RESEARCH IN ECONOMIC HISTORY

Edited by Shawn Kantor
and Carl T. Kitchens

RESEARCH IN ECONOMIC HISTORY

VOLUME 38

RESEARCH IN ECONOMIC HISTORY

This page intentionally left blank

RESEARCH IN ECONOMIC HISTORY VOLUME 38

RESEARCH IN ECONOMIC HISTORY

EDITED BY

SHAWN KANTOR
Florida State University, USA
National Bureau of Economic Research, USA

AND

CARL T. KITCHENS
Florida State University, USA
National Bureau of Economic Research, USA

emerald PUBLISHING

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Emerald Publishing, Floor 5, Northspring, 21-23 Wellington Street, Leeds LS1 4DL

First edition 2025

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-83608-929-2 (Print)
ISBN: 978-1-83608-928-5 (Online)
ISBN: 978-1-83608-930-8 (Epub)

ISSN: 0363-3268 (Series)

CONTENTS

This page intentionally left blank

LIST OF CONTRIBUTORS

Burton A. Abrams	University of Delaware, USA
Kirsten de Beurs	Wageningen University & Research, The Netherlands
James L. Butkiewicz	University of Delaware, USA
Gregori Galofré-Vilà	Universitat de València, Spain
María Gómez León	Universitat de València, Spain
Kyle Harper	University of Oklahoma, USA
Aaron M. Honsowetz	Bethany College, USA
Bárbara Tundidor	University Carlos III of Madrid, Spain
Le Wang	Virginia Tech, USA
Luis Felipe Zegarra	Pontificia Universidad Católica del Perú, Lima, Peru; CENTRUM Católica Graduate Business School, Lima, Peru

This page intentionally left blank

PROVINCIAL HEALTH INEQUALITIES IN SPAIN SINCE 1860

Gregori Galofré-Vilà and María Gómez León

Universitat de València, Spain

ABSTRACT

Using annual mortality rates at the provincial level for men and women, we construct a Gini index to estimate changes in regional health inequalities since 1860 in Spain. We find a long steady decline in health inequality across provinces from 1860 until today, interrupted by World War I and the Spanish Civil War. Over the 40 years of Franco's rule, health inequality stopped its downward trend and rose. Today, regional differences across provinces are at their lowest historical levels.

Keywords: Health; mortality; regional; Gini; inequality; Spain

JEL codes: J11; N23; N24; N93; N94

1. INTRODUCTION

Before the COVID-19 pandemic, Spain ranked among the healthiest places to live. Around 1880, the average life expectancy of a newborn boy in Spain was 30 years, compared to 33 years in Italy, 39 in the United States, and 43 in France. Yet, by 1980, Spain had caught up with Italy, France, and the United States, and in 2019, it only trails Japan globally (Riley, 2005). However, a close look at Spanish life expectancy in recent decades also shows heterogeneity across provinces with some places falling behind (Fig. 1). For instance, in 2019, the average life expectancy at birth in Alava, Guadalajara, Navarra, Salamanca, Segovia, Soria, and Valladolid was above 84.5 years and exceeding 85 years in the case of Madrid. By contrast, a newborn in Almería, Cádiz, Huelva, and Sevilla could expect to live less than 82 years, reaching only 81 in the case of Ceuta and Melilla. These patterns draw a north/south divide, with higher values in the center and northern areas and lower in the south and southwestern regions.

