof. Alessia Naccarato
Correlatore:
Dott. Federico Benassi
Dott. Andrea Pierini

Anno Accademico 2013-14
Sessione di Ottobre 2014

## Introduction

This paper tests a spatial Taylor's law (TL) using high-quality data on human population density, and proposes a simple model of exponential growth in population density that show an empirical linkage with TL.

In our work, we follow the approach of the article of Cohen J.E., Xu M. and Brunborg H., (2013) to apply Tayolor's law to human demography, with the ultimate aim of adding a new technique to demographic models. We test TL using data from the territorial aggregation of "Southern Europe", in the period $1950 - 2010$. The macro area of "Southern Europe" consists of 16 countries listed in this paper.

We suppose that the population density of each country had an exponential growth, following the exponential model proposed by Cohen J.E. (2013), and that each country had its own growth rate of the population density per capita, denoted by $r_i$ for the country $i$.

The values of the parameters $r_i$ have to be estimated from our data, and the intuitive method is to apply the linear regression model to each country separately. Otherwise we reasonably supposed that each country is tied to the other countries for different reasons, as migration flows, demographic causes or different social reasons. Therefore, if we suppose our system of equations could be correlated, the seemingly unrelated regressions (SUR) model was chosen to estimate the growth rates.

Results show improvement in the estimates of the growth rates, except that for two countries, Bosnia and Hertzegovina, and Croatia, that have a low coefficient of determination $R^2$. To begin drawing graphics of these two countries, we observed that have an evolution that recalled a third-order polynomial. Hence we estimated a new SUR model, imposing that growth rates of Bosnia and Herzegovina, and Croatia had to be estimated by the cubic regression model.

The model showed an increase of the coefficient of determination $R^2$ of these two countries, and the coefficients were all significant. We put the $r_i$ estimated by the cubic regression model for Bosnia and Herzegovina and Croatia in the weighted mean and in the weighted variance of the exponential model to build the relashionship of Taylor's law, but the result doesn't contribute to prove TL. So, we made our analysis with the old values of the growth rates, obtaining that Taylor's law describes remarkably well the spatial variation of population of the territorial aggregation of ”Southern Europe”, and that could offer a new empirical instrument in the ensemble of demographic techniques and models shared by demographers.

# 1 Theory, data and methods

## 1.1 Taylor's law

The ecologist Lionel Roy Taylor (1961) relates the logarithm of the variance of the density of a set of organisms in a specified area, to a linear function of the logarithm of the mean density: log(variance of population density) = $\log(a) + b \times \log$(mean population density), $\exists\, a, b > 0$ (Taylor *et al.*, 1978). In the early 1980s his colleagues J. N. Perry and I. P. Woiwod made a positive contribution to the development of that theory that became known as Taylor's law (Taylor *at al*, 1978), henceforth TL, or Taylor's power law of fluctuation scaling, and it has found applications in many different fields. Taylor's law was consequently applied to demography, and the first careful test was reported by Cohen J.E., Xu M. and Brunborg H., (2013) using human demographic data from Norway, in the period $1978 - 2010$.
Following the approach of that article of Cohen J.E. (2013), in this thesis we apply TL to the spatial distribution of the population density of the territorial aggregation of "Southern Europe", in the period $1950 - 2010$.

**Definition 1 (Taylor's law).** *We say that Taylor's law applies to population density D(t) for all times t, if and only if there exist real constants* $a > 0$ *and* $b$ *such that, for all t, Var(D(t)) = a* $(E(D(t)))^b$.
*Equivalently, Taylor's law (TL) applies to D(t) for all times t if and only if there exist constants* $a > 0$ *and* $b$ *such that:*

$$\log Var(D(t)) - b \log E(D(t)) = \log a. \tag{1.1}$$

We say that TL applies to *D(t)* holds in the limit as $t \to \infty$ as shown in the following definition.

**Definition 2.** *We say that TL applies to* D*(t) in the limit as* t *increases if there exist real constant* a $> 0$ *and* b *such that:*

$$\lim_{t\to\infty} \log Var(D(t)) - b \log(E(D(t))) = \log(a). \tag{1.2}$$

We focus on the parameter $b$ of Taylor's law, the slope of log variance as a linear function of log mean population density, because $b$ of TL is independent of the unit used to measure population density. For example, if TL holds when population density is measured by people per km$^2$, then it will also hold with the same value of $b$ when population density is measured by thousands of people per square mile. By contrast, the value of the parameter $a$ of TL depends on the units of measurement.

We assume that log$E$(D(t)) and log $Var$(D(t)) are differentiable functions of time $t$. We can take the derivative with respect to $t$ of both sides

$$\frac{d}{dt}\big[\lim_{t\to\infty}\log Var(D(t)) - b\log E(D(t))\big] = 0,$$

and we can switch the order of the derivative

$$\lim_{t\to\infty}[\frac{d\log Var(D(t))}{dt} - b\frac{d\log E(D(t))}{dt}] = 0,$$

or equivalently

$$\lim_{t\to\infty}[\frac{\frac{dVar(D(t))}{dt}}{Var(D(t))} - b\frac{\frac{dE(D(t))}{dt}}{E(D(t))}] = 0.$$

If $E$(D(t)) is not a constant function, which means that population density can grow or decrease at $t$ , then $\frac{dE(D(t))}{dt} \neq 0$ and we can divide the expression of $b$(t) on both sides

$$b(t) = \frac{\left(\frac{\frac{dVar(D(t))}{dt}}{Var(D(t))}\right)}{\left(\frac{\frac{dE(D(t))}{dt}}{E(D(t))}\right)}. \tag{1.3}$$

Let $b$(t) be the local slope, when it is definded, that is the slope at a finite time $t$ of log $Var$(D(t)) as a function of log $E$(D(t)).

