Determinants of Health in the City of Sao Paulo: A Spatial Approach

Marislei Nishijima[♣]

Alexandre Sartoris[ϒ]

Abstract

Several studies have shown the relationship between relative income and health, in a way that the bigger the disparity in relative income, the worse health population is. However, studies establishing this relationship in its geographical effects are still uncommon. Taking that into account, this paper aims to investigate health determinants of the inhabitants of the 96 districts in which the City of Sao Paulo is divided, by considering the geographical location of those inhabitants residence in the year 2000. To attain this goal, models using spatial techniques will be estimated using data from the government of the City of Sao Paulo, in order to evaluate the effects of geographical location in the relationship between relative income and health in the City of Sao Paulo. The purpose is to investigate not only the regional effects, but geographical spillover effects and diffusion among regions.

Key words: health, income, inequality, mortality, spatial econometrics

JEL: I10, D31, D63, C21

[♣] University of Sao Paulo – USP.

[ϒ] Sao Paulo State University – UNESP.

1. Introduction

According to Arrow (1963), the individual health has, among its determinants, the medical care market, the food, the living conditions, the life style and the social-economic conditions. Besides these determinants, there are the innate conditions, such as genetic characteristics and other random effects that affect health, such as environment and institutional conditions.

Income and the wealth have been often traditionally to relate socio-economic variables with health status in the international literature. In general, per capita income has been related directly to the health in a way that the population of the rich countries is healthier than in the poorest ones. These results are evaluated considering that the richest countries have lower mortality and infant mortality rates and higher life expectancy. Many studies, including some more recent ones using microdata, have shown that income is an important determinant of the health (Case et al., 2002; Lawson, 2004).

Preston (1975) compared many countries, in a cross countries study, and showed the life expectancy increases as per capita income grows, but this effect is bounded since, once a high level of per capita income is reached, an extra growth in per capita income does not improve the life expectancy any more. So, as an extra outcome of his empirical analysis, he concludes that countries with better pattern of income distribution have healthier individuals than countries with unequal income distribution.

Following the last statement of Preston, many studies have been done to evaluate empirically this hypothesis, such as Wilkinson (1992 e 1996) and Waldmann (1992). These studies were drove to empirically verify this hypothesis – known in the literature as relative income hypothesis. These authors found evidences of a negative significant relation between income inequality and health. These studies generated more studies and debates about this hypothesis, which became very controversial, since the validity of this proposition implies that economic policy of income equality is also health policy.

Deaton (1999 e 2001) is one of the authors which strongly argued against the relative income hypothesis. He found empirical evidences against it, but his main argumentation, however, is the lack of theoretical basis for the relative income hypothesis.

Considering the large size of Sao Paulo city, with a population of nearly 11 million inhabitants, this paper aimed to verify if the relative income hypothesis is valid for its 96 districts. The use of spatial econometric techniques allows evaluating if the income distribution has a role in explaining the health of the population. So, this study contributes with the empirical debate about relative income hypothesis bringing an additional element to the analysis as did Shin (2006). This analysis is interesting because allows controlling by spatial correlation among the districts of Sao Paulo city, which are easily classified according their socio-economic conditions. For example, the Morumbi district, although shelters slums, has surely a much better socio-economic standard than Cidade Líder district, a very poor region. The paper also verified the effects of average income and of the wealth on the health of the Sao Paulo habitants.

For the city of Sao Paulo, these results are very important because they could be able to guide public policies. Especially because Brazil has one of the worst income distribution pattern in the world and this pattern is seen in different regions of the country, with slight differences.

The next part of the paper presents a brief survey of the applied and theoretical literature about the relative income hypothesis. In part 3 we discuss the method of spatial analysis and its results. In the final part we present the conclusions.

2. Bibliographic survey

Preston (1975), in its cross countries study, showed that life expectancy increases as per capita income increases, but there is a superior bound that, once reached, increases in per capita income don't affect life expectancy anymore. This author interpreted this empirical result as being consequence of the non-linearity of the relationship between income and health. Besides this interpretation, his study – which becomes seminal in income inequality and health literature – is now the basic reference for the conclusion that countries with better income distribution have healthier individuals.

Wilkinson (1992) related empirical evidences that support the relative income hypothesis by using the mortality rate in different countries. The author argues that, besides the correlation between average income and mortality rate being low in developed countries, the mortally rate is highly correlated with the pattern of income distribution. So, countries with less social inequality tend to present better indicators of life expectancy.

Wilkinson (1996) did an empirical study, using data from the United Kingdom and the United States between 1970 and 1980, and concluded that income inequality was statistically significant to explain mortality rates.

Waldmann (1992) also found empirical evidences that infant mortality is strongly related to income distribution and, more specifically, that the level of income of the richest is positively correlated with infant mortality. The author did empirical studies for two groups of countries: one which includes all available countries in the *Word Table* of 1976 and in the *Population Statistics Report* of 1997; and other using only developing countries from the same sources. His results are robust even when the author controls by real income of the poorest, by the supply of services of medical care available for the poorest and by other variables. His interpretation of the phenomenon, which he considered anomalous, is that the income is not a good measure of the standard of living of individuals.

Deaton (1999) did an empirical study for the OCDE countries and another using data from American individuals in cohorts and didn't find evidences of income distribution being a risk factor to health. His results were contrary to Wilkinson's proposition.

Deaton (2001) argued against the validity of the relative income hypothesis, by stating that favorable empirical results of other researchers must reflect other effects than income inequality. Deaton claims that there is no robust empirical evidence favorable to the relative income hypothesis in developed countries and for the others countries the hypothesis has been empirically verified just because the unequal income distribution is an indicator of poverty. According to him, the income is a determinant of health, not its distribution. The author points out to the problem of lack of quality in the income distribution data used in the cross-countries studies. These data are not adequate because they are calculated using different methods for each country. And even if the study uses data from the same region, still remains the problem of the correlation between income inequality and health not being robust through the time, because empirical studies usually use cross-section data. Else, still remains the uncertainty of whether mortality comes from income inequality or from other factors correlated to the standard of living.