Research in Economic History, Volume 38, 1–25

ISSN: 0363-3268/doi:10.1108/S0363-326820250000038001

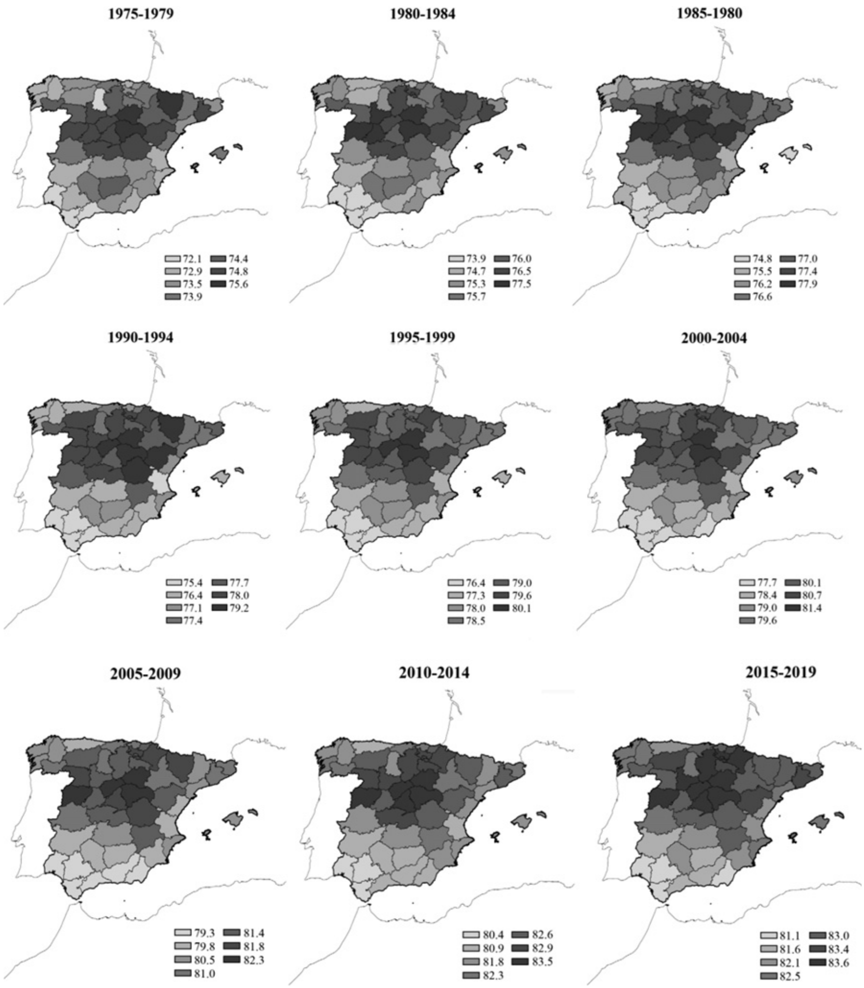

Fig. 1. Life Expectancy at Birth Across Spanish Provinces. *Notes:* Life expectancy at birth is defined as the average number of years that a newborn could expect to live if they were to pass through life subject to the age-specific mortality rates of a given period. As explained in Section 2, we are not showing data from Ceuta, Melilla, Santa Cruz de Tenerife, and Las Palmas. To display the life expectancy data across space, we use seven breaks in the data by the different Jenks. Jenk classes are based on natural groupings inherent in the data, whose boundaries are set where there are relatively big differences in the data values. *Sources:* Instituto Nacional de Estadística.

Given these gaps in recent decades, in this chapter, we document, for the first time, how health has varied across Spanish provinces in the last 150 years. Using official sources of death registers, we created a new data set with the number of yearly deaths per 1,000 population at the provincial level by sex. We focus on crude death rates, as historical age-adjusted population data to estimate life expectancies at birth are only available at census years (about every 10 years). While, in essence, our work is highly descriptive, we compute a Gini index and compare it with additional inequality indicators such as the Theil index and sigma- or beta-convergence.

By exploring these indices, we find that since 1860, there has been a process of health inequality reduction (i.e., a process of convergence) between Spanish provinces, which was interrupted in time during the Spanish Civil War and Franco's dictatorship (1936–1975) and resumed from 1980 onward. Using cause-specific mortality, we find that this last period of convergence is likely to be driven by advances in preventive mortality (i.e., deaths from cancer) and communicable and infectious diseases such as influenza and pneumonia. We also find divergence in recent decades in deaths from coronary and cardiovascular diseases. Finally, throughout, we also highlight some historical junctures and policies that can be behind periods of convergence or divergence.

While there is a wealth of research on health inequalities driven by socio-economic and educational groups (Case & Deaton, 2015; Currie & Schwandt, 2016), our paper is among the few exploring adult health inequalities historically. One notable exception is the recent work of Bonnet and D'Albis (2020), which explore convergence and divergence in life expectancy across French departments since the beginning of the 19th century. Between 1800 and 1880, they find zigzagged periods of convergence and divergence, followed by a century-long period of convergence. Another related paper is the work from Feigenbaum et al. (2019), showing that in the first half of the 20th century, not only were deaths from infectious diseases higher in the US South, but they also started to decline later than elsewhere in the country.