If $\lim_{t\to\infty} b(t)$ exists and it is a finite constant, that constant is necessarily $b$ in TL

$$\lim_{t\to\infty} b(t) = b \tag{1.4}$$

and we say that TL holds in the limit of large time $t$.

If $\lim_{t\to\infty} b(t)$ does not exist or is not a finite constant, TL does not hold in the limit as $t \to \infty$.

## 1.2 Data

The data used in this paper are provided by the United Nations (UN), Department of Economic and Social Affairs of the United Nations Secretariat. In particular, we have taken two variables from the database of UN: ”population” and ”population density”, for the years $1950 - 2010$ (annual stock) for each single country of the territorial aggregtion Southern Europe. This territorial division is shown in the table below.

| **Southern Europe** |
|---|
| Albania |
| Andorra |
| Bosnia and Herzegovina |
| Croatia |
| Gibraltar |
| Greece |
| Holy See |
| Italy |
| Malta |
| Montenegro |
| Portugal |
| San Marino |
| Serbia |
| Slovenia |
| Spain |
| TFYR Macedonia |

## 1.3 Theory

We shall present the method of approaching data used in this thesis. We suppose that the population density of each country changes exponentially (increasing or decreasing) following the exponential model proposed by Cohen J.E. (2013).

### 1.3.1 Methods

Let $n$ be the number of countries of "Southern Europe". We label with $i$ the index for the countries ($i = 1, \ldots, n$) and with $t$ the index for time, $t = 1950, \ldots, 2010$. We denote by $D(t)$ the population density among countries (population size/area, number per km$^2$), and by $E$(D(t)) and $Var$(D(t)) the weighted mean and the weighted varianceof population density in year $t$. Under these assumptions, the weighted average over the countries of population density in year $t$ is

$$E(D(t)) = \sum_{i=1}^{n} \omega_i(t) D_i(t) = \frac{\sum_{i=1}^{n} n\omega_i(t) D_i(t)}{n}.$$

The weighted variance among countries is

$$\begin{aligned} Var(D(t)) &= \sum_{i=1}^{n} \omega_i(t)(D_i(t) - E(D(t)))^2 \\ &= \sum_{i=1}^{n} \omega_i(t) D_i(t)^2 - \left( \sum_{i=1}^{n} \omega_i(t) D_i(t) \right)^2. \end{aligned}$$

We assign to each country $i$ the weight $\omega_{i,t}$ at time $t$. We consider three weighting methods specified as follows

- equally weighted $\omega_{i,t} = \dfrac{1}{n}$.
- areally weighted $\omega_{i,t} = \dfrac{a_i}{a}$, where $a_i$ is the area of country $i$ and $a$ is the sum of the areas.
- population weighted $\omega_{i,t} = \dfrac{p_{i,t}}{p_t}$, where $p_{i,t}$ is the population of country $i$ in year $t$ and $p_t$ is is the sum of the populations of the $n$ countries considered.

### 1.3.2 An exponential model

We suppose that the population density of each country changes exponentially (increasing or decreasing), with an an exponential growth of population density at a constant rate in continuous time. Let $D_i(0)$ be the initial population density of country $i$, $D_i(0) > 0$, $\forall i$. We shall need the definition of the population growth rate of country $i$.

**Definition 3.** *The population density $D_i(t)$ of country $i$ at time $t$ satisfies:*

$$\frac{dE(t)}{dt} = r_i D_i(t), \tag{1.5}$$

$-\infty < t < +\infty$ *and* $-\infty < r_i < +\infty$, *i= 1, ..., n.*
*The coefficient* $r_i = (dD_i(t)/dt)/D_i(t)$ *is the rate of change of the population density per capita, or the population growth rate of country i.*

We assume that $r_i \neq r_j$ for $i \neq j$, and that some $r_i$ might be positive, some negative and at most one $r_i$ might be zero. This assumptions held true for the estimated values of $r_i$ of "Southern Europe".
We label the countries in decreasing order of population growth rate so that $+\infty > r_1 > r_2 > \ldots > r_n > -\infty$.

**Definition 4 (The exponential model).** *Let* $r_i \neq r_j$ *for* $i \neq j$, $D_i(0) > 0$ *at* $t = 0$ *for* i = *1, ...,* n *and* $0 < \omega_{i,t} < 1$, $\omega_{i,t} + \ldots + \omega_{i,t} = 1$. *Then the population density* $D_i(t)$ *of the exponential model for the country* i *at time* t *is*

$$D_i(t) = D_i(0)e^{r_i t}. \tag{1.6}$$

**Definition 5.** *Under the conditions of the exponential model, the mean among countries of the exponential model of population density is*

$$E(D(t)) = \sum_{i=1}^{n} \omega_i(t) D_i(t) = \sum_{i=1}^{n} \omega_i(t) D_i(0) e^{r_i t} \tag{1.7}$$

*at every time* t.