Additionally, observes Deaton (2001), the empirical studies about income inequality and health, in general, don't consider the double direction of causality between these variables. Since income affects health, it is true that health also affects the capacity of obtaining income. This causes endogeneity problems for the explanatory variable and biased estimations. The author also presents a discussion about the inadequacy of income inequality indicators in reflecting these measures effectively.

Deaton (2001) reports Mellor and Milyo (2001) to illustrate a study which reproduces Wilkinson (1996) using a larger sample, between 1950 and 1990, from the census of the 48 American states. Their results show high correlation between Gini coefficient and all mortality causes when only year dummies, age composition of the state and the median income are included in the model as controls. However, when they control by the average level of education in each state, the coefficient associated with the Gini index became not statistically significant.

For Deaton (2001) the income distribution itself is not a risk factor for health, but broader socio-economic issues associated with inequalities. He does not deny, however, the importance of income as a determinant of health. Throughout his argumentation, he performs a theoretical and empirical analysis about the health determinants and concludes that most empirical studies don't have a theoretical model able to support the relative income hypothesis. Doing this analysis, the author generates a quite adequate bibliography survey, which maps all important empirical contributions for this subject to the date.

Gravelle et al (1998) also disagree of the relative income hypothesis. They consider that the empirical literature pro the hypothesis use empirical "tricks" to ensure that income distribution may affect health.

Beckfield (2004) replicates a previous study among different countries, using broader data and more control variables, to verify if the relative income hypothesis is robust to these changes. He also tested different types of measure of income inequalities. His results shown: the correlation between health and income inequalities attenuates with the inclusion of the new control variables; and the correlation vanishes when he estimates fix effects models, which controls by non-observables heterogeneities of each country.

Shin (2006) explores the relative income hypothesis using spatial econometrics, which allows him to consider the geographical distribution of the income inequality, putting an additional dimension in the debate. His results point out the validation of the hypothesis and the importance of the geographical factor to the phenomenon.

Shin (2006) considers the geographic factor as been crucial for the empirical analysis because health, in its different concepts, also depends on the place were the individual lives, which is conditioned by climate, status of war, and other institutional characteristics. The author uses, as theoretical fundamentals of the relative income hypothesis, the idea of the two phases of epidemiologic transition from Preston (1975). This idea assumes that life expectancy and mortality rates are good measures for the population state of health. The epidemiologic phases are: the phase of infectious diseases, when there is high mortality due to infectious-contagious disease and high level of fertility among the younger population predominant in poor and developing countries; and the phase of chronic disease, when people die mostly by chronic diseases associated with advance of age, predominant in developed countries. His interpretation of the phenomenon is the same of Wilkinson (1996), namely, countries that become rich and get out the epidemiologic frontier change their mainly cause of mortality from material deprivation to social disadvantage. Taking in account that the passage between these two phases doesn't take place simultaneously through time and space, the spatial analysis is justifiable.

The core of the controversy in the literature about the relative income hypothesis, surveyed above, can be viewed as being basically by the search of robust empirical evidence favorable or unfavorable to the hypothesis. This because, according to Deaton (2001), the theoretical literature, except some punctual possibilities in models – such as the model of Alesina et al. (1999) – didn't support the idea of the income inequality being a risk factor for the population health. By this point of view, this work tried to contribute with an additional empirical study, taking into account the dimension of the geographic factor in the analysis, which is not well exploited in the empirical studies yet, with a few exceptions.

3. Spatial analysis

Inspired by the method of spatial analysis applied to the problem of relative income hypothesis by Shin (2006) and taking into account the controversy summarized above, this paper empirically investigates whether the hypothesis is valid to the city of Sao Paul. This city is divided in 96 districts placed in 5 macro-regions, with very different standards of living. The paper also analyses if income itself is a statistically significant determinant of the city population health.

To test the relative income hypothesis for the city of Sao Paulo we used data from the city government, from DIEESE (Inter-labor Unions Department for Socio-economic and Statistical studies) and IBGE (Brazilian Institute of Geography and Statistics) for the year 2000.

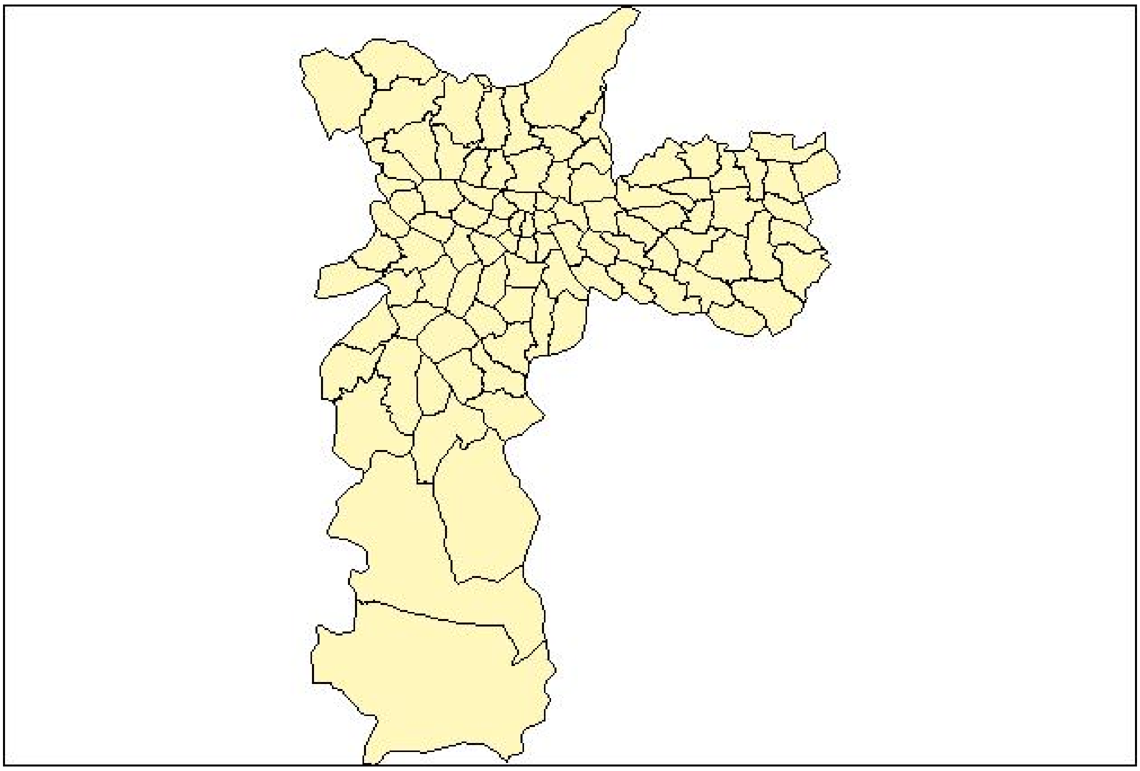

Figure 1: the city of Sao Paulo and its districts

As our analysis refers to distinct regions of the same place, the city of Sao Paulo, differently from cross-countries studies, the problem of the quality of the data pointed out by Deaton (2001) is minimized, since there are no differences in methodology for building the same variable among the city districts.