For the Spanish case, Pérez Moreda et al. (2015) also explored changes in mortality spatially, but their work is limited to urban regional differences during the first third of the 20th century. The work from Cussó Segura and Nicolau (2000), Gómez Redondo (1992), and Sanz Gimeno and Ramiro Fariñas (1999a, 1999b) is also very relevant here, but these studies aimed to explore infant and child mortality at some benchmark years. Our work directly connects with González and Rodríguez-González (2018), but while they begin to document regional health differences in 1990, we can go back to 1860.

There is also a broad and rich body of economic history literature looking at regional income inequality in Spain using gross domestic product (GDP) data (Díez Minguela et al., 2018; Martínez-Galarraga et al., 2018; Rosés et al., 2010; Rosés & Wolf, 2018), real wages (Rosés & Sánchez-Alonso, 2004) with some papers examining regional inequality and education (Beltrán Tapia et al., 2021; Beltrán Tapia & Martinez-Galarraga, 2018), and economic growth (Beltrán Tapia & Martinez-Galarraga, 2020; Martínez-Galarraga et al., 2015). Ultimately, our work directly contributes to this strand of scholarship, looking at the regional differences in health in the long run.

The paper continues as follows. We next present our data and methodology (Sections 2 and 3) to compute the mortality Gini index. Sections 4 and 5 describe the patterns of inequality in mortality over time and put them into an historical narrative. Section 6 concludes.

2. MORTALITY DATA: PATTERNS ACROSS PROVINCES AND ACROSS TIME

In this section, we will discuss data sources and limitations. We will also illustrate the heterogeneity across provinces and trends across time. Using a previously untapped data source to explore regional health inequalities, the *Movimiento Nacional de la Población*, edited by the *Instituto Geográfico y Estadístico*, we have manually transcribed the annual number of deaths for each of the 48 Spanish provinces since 1860 for men and women. For geographical reasons, we did not transcribe data from the provinces of Ceuta and Melilla (located in Africa) and Santa Cruz de Tenerife and Las Palmas, lying far away from the Spanish peninsula.

Because annual population figures at the provincial level are only available in census years (i.e., every 10 years), our outcome measure is simply the crude death rate (i.e., the number of deaths within a population). We computed it by dividing the annual deaths in the province i and year t by the population in the same province i and year t, where the annual population is interpolated between census years. To control for the process of ageing, after 1975, we use the official age-adjusted mortality statistics available from the *Instituto Nacional de Estadística*. Unfortunately, we have a gap in the data, as the *Movimiento Nacional de la Población* was not printed between 1871 and 1885, and mortality data by sex were also unreported between 1886 and 1899 (registering only total mortality). When we compute crude death rates for men and women, in the denominator, we use the male or female population in the province i and year t.

Fig. 2 displays the evolution of the crude death rates in each of the Spanish provinces from 1860 onward and compares it with the Spanish mean. Though each province follows a unique path, there are some general patterns. Mortality reached its peak in the 1870s and then started a long fall, reaching a plateau already in the 1960s until today. The demographic transition likely accounted for this decline (Omran, 1971). Starting from a regime of high birth and death rates and slow population growth, from 1880 onward, the death rate began to fall, accelerating the population increase. Then mortality and population were stabilized in the first half of the 20th century, when birth rates also fell, and the demographic transition was completed by 1960. It is also possible to see a short-lived peak of mortality in 1918, due to the 1918 influenza pandemic, and another increase of mortality during the second half of the 1930s, due to the Spanish Civil War.[1]

We also use the age-adjusted mortality data in 1975 and 2017 to show the distribution of the provincial mortality across space (Fig. 3). As with life expectancy, the mortality data also display a north south divide, with the provinces that registered higher mortality rates located in the South and Levante, including Almería, Cáceres, Granada, Málaga and València, and places with