**Definition 6.** *Under the conditions of the exponential model, the weighted variance of population density of the exponential model at time* t *is*

$$Var(D(t)) = E(D(t)^2) - [E(D(t))]^2 = \sum_{i=1}^{n} \omega_i(t) D_i(0)^2 e^{2r_i t} - [E(D(t))]^2. \tag{1.8}$$

We shall show that the exponential model predicts a spatial TL in the limit of large time.

# 2 Results

We show the result of testing TL to our demographic data. In particular, we apply TL just to 11 countries of the macro area of "Southern Europe", excluding Andorra, Gibraltar, Holy See, Malta and San Marino from our analysis. These countries represent by their very nature, outliner compared to the other member countries of the territorial aggregation. They are all small territorial entities, sparsely populated and not always apt to have a natural demographic situation.

## 2.1 Exponential model of country population density

We estimated the growths rates (parameters $r_i$ of the exponential model) of the 11 countries considered of "Southern Europe" by applying the SUR method, and the results are shown in Table 1 . The growth rates could be estimated by using the linear regression for each country separately, but we reasonably supposed to deal with a system of correlated equations, where it is convenient to use the seemingly unrelated regressions (SUR) model to estimate the parameters.
The SUR method was applyed to $\log_{10}$ of country population density $D_i$(t) across all censuses as a linear function of time

$$\log_{10} D_i(t) = \log_{10} D_i(0) + (\log_{10} e) r_i t.$$

The estimated slopes of all countries were divided by $\log_{10} e$ to derive $r_i$ for each country $i$.

As shown in Table 1, we have a low coefficient of determination $R^2$ of Bosnia and Herzegovina, and Croatia that is respectively $R^2 = 0.4204$, and $R^2 = 0.6062$. To begin drawing graphics of the two countries, we observed that have an evolution that recalled a third-order polynomial. So, we tried to apply the cubic regression model to improve the values of the growth rates of these countries. Hence we estimated a new SUR model of 11 countries imposing that Bosnia and Herzegovina and Croatia respected the cubic regression model. The model revealed an increase of the coefficient of determination $R^2$, and the coefficients were all significant.

Table 1: *Statistics of parameters for exponential model with the SUR method of country populations in "Southern Europe", excluding Andorra, Gibraltar, Holy See, Malta and San Marino, from 1950 to 2010 on the logarithmic scale in base 10* : $\log_{10} D_i(t) = \log_{10} D_i(0) + (\log_{10} e) r_i t$.

| Code | Country | Growth rate of country $i$, $r_i$ ($\times 10^3$) | $\log_{10} D_i(0)$ | Multiple R-squared | Adjusted R-squared |
|---|---|---|---|---|---|
| 01 | Albania | 17,5300 | 1,6987 | 0,8652 | 0,8629 |
| 02 | Bosnia and Herzegovina | 4,7156 | 1,7988 | 0,43 | 0,4204 |
| 03 | Croatia | 2,6594 | 1,8491 | 0,6128 | 0,6062 |
| 04 | Greece | 6,7557 | 1,7679 | 0,9853 | 0,9853 |
| 05 | Italy | 3,8591 | 2,2055 | 0,8969 | 0,8951 |
| 06 | Montenegro | 6,7442 | 1,5118 | 0,8591 | 0,8567 |
| 07 | Portugal | 4,0939 | 1,9606 | 0,9363 | 0,9352 |
| 08 | Serbia | 6,8801 | 1,9096 | 0,9067 | 0,9051 |
| 09 | Slovenia | 5,9107 | 1,8707 | 0,9338 | 0,9327 |
| 10 | Spain | 7,6791 | 1,7523 | 0,9698 | 0,9693 |
| 11 | TFYR Macedonia | 8,3123 | 1,7303 | 0,9125 | 0,9110 |

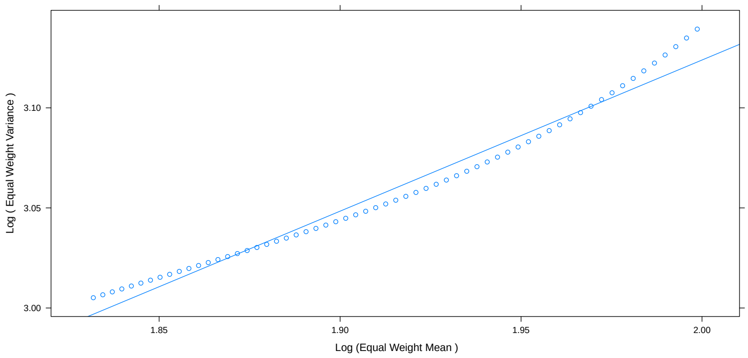
## 2.2 Taylor's law

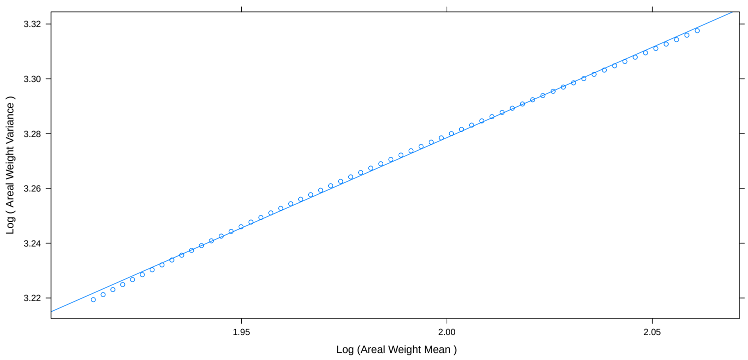
To see how well TL works, we show in Figure 1 the predicted mean and variance from the exponential model of population density, under the three weighting methods.
To obtain these graphics, we calculated for starting $E$(D(t)) and $Var$(D(t)) of the exponential model for $t = 0, \ldots, 60$, as follows