To consider the problem pointed out by Deaton (2001), which states that the income may not be a good measure for the standard of living, or the socio-economic variable, we also tested a model using the inequality of the wealth as a socio-economic variable. We used as proxy of the wealth inequality the percentage of occupation of districts land by high standard residences.

In order to empirically evaluate whether the relative income hypothesis is valid for the city of Sao Paulo, we are going to use spatial analysis methods that will allow as taking into account if the spatially distributed data have dependence or heterogeneity in their structure.

In terms of traditional Econometrics, the presence of such effects might require some modification or even invalidate its results. In many cases, it is necessary to create new techniques for the correct treatment of these effects. As pointed out by Anselin

(1988), these issues are often overlooked by the traditional Econometrics, hence the creation of a separated field of Spatial Econometrics.

From the methodological point of view, Spatial Econometrics deals with lattice data. A lattice of locations comes from the idea of regions connected to their neighbors, as shown in the figure below:

1	2	3
4	5	6

Figure 2: a regular lattice of 6 regions

Figure 2 shows a regular lattice, while figure 1, the map of the city of Sao Paulo is an irregular lattice.

Region 1 is neighbor of regions 2, 4 and 5, while region 5 is neighbor of all the other regions. These neighborhood relationships can be represented by the contiguity matrix **w**:

$$\mathbf{w} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Or by the matrix **W**, which is the matrix **w** row normalized in order to all rows sum up to 1.

$$\mathbf{W} = \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{bmatrix}$$

This is a binary matrix **W**. There are other ways to build a contiguity matrix by using some kind of measure that establishes the weight of neighbors, like, for instance, the distance between two neighbors. Conley and Lugon (2002) suggest a measure of economic distance instead of the geographic one. However, since we are dealing with relatively small regions (districts), we are going to choose the binary matrix.

3.1 Exploratory spatial data analysis – spatial autocorrelation

Moran's index (I), given below, is often used as a measure of (total) spatial autocorrelation.

$$\mathrm{I} = \left(\frac{n}{S_0}\right)\left(\frac{\mathbf{Z'wZ}}{\mathbf{Z'Z}}\right) \qquad (1)$$

Z is a vector of observations through space.

The I statistic ranges from -1 to +1 and provides a general measure of the linear association between the components of the **Z** vector and its neighbors' average, the spatial lags (**wZ)**. Its expected value is given by:

$$\mathrm{E(I)} = -\frac{1}{n-1}$$

As n tends to infinity, E(I) tends to zero. This means that if I is near zero, there is no significant spatial autocorrelation, as positive values indicate positive autocorrelation and vice-versa.

If we use a row normalized contiguity matrix, it can be shown that (1) becomes:

$$\mathrm{I} = \frac{\mathbf{Z'WZ}}{\mathbf{Z'Z}}$$

Moran's diagram compares normalized values of a variable in space with the normalized neighbors average, generating a two-dimensional plot of **Z** versus **WZ**, so we can visualize spatial dependence of such variable

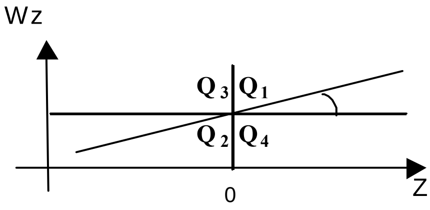

Figure 3: Moran's diagram.

Moran's I will be the slope of the regression of **WZ** against **Z**.

The first quadrant, Q1, the high-high quadrant, shows high values for the variable to both the region and its neighbors. Q2, the low-low quadrant, shows regions with low values, as well as their neighbors. If the region has low values, but is surrounded by neighbors with high values, it will be plotted in the Q3 quadrant (low-high), as well as regions with high levels surrounded by low levels regions will be in Q4 (high-low).

3.2 Empirical results for the spatial exploratory data analysis

Since it is considered an important determinant of health, Figure 4 shows the spatial distribution of income, measured in minimum wage units:

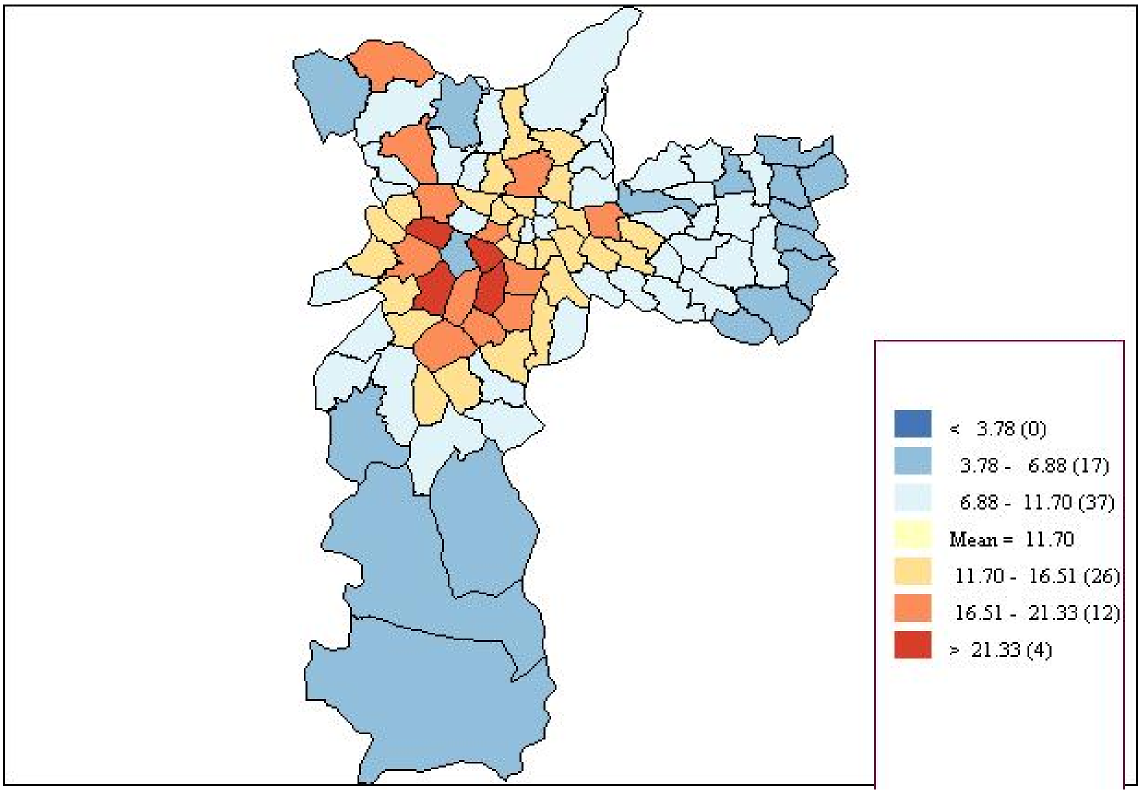

Figure 4: spatial distribution of income.

Since income grows towards the center (even if not to the center itself), figure 4 suggests the existence of a positive autocorrelation in the variable income. In fact, Moran's I is almost 0.5 and the Moran's diagram is show on figure 5.

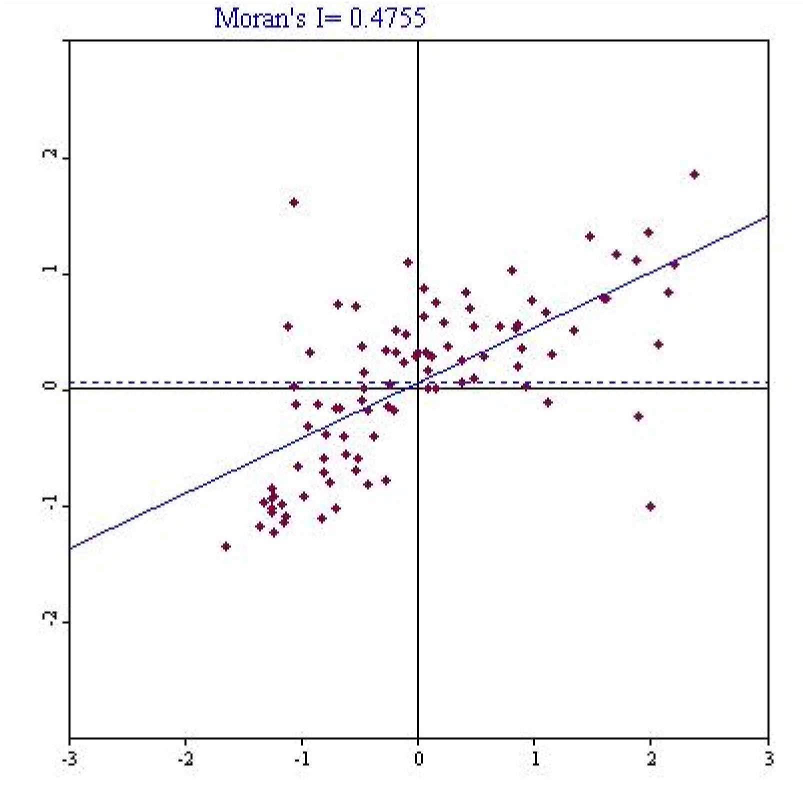

Figure 5: Moran's diagram for income.

Figure 6 shows the spatial distribution of Gini's coefficient through Sao Paulo districts:

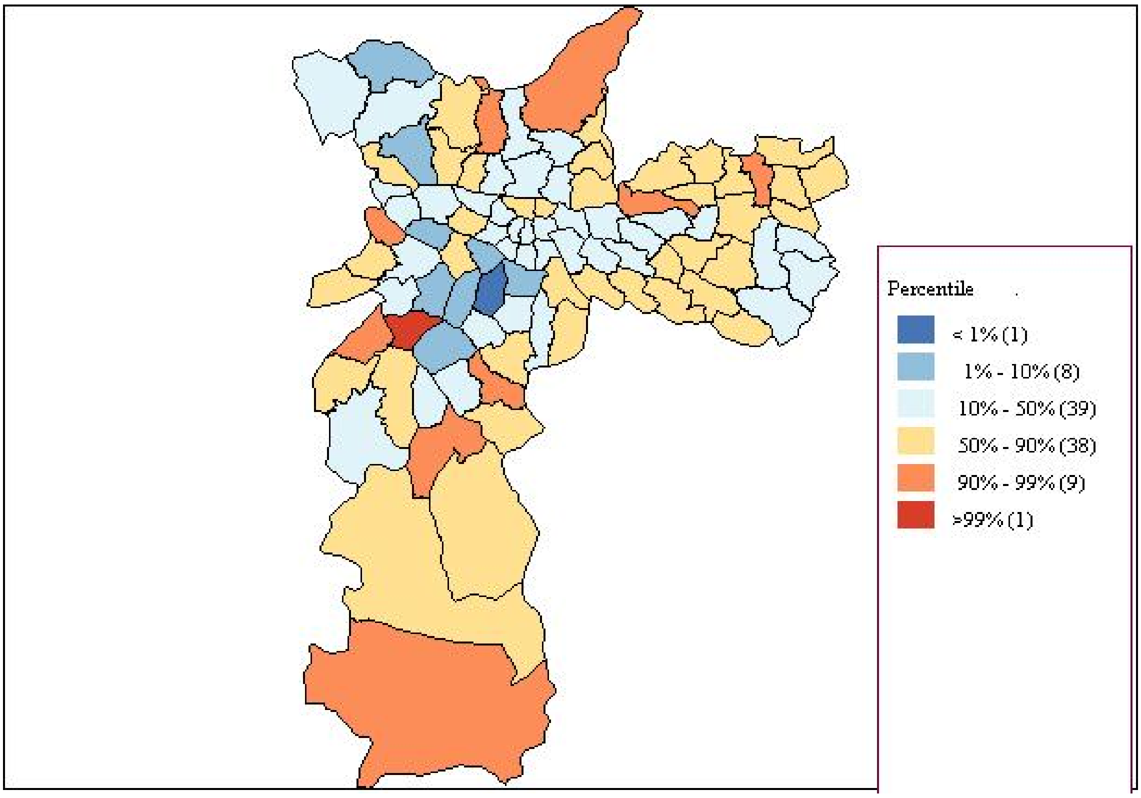

Figure 6: spatial distribution of Gini's coefficient

Gini's coefficient is also positively autocorrelated, but the situation is the reverse when compared to income, since higher Gini's coefficients tend to be on the more distant districts. In fact, there is a high correlation (in absolute value) between Gini's coefficient and income (-0.878). Moran's diagram for the Gini's coefficient is shown in figure 7.

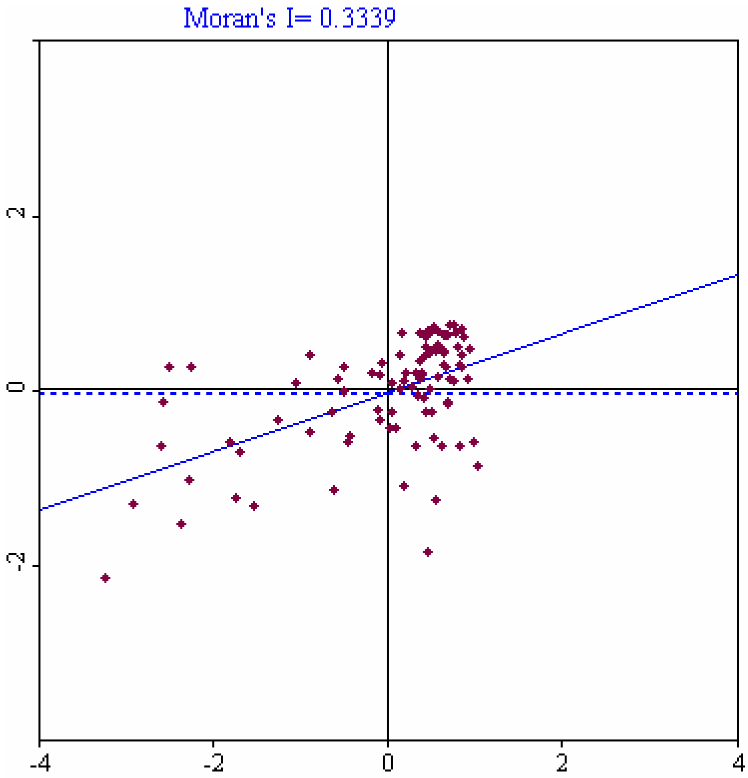

Figure 7: Moran’s diagram for Gini’s coefficient

The measure of the state of the population health, following several authors, as seen in section 2, will be mortality rate. In fact, we are going to use 3 of those rates: the general mortality rate, the infant mortality rate and also the aids mortality rate. Figures 8 to 10 show the spatial distributions of such rates.

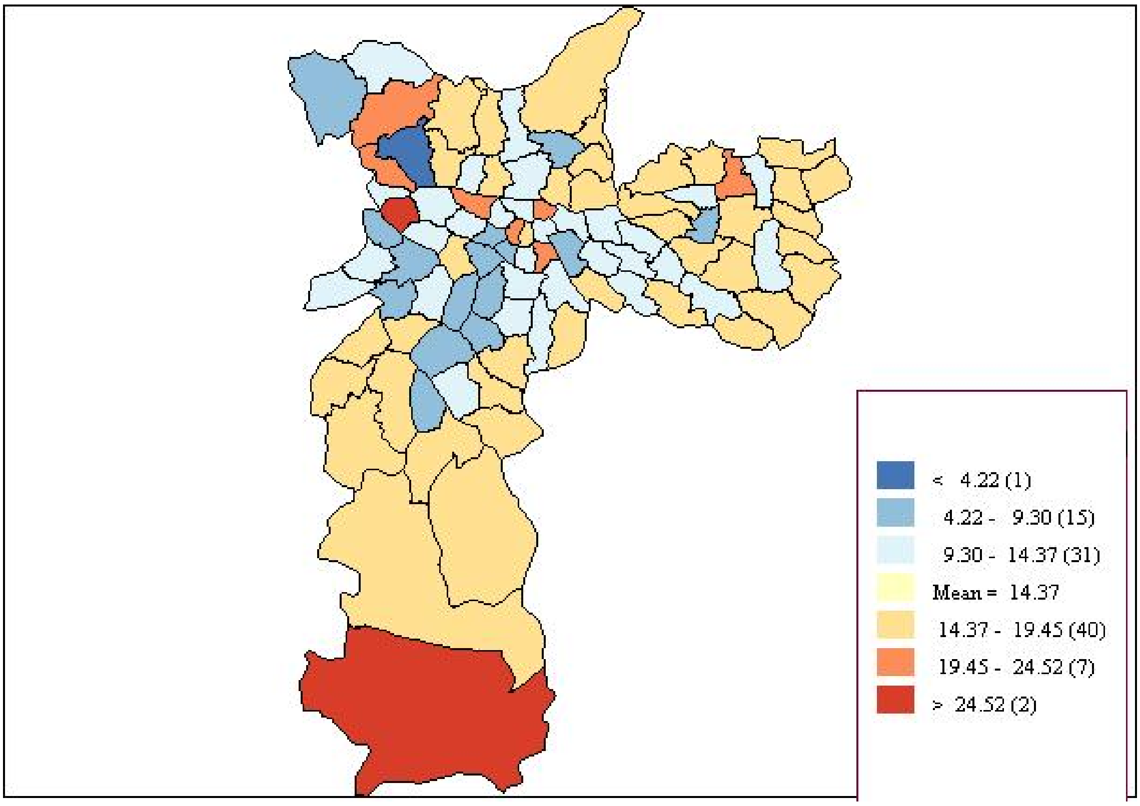

Figure 8: spatial distribution of infant mortality rate.