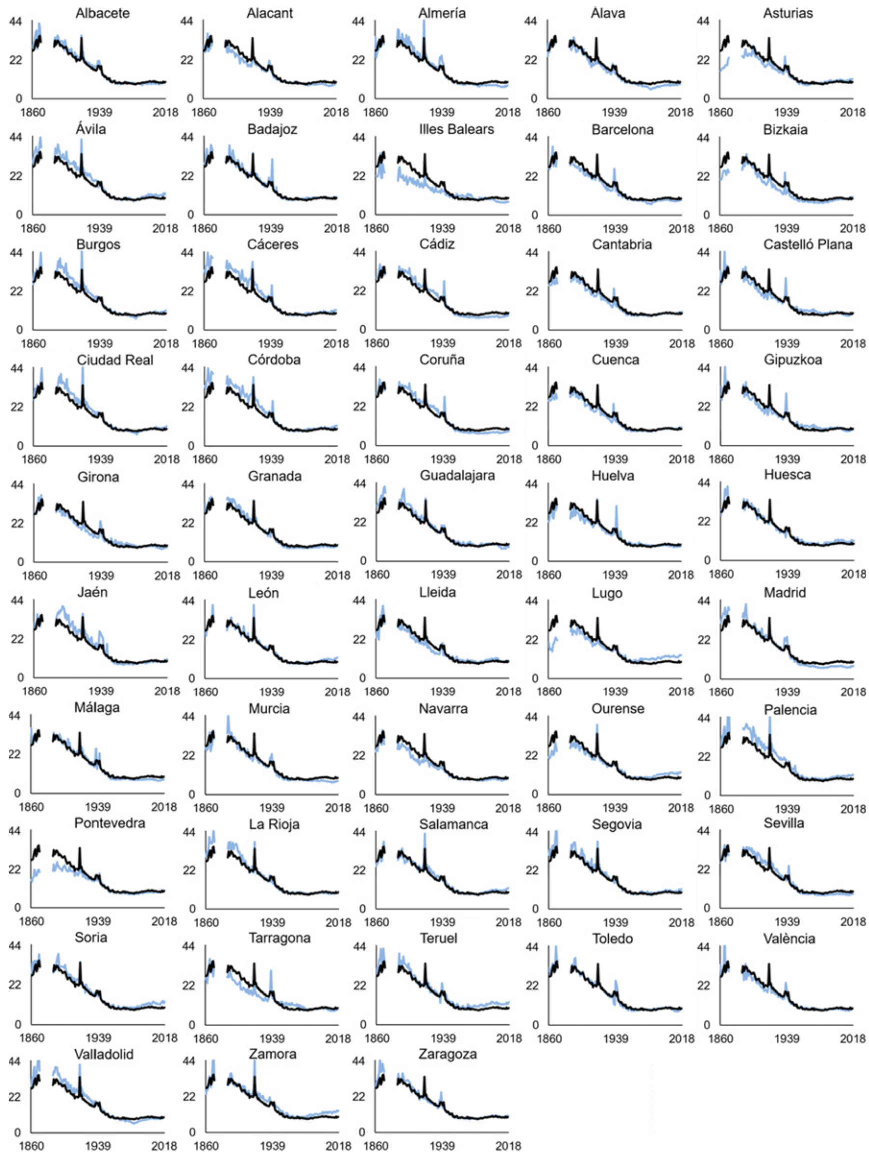

Fig. 2. Crude Death Rates in the Spanish Provinces. *Notes:* Light blue lines show annual average crude death rates in particular provinces. Dark black lines show annual average crude death rates in Spain. As explained in Section 2, we are not showing data from Ceuta, Melilla, Santa Cruz de Tenerife, and Las Palmas. *Sources:* Mortality data are from *Movimiento Nacional de la Población* and population data from census years; for more details, see Section 2.