$$E(D(t)) = \sum_{i=1}^{n} w_{i,t} D_i(t) = \sum_{i=1}^{n} w_{i,t} D_i(0) e^{r_i t},$$
$$Var(D(t)) = E(D(t)^2) - [E(D(t))]^2 = \sum_{i=1}^{n} w_{i,t} (D_i(0))^2 e^{2r_i t} - [E(D(t))]^2, \tag{2.1}$$

using the $r_i$ of Table 1.
We didn't use the $r_i$ estimated by the cubic regression model for Bosnia and Herzegovina, and Croatia, because putting these values in 2.1 the result doesn't contribute to prove TL.
Then we obtained the sixty values of $E$(D(t)) and $Var$(D(t)), $t$= $0, \ldots, 60$, for each of the three sets of weights. Finally the linear regressions were fitted by the least square method to the independent variable log$E$(D(t)) and the dependent variable log $Var$(D(t)).
On visual inspection the linearity of the relationship is evident. The coefficients of determination on the other hands, show this linearity.
Statistics of these regressions are summarized in Table 2. The solid line is the least-squares regression line. Points from the most recent years appears in the upper right corner. 95 % CI of the slope and the intercept from the model fell within the 95 % CI of the corresponding parameters predicted from the data.

Figure 1: *Predicted mean and variance of population density of the 11 countries considered of the territorial aggregation of ""Southern Europe"", on* $log_{10}$*-*$log_{10}$ *coordinates, from an exponential model based on fitting a straight line to the log population density of each country separately, in the period 1950 - 2010, using respectively equal weights, areal weights and population weights.*



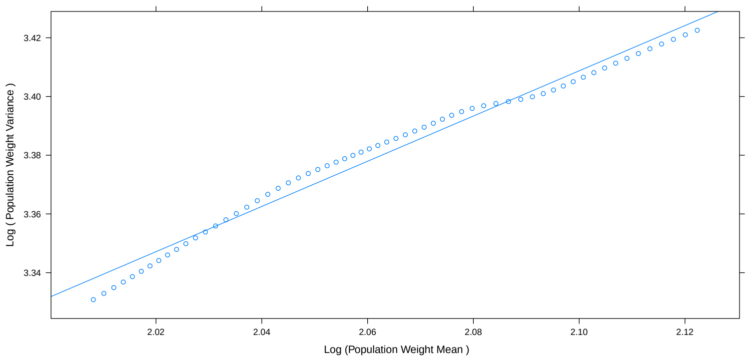



Table 2: *Statistics of linear regression of dependent variable* $\log_{10}$*(Variance of population density by country) on independent variable* $\log_{10}$*(Mean population density by country) from the data and the model, 1950 - 2010.*

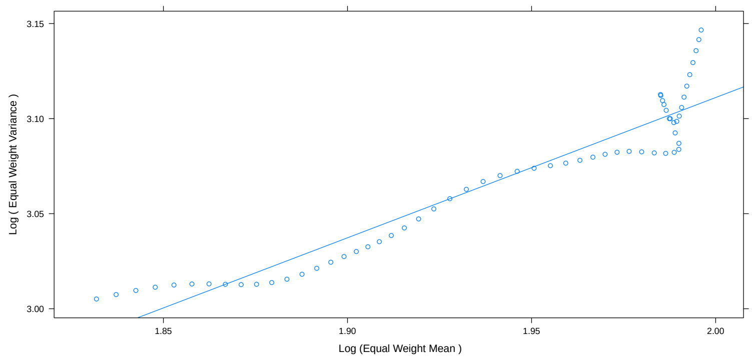
| Source | Weight | Slope | | | Intercept | | | Linear |
|---|---|---|---|---|---|---|---|---|
| | | Estimate | 95% CI | | Estimate | 95% CI | | $R^2$ |
| | | | Lower bound | Upper bound | | Lower bound | Upper bound | |
| **Data** | Equal | 0,73775 | 0,70669 | 0,76881 | 1,63560 | 1,26306 | 2,00814 | 0,9017 |
| | Area | 0,65426 | 0,62405 | 0,68446 | 1,98957 | 1,48185 | 2,49729 | 0,8049 |
| | Population | 0,80905 | 0,78085 | 0,83726 | 1,72972 | 0,91258 | 2,54686 | 0,7234 |
| **Model** | Equal | 0,75513 | 0,74137 | 0,76889 | 1,61363 | 1,43287 | 1,79439 | 0,9754 |
| | Area | 0,65944 | 0,65752 | 0,66137 | 1,95965 | 1,94589 | 1,99341 | 0,9989 |
| | Population | 0,77003 | 0,76448 | 0,77557 | 1,79170 | 1,61918 | 1,96421 | 0,9813 |