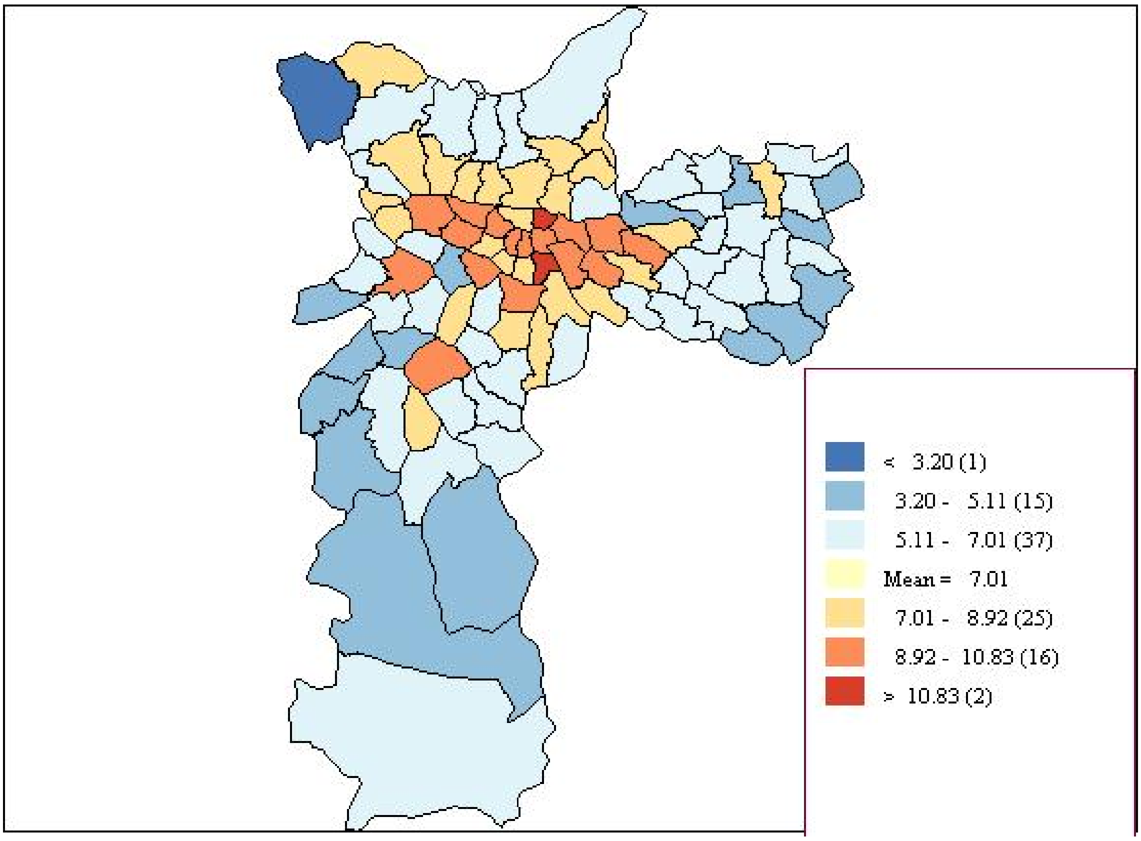

Figure 9: spatial distribution of mortality rate.

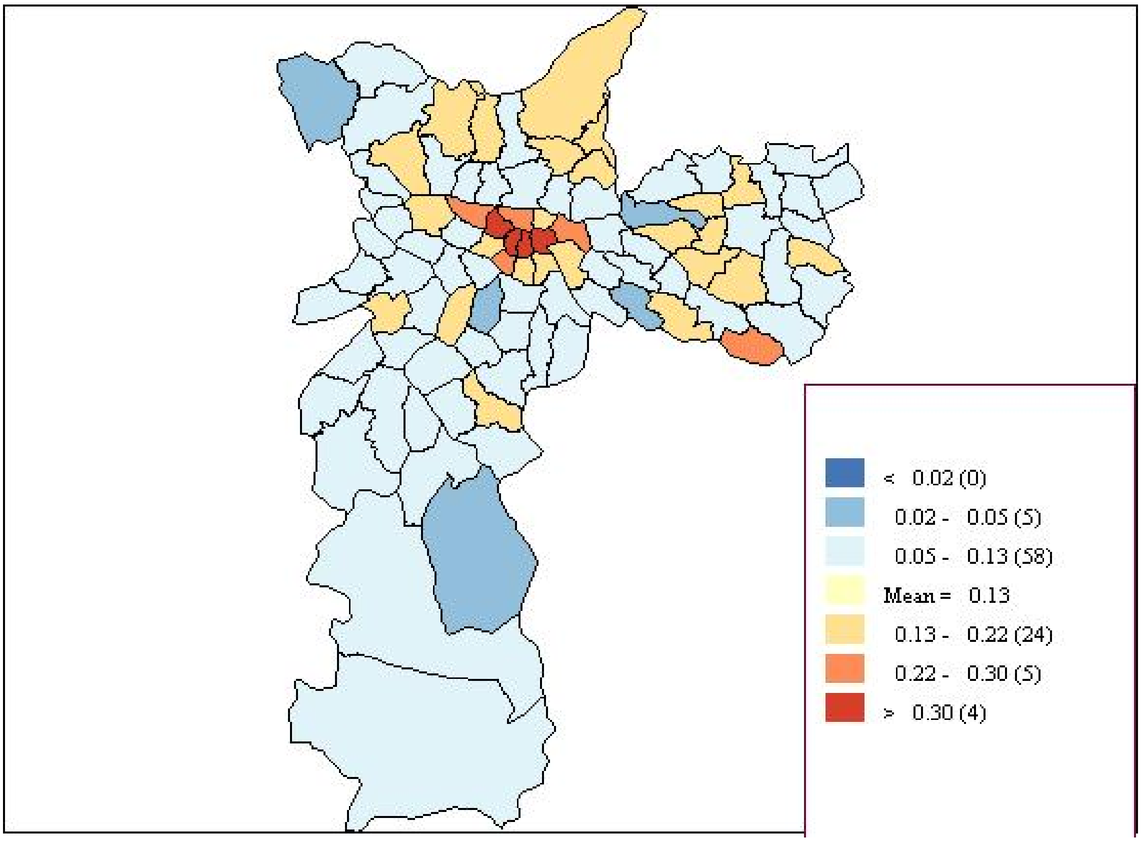

Figure 10: spatial distribution of aids mortality rate.