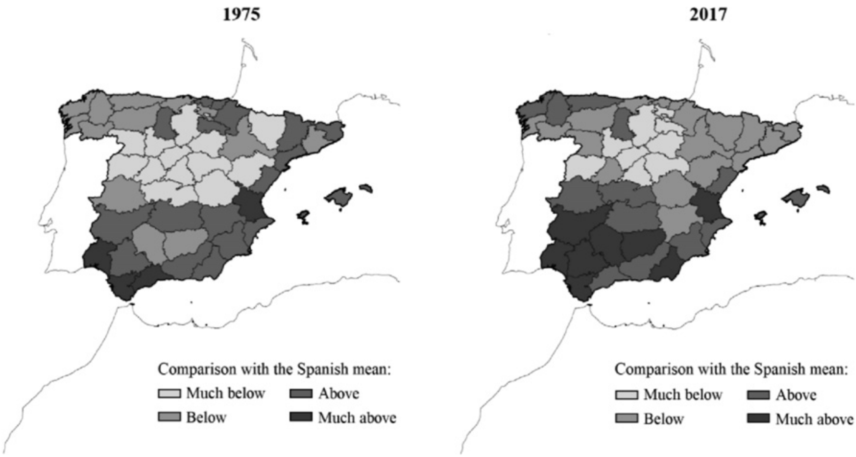

Fig. 3. Age-Adjusted Mortality Rates Across Spanish Provinces. *Notes:* As explained in Section 2, we are not showing data from Ceuta, Melilla, Santa Cruz de Tenerife, and Las Palmas. To display the age-adjusted mortality data across space, we use four breaks in the data by the different Jenks. Jenk classes are based on natural groupings inherent in the data, whose boundaries are set where there are relatively big differences in the data values. The categories of much below, below, above, and much above are calculated in relation to the Spanish average. *Sources:* Instituto Nacional de Estadística.

lower mortality rates clustered in the North, including Asturias, Bizkaia, Gipuzkoa, A Coruña, Lugo, and Pontevedra.[2] Using crude death rates, the north/south divide was already visible back in the 19th century.

Beyond differences in total mortality, we also computed sex mortality ratios (the ratio of male to female death rates per 100) across time in the different provinces (Fig. 4). Broadly speaking, women, due to biological factors, lifestyles, and risk behaviors, had nearly always achieved lower mortality rates than men (Goldin & Lleras-Muney, 2019). Nevertheless, in the last 50 years, there has been wide heterogeneity in the male to female mortality ratio across the different provinces, with some provinces registering low sex mortality ratios (such as Cáceres, Ciudad Real, and Jaén) when compared to others (Alava, Baleares, Burgos, and Gipuzkoa). Differences here are likely due to cause-specific mortalities (Pérez Moreda et al., 2015). Over time, the sex mortality ratios also show how during the Civil War, men who were sent to the front died prematurely. However, there is less evidence of gender discrimination in mortality during the 1918 influenza pandemic.

To control for unobservables in the crude death rates, such as the age distribution and the environmental and social characteristics of the provinces, we next examine the timing of the decline in mortality with a simple ordinary least squares (OLS) regression model, with the crude death rates as outcome and regressing the mortality on region and year dummies. Specifically, we organize