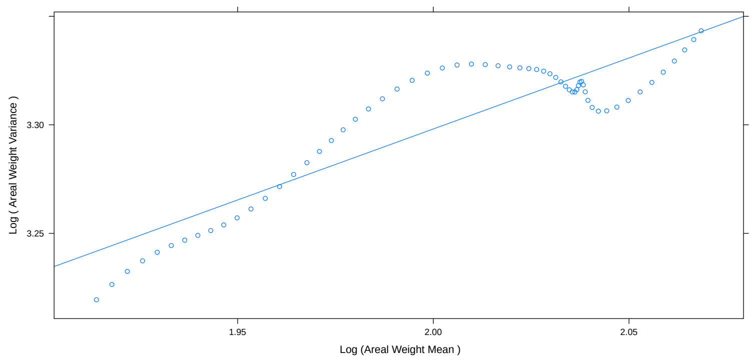
## 2.3 Model compared with Data

We show the observed mean and the observed variance of the population density, on a log scale, under each weighting method.

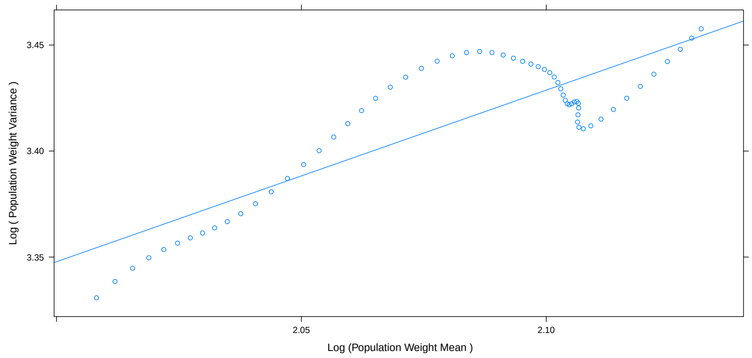
These graphics show the results in correspondence to 'Data' in Table

Figure 2: *Observed variance of population density of the 11 countries considered of the territorial aggregation of "Southern Europe" as a function of observed mean of population density, on $log_{10}$-$log_{10}$ coordinates, in the period 1950 - 2010, using respectively equal weights, areal weights and population weights.*




2, that is the observed mean and variance on a log scale. The solid line is the least-squares regression line. Using equal weights $R^2 = 0,9017$ from the data, compared with $R^2 = 0,9754$ from the model.Using areal weights $R^2 = 0,8049$ from the data, compared with $R^2 = 0,7234$ from the model. Using population weights $R^2 = 0,7234$ from the data, compared with $R^2 = 0,9813$ from the model. 95% CI of the parameters are given after the corresponding




point estimates.

# 3 Discussion and complements

The principal finding of this paper is that TL agreed remarkably well with the spatial variation of the population density of the 11 countries considered of the territorial aggregation of "Southern Europe". The plots of $\log Var(\mathrm{D(t)})$ as a linear function of $\log E(\mathrm{D(t)})$ are so close to straight lines that the coefficient of determination $R^2$ was significantly $> 0.95$ for each weighting method (see Table 2).
Can the exponential model predict parameters of Taylor's law? We shall show the theoretical behaviour of the slope $b(t)$ of TL in the limit of large time. Anyway, the results depend on the choose of weighting methods, as also shown by the theoretical results. In particular in this work, we report a comparative analysis of the consequences of different methods of weighting in computing the mean, the variance, and also the linear regressions, obtaining that the parameters of TL are sensitive to the weighting used.

## 3.1 Asymptotic analysis of the local slope $b(t)$ of Taylor's law under the three sets of weights

We show the theoretical results concerning the value of $b(t)$ of TL in the limit of large time. We focus on the parameter $b$ of Taylor's law, because $b$ is independent of the unit used to measure population density. By contrast, the value of the parameter $a$ of TL depends on the units of measurement.

The theoretical results will depend on the choose of weighting methods.

### 3.1.1 Limiting behaviour of $b(t)$ when countries are equally or areally weighted

In the simple case in which the weights of countries are time independent, the local slope $b$(t) of Taylor's Law approaches 2 as t $\to \infty$, adopting the exponential model.
The limiting exponent $b = 2$ of TL is independent of the population growth rate of the countries (provided the population growth rates are distinct) and independent of the proportions of countries (provided the proportions are positive).

**Theorem 3.1.** *Let all the weights $w_i(t) = w_i$ be constant in time.*
*If $r_i \neq r_j$ for $i \neq j$, and $w_i$ and $D_i(0)$ are all positive, $i$ = 1,...,n, then in the limit of large time, the spatial TL (1) holds with $b = 2$.*
*And this limit*

$$\lim_{t\to\infty} b(t) = 2 = b.$$

*is independent of the initial densities $D_i(0)$, the growth rates $r_i$ and the weights $w_i$, $\forall i$.*

### 3.1.2 Asymptotic analysis of $b(t)$ when countries are weighted by their population size or population density

A different result is obtained when the weights are function of time $t$. Countries are weighted by their population size or population density at time $t$ if, for positive constants $a_i$,

$$\omega_i(t) = \frac{D_i(t)a_i}{\sum_{i=1}^{n} D_i(t)a_i}, \tag{3.1}$$

where $a_i$ is interpreted as the land area of country $i$.
When $a_i$ are identical we say that countries are weighted by their population density.
We prove the empirical finding that the local slope $b$(t) of Taylor's Law, when the weights $w_i(t)$ depend on time – that is $w_i(t)$ is defined as in (3.1)– approaches $1 + r_2/r_1$ as t $\to \infty$.