Figures 9 and 10 suggest positive autocorrelations for both mortality rate and aids mortality rate. In fact, Moran's I for those variables are 0.5069 and 0.4832. In figure 8, it seems there is no significant spatial correlation for infant mortality rate, and its Moran's I is 0.0785.

3.3 Spatial Econometrics

A typical Spatial Econometrics model is a regression equation that takes in account possible spatial dependence (and/or spatial heterogeneity).

A general model for spatial dependence is given below:

$$\mathbf{Y} = \rho\mathbf{WY} + \mathbf{X\beta} + \boldsymbol{\varepsilon},$$

$$\boldsymbol{\varepsilon} = \lambda\mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\mu}.$$

$\mathbf{Y}$ is the vector of dependent variables, $\mathbf{X}$ is a matrix of explanatory variables, $\boldsymbol{\varepsilon}$ is a random terms vector and $\mathbf{W}$ is the contiguity matrix.

We are going to consider two particular cases: $\rho = 0$ and $\lambda = 0$, each one with specific econometric problems. If both ρ and λ equal zero, we have a classical linear regression model.

In the event of $\lambda = 0$, we have a spatial lag model given by:

$$\mathbf{Y} = \rho\mathbf{WY} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

In this case, we have an endogenous explanatory variable, **WY**. OLS estimation will lead to biased and inconsistent estimators.

If, however, ρ equals zero, then we have a model with spatial error autocorrelation:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$$\boldsymbol{\varepsilon} = \lambda\mathbf{W}\,\boldsymbol{\varepsilon} + \boldsymbol{\mu}.$$

Autocorrelation in the random term, no matter if temporal or spatial, leads to inefficient estimators when OLS are used.

A consistent and efficient estimator for such models will be given by maximum likelihood estimation.

Note that a model that includes spatial lags of the exogenous variables, such as:

$$\mathbf{Y} = \mathbf{X}\,\boldsymbol{\beta}_1 + \mathbf{WX}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon},$$

This model is, everything else remaining the same, a classical linear regression model because, since **X** is a set of exogenous variables, **WX** will be exogenous as well.

3.4 Model results

The variables used in the estimation are: INCOME (measured in minimum wage units); GINI (Gini's coefficient); INFANT (infant mortality rate); AIDS (AIDS mortality rate); MORTAL (general mortality rate); HOSPITAL (number of hospitals in the district); HSTAND (high standard residences in percentage of land use); FAVELA (area occupied by "favelas" – communities of very poor people who live in poorly built materials – in the district) and INFEC (number of cases of infectious diseases – not necessarily fatal – in the district). The notation W_variable indicates a spatial lag of such variable, i.e., W_variable = **W**×variable.

Table 1 shows the estimations for mortality rate as the dependent variable:

Table 1- dependent variable: mortality rate (standard errors in parenthesis)

variables	*(1)*	*(2)*	*(3)*	*(4)*
INCOME	–	0.3137*** (0.6070)	–	0.3468*** (0.0700)
GINI	-14.729*** (4.0391)	-1.5810 (4.4153)	-17.640*** (3.9628)	-0.1617 (5.1460)
HOSPITAL	0.0277 (0.0536)	–	0.04666 (0.05717)	–
FAVELA	-2.171e-5** (9.571e-6)	-1.328e-5* (7.867e-6)	-2.269e-5** (1.0927e-5)	-1.451e-5** (7.566e-6)
HSTAND	-7.5051** (3.3980)	-12.136*** (3.1470)	-10.423*** (3.4705)	-12.921*** (3.2011)
INFEC	0.00148 (0.00809)	–	-0.00132 (0.00860)	–
W_HOSPITAL	0.1253 (0.1012)	–	0.1903 (0.1337)	–
W_FAVELA	-4.354e-6 (2.259e-5)	–	-7.983e-6 (2.741e-5)	–
W_HSTAND	-11.361*** (3.9989)	-12.378*** (3.2643)	-11.475** (5.6424)	-9.1942** (4.5466)
W_INFEC	-0.0192 (0.0180)	–	-0.02312 (0.02117)	–
Ρ	0.5088*** (0.1039)	0.4484*** (0.0884)	–	–
Λ	–	–	0.6063*** (0.0970)	0.5790*** (0.1011)
AIC	334.452	305.535	332.257	307.970

(1) spatial lag model with all variables, except income.
(2) spatial lag model with few variables, but including income.
(3) spatial error model with all variables, except income.
(4) spatial error model with few variables, but including income.
* significant at 10% level
** significant at 5% level
*** significant at 1% level

The results suggest spatial dependence in the mortality rate.

The number of hospitals, whether they are in the district or nearby, doesn't seem to be relevant to the mortality rate. The same happens with the number of infectious diseases cases.

The area occupied by "favelas" and high standard residences are significant, including high standard residences in surrounding districts. However, the expected sign for "favelas" would be positive, since it represents more poverty in the district.

Gini's coefficient was significant only when income was excluded. That result seems to reinforce Deaton's (2001) claim that is the income, not income inequality, a factor that explains health levels. The signs for income and Gini were not according to expectation, though.

Table 2 shows the estimations for infant mortality.