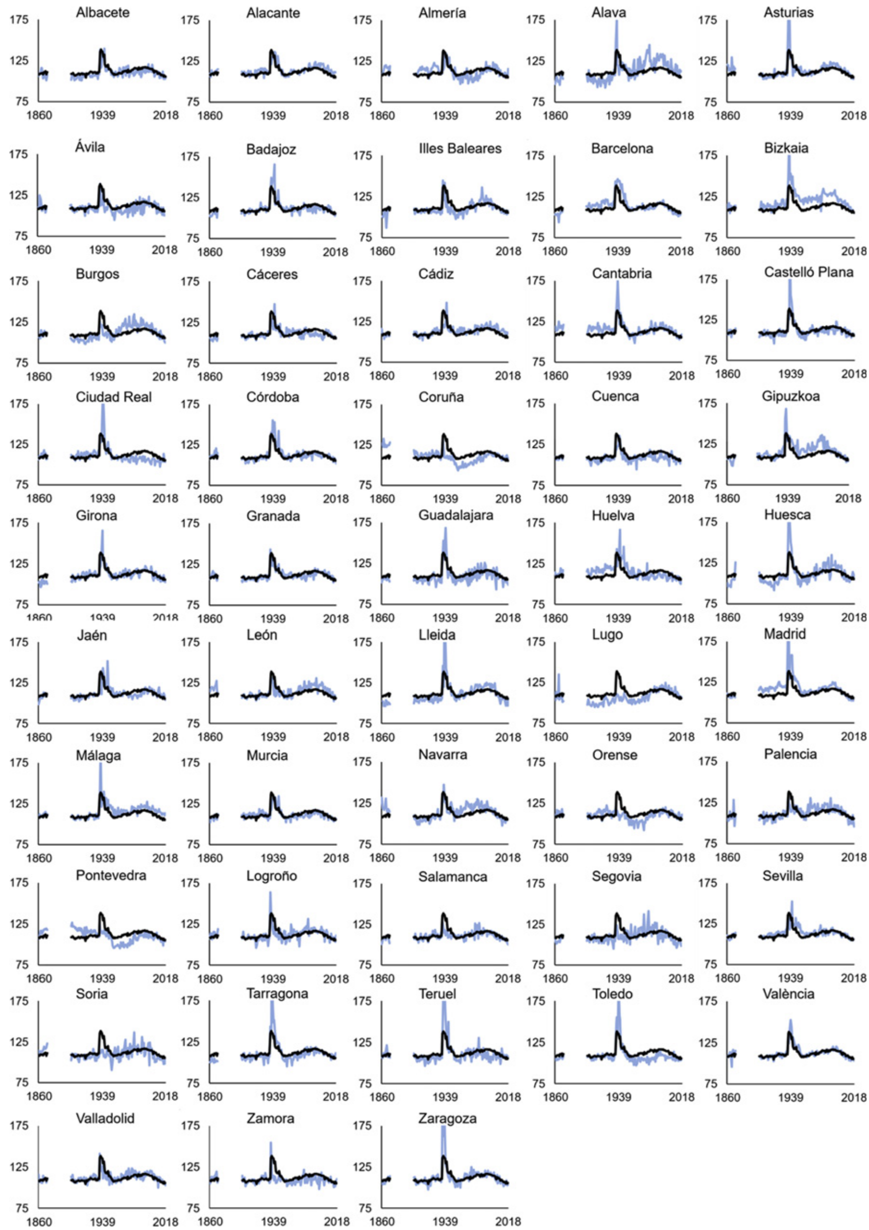

Fig. 4. Sex Mortality Ratios in the Spanish Provinces. *Notes:* Light blue lines show annual sex mortality ratios, and dark black lines show annual average sex mortality ratios in Spain. As explained in Section 2, we are not showing data from Ceuta, Melilla, Santa Cruz de Tenerife, and Las Palmas. *Sources:* Mortality data are from *Movimiento Nacional de la Población* and population data from census years; for more details, see Section 2.

the annual provincial crude death rates into groups of 20 years (1900–1919, 1920–1939, etc.) and use the following equation:

$$\mathrm{CDR}_{p_t} = \alpha + \gamma_r + \delta_t + \epsilon_{p_t} \tag{1}$$

where p denotes provinces ($p = 1,\ldots, 48$), t years (for instance, the group 1900/19 will include the years 1900, 1901, …, 1919, the group 1920/1939 the years 1920, 1921, …, 1939, and so on), *CDR* denotes our outcome variable (crude death rates), γ_r are region fixed effects (for region here, we deploy the *Comunidad Autónoma* distinction, one of the highest levels of aggregation and above provinces, $r = 1,\ldots, 16$), δ_t are year fixed effects within the years of the different time groups (for instance, the group 1900/1919 adds dummies for the years 1900,…,1919, the group for 1920/1939 adds dummies for the years 1920,…,1929, etc.), and ϵ_{p_t} is the error term, clustering the standard errors at the provincial level.[3]

Point estimates in Fig. 5 represent the slope coefficients of the different time periods, measuring the pace of the mortality decline. Grey areas plot 95% confidence intervals. For comparison, in the first place, we present the slopes without region or time fixed effects (unconditional models). Looking down the figure, the means declined gradually, showing that mortality declined over time. This pattern appears to hold for men and women, but means are nearly always higher for men and, in the period 1940–1959, differences in sex mortality get bigger, despite being statistically significant only when adding the regional and year fixed effects. We next investigate how these trends evolved across time and between Spanish regions with a range of inequality measures.

3. GINI AND OTHER MEASURES OF INEQUALITY

Using the annual provincial crude death rates, we next compute the Gini index to measure health inequality. The Gini index is a measure of inequality, which measures the extent to which a variable "*y*" (usually income or wealth) is equally or unequally distributed among individuals. The Gini (expressed in percentages) ranges from 0 to 100, where 0 would denote a perfect egalitarian distribution of "*y*" among individuals and 100 would indicate that "*y*" is fully concentrated by one individual. While the Gini index is generally used to study income and wealth concentration, it has also proved to be useful to assess inequalities in other dimensions, such as life expectancy and mortality (Peltzman, 2009). For its computation, we use the following equation:

$$\mathrm{Gini} = \frac{1}{2n^2\overline{y}} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| y_i - y_j \right| \tag{2}$$

where n is the number of provinces, $\overline{y}$ is the overall mean crude death rate in all provinces, and y_i is the mean crude death rate of people belonging to the i-th province, with provinces ranked in ascending order ($y_j > y_i$). As Peltzman (2009)