This limit is independent of $D_i(0)$ and $a_i$ , $i = 1, \cdots, n$, when the $r_i$ are all distinct and the $a_i$ and $D_i(0)$ are all positive and constant in time.

**Theorem 3.2.** *Let $w_i(t)$ be the weights by population size or population density at time t, defined as*

$$\omega_i(t) = \frac{D_i(t)a_i}{\sum_{i=1}^{n} D_i(t)a_i}. \tag{3.2}$$

*Let $D_i(0)$ and $a_i$ be all positive and constant in time, $i$ = 1,...,n.*
*If $r_i \neq r_j$ for* i $\neq j$*, then in the limit of large time, the spatial TL (*1*) holds and*

$$\lim_{t\to\infty} b(t) = 1 + \frac{r_2}{r_1}. \tag{3.3}$$

*This limit is independent of the initial densities $D_i(0)$, the growth rates $r_i$ and $a_i$, $\forall i = 1, ..., n$.*

## 3.2 The growth-rate theorem (GR)

In addition, we allow us a little digression about the main theoretical result concerning the exponential model, and in particular the growth rates of the population, the so called growth-rate theorem (GR), proposed by Cohen J.E. (in press a, 2013).

**Theorem 3.3** (**The growth-rate theorem**)**.** *Let the growth rate of the population E(r(t)) at time t be the mean of $r_i$, i = 1,. . . ,n,*

$$E\big(r(t)\big) = \frac{\sum_{i=1}^{n} \omega_i(t) r_i D_i(t)}{\sum_{i=1}^{n} \omega_i(t) D_i(t)} = \frac{\sum_{i=1}^{n} \omega_i(t) r_i D_i(0) e^{r_i t}}{\sum_{i=1}^{n} \omega_i(t) D_i(0) e^{r_i t}} \tag{3.4}$$

*under the conditions of the exponential model.*
*The growth-rate theorem (GR) states that, in a subdivided population, the growth rate of the overall growth rate is proportional to the variance of the subpopulations' growth rates:*

$$\frac{dE\big(r(t)\big)}{dt} = Var\big(r(t)\big).$$

*Remark 1.* The mean population growth rate $E\big(\text{r(t)}\big)$ is also the denominator of the local slope $b$(t) of Taylor's law.

*Remark 2.* The upper limit of the mean population growth rate $E\big(\mathrm{r(t)}\big)$ is $r_1$. Simbolically, as t $\to \infty$,

$$E\big(r(t)\big) = \frac{d}{dt}\big(logE\big(r(t)\big)\big) \to r_1.$$

The intuitive reason for this conclusion is that the country 1 has the highest population growth rate, then as time increases country 1 increasingly dominates the population, so the per capita growth rate of the overall population density gets increasingly close to the per capita growth rate of country 1, which is $r_1$.
Moreover, we remark the importance of the interpretation of certain quantities in GR.

**Theorem 3.4.** *The overall growth rate of population density $E\big(r(t)\big)$ is positive for all t if all $r_i \geq 0$.*
*Let $\omega_i(t) = \omega_i$ be constant in time, all positive and $\sum_{i=1}^{n} \omega_i(t) = 1$. If $r_i \neq r_j$ for $i \neq j$ and $t \in (-\infty, +\infty)$, then Var(r(t)) $> 0$.*

**Theorem 3.5.** *Both Var$\big(r(t)\big)$, the variance of population growth, and $\frac{d}{dt}Var\big(r(t)\big)$, the growth rate of the variance of population growth rate, approach zero fo r large (positive or negative) time.*

Specifically, the mean population growth rate $E(\mathrm{r(t)})$ is the first derivative with respect to time of the logarithm of the mean population density, and the variance of population growth rate *Var*(r(t)) is the second time derivative of the mean population density.

## 4 The seemingly unrelated regressions (SUR) model

In our work, we deal with a system of equations that are different but that we suppose to be correlated. Countries are connected for different reasons, like migration flows, demographic causes or different social reasons, yielding to introduce the seemingly unrelated regressions (SUR) model to estimate the growth rates of countries. The name suggests that the error components are assumed to be correlated across the equations.
We assume that a total of $M$ observations are used in estimating the parameters of the $n$ equations. Each equation involves $k_i$ regressors, for a total of

$K = \sum_{i=1}^{n} k_i$.
The SUR model applies to all $n$ observations is

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \qquad i = 1, \dots, n. \tag{4.1}$$

where,

a) $\mathbf{y}_i$ is the $M \times 1$ vector of endogenous variables in the $i$-th equation,

b) $\mathbf{X}_i$ is the $M \times k_i$ vector of parameters in the $i$-th equation,

c) $\boldsymbol{\beta}_i$ is the $M \times 1$ vector of parameters in the $i$-th equation,

d) $\boldsymbol{\varepsilon}_i$ is the $M \times 1$ vector of error components in $i$-th equation.