Table 2 - dependent variable: infant mortality rate (standard errors in parenthesis)

variables	*(1)*	*(2)*	*(3)*	*(4)*
INCOME	–	–	–	-0.4648** (0.1894)
GINI	32.4582** (14.0551)	31.6619** (13.9844)	33.2604** (14.9091)	10.8544 (14.9117)
HOSPITAL	-0.2099 (0.1861)	-0.2146 (0.1874)	-0.2155 (0.1974)	–
FAVELA	5.595e-5* (3.308e-5)	5.502e-5* (3.295e-5)	5.558e-5 (3.509e-5)	–
HSTAND	-2.9109 (11.7352)	-2.7923 (11.7703)	-1.7351 (12.4296)	–
INFEC	-0.01415 (0.02823)	-0.01368 (0.02840)	-0.1190 (0.0299)	–
W_HOSPITAL	0.1317 (0.3423)	0.1839 (0.3280)	0.1656 (0.3587)	–
W_FAVELA	4.4190e-5 (7.5737e-5)	4.3282e-5 (7.3367e-5)	3.143e-5 (7.989e-5)	–
W_HSTAND	-22.5251 (14.5693)	-19.1572 (13.4711)	-19.3768 (14.6084)	–
W_INFEC	-0.07496 (0.06243)	-0.07233 (0.06161)	-0.07653 (0.06619)	–
P	-0,1207 (0.1579)	–	–	–
Λ	–	-0.1006 (0.1594)	–	–
AIC	568.346	566.549	566.864	553.337

(1) spatial lag model with all variables, except income.
(2) spatial error model with all variables, except income.
(3) classical model with all variables, except income.
(4) classical model with few variables, but including income.
* significant at 10% level
** significant at 5% level
*** significant at 1% level

Since there was no kind of spatial dependence – neither in the dependent variable, nor in the error – it makes sense to estimate a classical linear regression model. But it seems that infant mortality rate does not respond to anything but income/income distribution.

Once again, nevertheless, Gini's coefficient only was significant when income was excluded.

Table 3 shows the results for AIDS mortality rates.

Table 3 - dependent variable: AIDS mortality rate (standard errors in parenthesis)

variables	*(1)*	*(2)*	*(3)*	*(4)*
INCOME	–	-0.001389 (0.001509)	–	0.001472 (0.002512)
GINI	-0.1757 (0.1930)	–	-0.2741 (0.1869)	–
HOSPITAL	-0.00278 (0.00255)	–	0.00344 (0.00291)	–
FAVELA	-7.079e-7 (4.569e-7)	–	-1.092e-6** (5.571e-7)	-5.857e-7 (4.048e-7)
HSTAND	-0.2467 (0.1615)	–	-0.3749** (0.1742)	-0.1877 (0.1491)
INFEC	0.000737* (0.000387)	-0.000444 (0.000367)	0.000762* (0.000432)	0.000617* (0.000374)
W_HOSPITAL	0.003167 (0.3423)	–	-0.00349 (0.00716)	–
W_FAVELA	-3.913e-7 (1.0576e-6)	–	-1.591e-6 (1.410e-6)	–
W_HSTAND	-0.2282 (0.1944)	–	-0.5158 (0.3165)	–
W_INFEC	-0.000972 (0.00085)	-0.001191* (0.000684)	-8.105e-6 (0.00108)	–
Ρ	-0,6897*** (0.08184)	0.7356*** (0.0751)	–	–
Λ	–	–	0.7528*** (0.07215)	0.7377*** (0.07500)
AIC	-242.719	-246.546	-245.462	-248.715

(1) spatial lag model with all variables, except income.
(2) spatial lag model with few variables, but including income.
(3) spatial error model with all variables, except income.
(4) spatial error model with few variables, but including income.
* significant at 10% level
** significant at 5% level
*** significant at 1% level

AIDS mortality rate seems to have little to do with income or its distribution. The only thing that is clearly present in all models is the spatial dependence. It seems to have a correlation with the number of cases of infectious diseases, which is a reasonable result.

4. Conclusions

Empirical results clearly show the existence of spatial dependence for the mortality rate and also for AIDS mortality rate. Neglecting those effects while estimating econometric models for these variables might lead to wrongful estimates.

There seems to be no spatial dependence for infant mortality rate. The results suggest that what is really determinant for infant mortality is income (or income inequality), and other random factors.

As for the controversy of the relative income hypothesis, the results seem to reinforce Deaton's claim that it is the income that matters to health, not income inequality. However, it is possible that this behavior could be partially due to the high correlation (in absolute value) between income and income inequality that is observed throughout the districts of the city of Sao Paulo.

References

ALESINA, A.; R. BAQIR; and W. EASTERLY (1999), "Public goods and ethnic divisions," *Quarterly Journal of Economics*, 114, 1243–84.

ANSELIN, Luc (1988) **Spatial Econometrics: Methods and Models**. Kluwer Academic Publishers, Dordrecht.

ARROW, Kenneth. J., (1963) "Uncertainty and the Welfare Economics of Medical Care". *The American Economic Review*, vol. 53, n.5.

BECKFIELD, J. (2004), "Does Income Inequality Harm Health? New Cross-National Evidence". *Journal of Health and Social Behavior*, Vol. 45, No. 3., pp. 231-248.

CONLEY, T.G. e E. LUGON (2002) "Economic Distance and Cross-Country Spillovers", Journal of Economic Growth, n.7, p.157-87.

DEATON, A. (1999) "Inequalities in income and inequalities in health" *NBER Working Paper* No. 7141.

DEATON, A. (2001) "Relative deprivation, inequality, and mortality," Princeton University, Research Program in Development Studies and Center for Health and Wellbeing, *Working paper*, Princeton NJ. (*NBER Working Pape*r No. 8099)

GRAVELLE, H. (1998) "How much of the relation between population mortality and unequal distribution of income is a statistical artefact?" *British Medical Journal*, 316, 382–85.

MELLOR, J. M., and J.MILYO (2001), "Re-examining the evidence of an ecological association between income inequality and health," *Journal of Health Politics*, Policy, and Law.

PRESTON, S. H. (1975) "The changing relation between mortality and level of economic development," *Population Studies*, 29, 231–48.

SHIN, M. E. (2006). "Income Inequality and Population Health in Geographic Perspective". California Center for Population Research, *On-Line Working Paper Series*, CCPR-028-06.

WALDMANN, Robert J., 1992, “Income distribution and infant mortality,” *Quarterly Journal of Economics*, 107, 1283–1302.

WILKINSON, Richard G., 1992, “Income distribution and life-expectancy,” *British Medical Journal*, 304, 165–68.

WILKINSON, Richard G. (1996). **Unhealthy societies: the affliction of inequality**. London. Routledge.

WILKINSON, Richard G. (2000). **Mind the gap: hierarchies, health, and human evolution.** London. Routledge