More explicitly, in matrix form, the SUR model is

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & 0 & \dots & 0 \\ 0 & \mathbf{X}_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \mathbf{X}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_n \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_n \end{pmatrix} := \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4.2}$$

We shall generally use boldface type to indicate rows and column vectors, and bold capital letter to indicate matrices. Subscripts $i$ and $j$ will be used to denote columns (variables).

## 5 Generalized least squares

Let us assume

$$E(\boldsymbol{\varepsilon_i}) = 0, \qquad \forall i$$

and the error components are correlated across equations (but uncorrelated across observations), so that

$$E(\boldsymbol{\varepsilon_i}\boldsymbol{\varepsilon_i}^T) = \sigma_{ii}\mathbf{I}_M, \qquad \forall i$$
$$E(\boldsymbol{\varepsilon_i}\boldsymbol{\varepsilon_j}^T) = \sigma_{ij}\mathbf{I}_M, \qquad \forall i,j \qquad i \neq j.$$

Let $\mathbf{I}_M := \mathbf{I}$. For the $i$-th observation, the $n \times n$ variance-covariance matrix of the error components is $\mathbf{\Sigma}_{n,n} := \mathbf{\Sigma}$ defined as

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{pmatrix},$$

and the variance-covariance matrix of the error components of the entire system(4.2) is $\mathbf{\Omega}_{nM,nM} := \mathbf{\Omega}$, therefore,

$$\begin{aligned} \mathbf{\Omega} &= \mathbf{\Sigma} \otimes \mathbf{I} \\ &= E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) \end{aligned}$$

where we denote by $\otimes$ the Kronecker product.
Writing the inverse of $\mathbf{\Sigma}$ in the following way

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} \sigma^{11} & \sigma^{12} & \dots & \sigma^{1n} \\ \sigma^{21} & \sigma^{22} & \dots & \sigma^{2n} \\ \vdots & \vdots & & \vdots \\ \sigma^{n1} & \sigma^{n2} & \dots & \sigma^{nn} \end{pmatrix},$$

(i.e. denoting the $ij$-th element of $\mathbf{\Sigma}^{-1}$ by $\sigma^{ij}$), we have

$$\mathbf{\Omega}^{-1} = (\mathbf{\Sigma} \otimes \mathbf{I})^{-1} = \mathbf{\Sigma}^{-1} \otimes \mathbf{I}.$$

So that the parameters should be estimated with GLS (Generalized Least Squares) estimators. It is

$$\hat{\boldsymbol{\beta}}_{GLS} = [\mathbf{X}^T\mathbf{\Omega}^{-1}\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{\Omega}^{-1}\mathbf{y} = [\mathbf{X}^T(\mathbf{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{X}]^{-1}\mathbf{X}^T(\mathbf{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y}. \quad (5.1)$$

Or, the GLS estimator is explicitly given by

$$\hat{\boldsymbol{\beta}}_{GLS} = \begin{pmatrix} \sigma^{11}\mathbf{X}_1^T\mathbf{X}_1 & \sigma^{12}\mathbf{X}_1^T\mathbf{X}_2 & \dots & \sigma^{1n}\mathbf{X}_1^T\mathbf{X}_n \\ \sigma^{21}\mathbf{X}_2^T\mathbf{X}_1 & \sigma^{22}\mathbf{X}_2^T\mathbf{X}_2 & \dots & \sigma^{2n}\mathbf{X}_2^T\mathbf{X}_n \\ \vdots & \vdots & & \vdots \\ \sigma^{n1}\mathbf{X}_n^T\mathbf{X}_1 & \sigma^{n2}\mathbf{X}_n^T\mathbf{X}_2 & \dots & \sigma^{nn}\mathbf{X}_n^T\mathbf{X}_n \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^n \sigma^{1j}{\mathbf{X}_1}^T\mathbf{y}_j \\ \sum_{j=1}^n \sigma^{2j}{\mathbf{X}_2}^T\mathbf{y}_j \\ \vdots \\ \sum_{j=1}^n \sigma^{nj}{\mathbf{X}_n}^T\mathbf{y}_j \end{pmatrix}. \quad (5.2)$$

This estimator is obviously not coicident with ordinary least square (OLS) estimator.

### 5.0.1 Seemingly unrelated regressions with identical regressors

The special case of identical regressors – that is, $\mathbf{X}_i = \mathbf{X}_j = \mathbf{X}$ – goes under the name of multivariate regression model. In this special case, generalized least squares (GLS) estimator coincides with ordinary least squares (OLS) estimator.
Therefore, the OLS estimators of the parameters of the whole system coincide with OLS estimator equation by equation, which are best linear unbiased estimators (BLUE), consistent and asymptotically normally distributed (CAN).
The OLS estimator is

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{OLS} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \begin{pmatrix} (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T & 0 & \dots & 0 \\ 0 & (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} \\ &= \begin{pmatrix} (X^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}_1 \\ (X^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}_2 \\ \vdots \\ (X^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}_M \end{pmatrix}. \end{aligned} \tag{5.3}$$

If the equations have identical explanatory variables, it means that $\mathbf{X}_i = \mathbf{X}_j = \mathbf{X}$, where the matrix $\mathbf{X}$ is

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_{n,k} & 0 & \cdots & 0 \\ 0 & \mathbf{X}_{n,k} & \cdots & 0 \\ 0 & 0 & \cdots & \mathbf{X}_{n,k} \end{pmatrix} = \mathbf{I} \otimes \mathbf{X}.$$

Using the Kronecker products it can be easily to show that

$$\hat{\boldsymbol{\beta}}_{GLS} = \hat{\boldsymbol{\beta}}_{OLS}.$$

Infact

$$\hat{\boldsymbol{\beta}}_{GLS} = [\mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{X}]^{-1}\mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{y} = [\mathbf{X}^T\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{X}]^{-1}\mathbf{X}^T(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y}$$

$$= [(\mathbf{I} \otimes \mathbf{X}^T)(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{X})]^{-1}(\mathbf{I} \otimes \mathbf{X})^T(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y}.$$

Hence

$$[(\mathbf{I}\otimes\mathbf{X}^T)(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})(\mathbf{I}\otimes\mathbf{X})]^{-1} = [(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}^T)(\mathbf{I}\otimes\mathbf{X})]^{-1}] = [\boldsymbol{\Sigma}^{-1}\otimes\mathbf{X}^T\mathbf{X}]^{-1} = \boldsymbol{\Sigma}\otimes(\mathbf{X}^T\mathbf{X})^{-1},$$

which has to be muliplied through $(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{X}^T)\mathbf{y}$, to obtain

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{GLS} &= [\boldsymbol{\Sigma} \otimes (\mathbf{X}^T\mathbf{X})^{-1}][\boldsymbol{\Sigma}^{-1} \otimes \mathbf{X}^T]\mathbf{y} \\
&= [\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} \otimes (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T]\mathbf{y} \\
&= [\mathbf{I} \otimes (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T]\mathbf{y} \\
&= \begin{pmatrix} (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T & 0 & \dots & 0 \\ 0 & (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} = \hat{\boldsymbol{\beta}}_{OLS}.
\end{aligned}$$

# 6 The cubic regression model

Polynomial functions are usually used as regression function of the linear regression model, because of their capacity to model a large number of different forms of the unknown function, included straight lines or quadratic and cubic polynomials.

Le $k$ be the highest power of $X$, the polynomial regression model with one predictor is

$$y_i = \beta_0+\beta_1 X_i+\beta_2 X_i^2+\beta_3 {X_i}^3+\dots+\beta_k X_i^k+\varepsilon_i, \qquad i = 1, \dots, n. \tag{6.1}$$

or

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_i^j + \varepsilon_i. \tag{6.2}$$

If $k = 2$, the equation 6.1 is the quadratic regression model, that is a second order polynomial model

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i. \tag{6.3}$$

If $k = 3$, the equation 6.1 is the third-order polynomial model, or the **cubic regression model**, that is

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + \varepsilon_i, \tag{6.4}$$

so $X^3$ is the highest power of $X$ included in the model.
The cubic regression model is similar to the multiple regression model excepting that the regressors are powers of the same dependent variable $X$, i.e. $X$, $X^2$, $X^3$, whereas in the multiple regression regressors are different independent variables. Therefore the techniques of extimation developped for the multiple regression can be the same. In particular, the coefficients $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$ of the equation 6.1 can be extimate with OLS regression of $y_i$ over $X_i$, $X_i^2$ and $X_i^3$.
Referred to our work, we tried to aproximate the growth rates of Bosnia and Herzegovina, and Croatia with the cubic regression model. Infact, by applying the SUR method to the 11 countries of "Southern Europe", we had a low coefficient of determination $R^2$ of Bosnia and Herzegovina, and Croatia. To begin drawing graphics of the two countries, we observed that have an evolution that recalled a third-order polynomial. Hence we estimated a new SUR model of 11 countries imposing that Bosnia and Herzegovina and Croatia respected the cubic regression model.
The model reveals an increase of the coefficient of determination $R^2$ of these two countries, and the coefficients were all significant. The relative slopes were divided by $\log_{10} e$ to derive the new growth rates of the two countries, $r_{Bosnia}= 0,0285896$ and $r_{Croatia}= -0,0008834$. We put the $r_i$ estimated by the cubic regression model for Bosnia and Herzegovina and Croatia, in the weighted mean $E(\text{D(t)})$ and in the weighted variance $Var(\text{D(t)})$ of the exponential model to build the relashionship of Taylor's law, but the result doesn't contribute to prove TL. Infact, the coefficients of determination $R^2$ of the 'Model' are lower than $R^2$ of the simple linear regression, for the three sets of weights.

# 7 Conclusions

The principal finding of our work is that TL agreed reasonably well with the time course of spatial variation in the population density of the 11 countries considered of Southern Europe.
We showed that the exponential model can predict with useful accuracy parameters of Taylor's law in the limit of large time. The limiting exponent

$b = 2$ of TL, when the weights of countries are time independent, is independent of the population growth rate of the countries, and independent of the proportions of countries.
To evaluate TL, the choice of any spatial scale of analysis is in part arbitrary; we studied population density in Southern Europe at a scale of countries, but could be analyzed at different spatial scales (e.g. municipalities, counties, regions).
A practical application of TL would be to compare the parameters of TL estimated from historical data with the parameters of TL estimated from projections of future county populations. In such a case, TL might help evaluate the relative plausibility of alternative projections. If TL is successful, it could offer a new empirical instrument in human demography and could be added to the ensemble of demographic techniques and models shared by demographers.