A Letters Journal Exploring
the Frontiers of Physics

OFFPRINT

Impact of personal income on mortality. Decomposition into biological vs. socio-economic effects

Peter Richmond, Wonguk Cho, Beom Jun Kim
and Bertrand M. Roehner

EPL, **135** (2021) 14002

Please visit the website
www.epljournal.org

A Letters Journal Exploring the Frontiers of Physics

July 2021

EPL, **135** (2021) 14002 www.epljournal.org
doi: 10.1209/0295-5075/ac115e

Impact of personal income on mortality: Decomposition into biological vs. socio-economic effects

Peter Richmond[1(a)], Wonguk Cho[2(b)], Beom Jun Kim[2(c)] and Bertrand M. Roehner[3(d)]

[1] *School of Physics, Trinity College - Dublin, Ireland*
[2] *Physics Department, Sungkyunkwan University - Seoul, South Korea*
[3] *Institute for Theoretical and High Energy Physics (LPTHE), Pierre and Marie Curie Campus, Sorbonne University, National Center for Scientific Research (CNRS) - Paris, France*

received 21 March 2021; accepted in final form 5 July 2021
published online 3 September 2021

Abstract – Humans have two facets, biological and social. Whereas it is a common idea that poor social conditions affect the biological facet for instance by reducing life expectancy, there are few known cases where an economic effect is affected by the biological facet in a way which can be predicted. The purpose of the paper is to present such a case. In other words, we are going to decompose an economic phenomenon into its biological and social components, a step which provides a marked conceptual simplification. The economic phenomenon that we consider here is one of the most basic that one can think of, namely the relationship between personal income and life expectancy. Intuitively, one is not really surprised that wealthy people live longer than poor people. Here, however, we show an effect which is far less obvious, namely the fact that this relationship does not hold at both extremities of the lifespan interval. The disconnection between income and neonatal (*i.e.*, in the first 28 days after birth) mortality is quite impressive. This observation is particularly significant on account of the fact that the infant mortality (*i.e.*, in the first 365 days after birth) is often taken by economists as a proxy of development when no reliable income data are available. This indicator may be valid for very poor countries where the burden of death due to infectious diseases remains very high, but it is not valid in developed countries such as those considered in this paper. More specifically, we explore the influence of income on mortality by age, separately in France, the United States and South Korea. The same pattern appears for the curves of the (income-mortality) correlation as a function of age. We conjecture that this pattern will be observed in any developed country where the test can be performed. For the test to be possible the main requirement is the availability of income and age-specific mortality data at regional level.

Introduction. –

Interest for econophysicists. Let us start by explaining the possible interest of the present study for physicists.

- To our knowledge it is the first instance in which a well-known economic effect can be decomposed into its biological and social components. We think this is a promising way for decomposing intricate economic or social phenomena into simpler components.

- The decomposition that we describe here was observed in different countries across three continents and despite the idiosyncrasies of their health care systems. In other words one is now in the same situation as in physics when a new phenomenon has been discovered. Usually, in such cases, additional experiments are performed in order to ascertain its occurence requirements. If similar studies are carried out in (for instance) Japan, the Netherlands or Sweden, we posit that they will lead to the observation of the same pattern; cases which do *not* follow the same pattern will raise interesting questions.

What made us focus on this effect? In the abstract we said that we could predict the effect of the age factor even before collecting the data and performing the observation. Why?

In previous papers (*e.g.*, [1,2]) a careful analysis of mortality data led us to the conclusion that despite major medical progress over the past centuty, the mortality rates

(a) E-mail: peter_richmond@ymail.com
(b) E-mail: wonguk.cho@gmail.com
(c) E-mail: beomjun.kim@gmail.com
(d) E-mail: roehner@lpthe.jussieu.fr (corresponding author)

at very young and very old age were almost unaffected. At first we were surprised but after a while the reason became clear.

– The main factor in neonatal mortality (and also in infant mortality albeit to a smaller degree) is represented by congenital defects. Heart, brain, lung or kidney malformations can hardly be corrected by medicine. However, the weak connection between income and death rate shows something more, namely that *in utero* development is fairly well protected from external factors. Drinking alcohol or smoking does have a detrimental effect but may not necessarily result in infant deaths which means that these effects will not be reflected in our death rate data.

– As is well known, adult mortality follows Gompertz's law which states that the death rate is multiplied by a factor two every time the age increases by 9 or 10 years. In other words, the death rate is an exponential function of age. Around the age of 25–30 the mortality rate is so low that the first doublings go unoticed. However, at the age of 85–90 the death rate becomes substantial so that the doublings become quite significant.

  As for the case of young-age mortality, a careful analysis of historical data convinced us that in the course of the past century the death rates of persons over 90 did not change much. This is because at the age of 90 or 100 one is near the biological upper bound of human life (which is about 120 years, see [2]).

The fact that medicine was unable to reduce mortality at young and old age, led us to the fairly natural conjecture that improved living conditions, which come along with higher income, have also little effect on the mortality of these age groups.

Few previous studies. It is known that personal income has a substantial influence on the life expectancy of adults (see, for instance, [3], fig. 1).

The studies published in recent times on the relation between mortality and personal income [4–7] are not much concerned with how this relation changes with age which is the topic of the present letter.

The only studies of this question seem to go back to the 1960s; for instance, Frederiksen [8,9] studied this point as an aside of a broad study of population growth.

Answer to a possible objection. In [10] (p. 484) one reads that the correlation between infant mortality and level of income is commonly found to be of the order of -0.8. This shows a close relationship; actually it is sufficiently strong for infant mortality sometimes to be used as a proxy of income level. How is this compatible with the low correlation mentioned above?

The reason is very simple.

Correlations as high as -0.8 are found in cross-sectional studies involving a broad set of countries ranging from developing to highly developed countries. For such samples the correlation reflects the fact that in developed countries the death toll of infectious diseases has almost vanished, whereas in developing countries it still represents a major cause of death in early infancy. This difference is demonstrated very clearly by the levesl of the annual infant mortality, namely around 3 per 1000 in advanced countries and 10 times more in many developing countries.

Outline of the paper. In the present paper, we consider fairly homogeneous data, namely cross-regional data within three developed countries, France, the United States and South Korea[1].

The paper proceeds as follows.

– As already said, France is the first case that we consider.

– In the second part of the paper we examine the case of the United States. Our main goal is of course to see if the typical pattern observed for France will also be found in the US. As subunits the most appropriate choice consisted in the 50 states (plus the District of Columbia). Counties would be too small especially in the east and regions would be too inhomogenous. In addition to the methodology used in part one we propose what may be called a fast methodology in which one concentrates on "extreme" cases, that is to say on the states which have the lowest and highest incomes. This methodology is based on the realization that if it were possible to conduct this observation as an experiment (in other words, as a prospective investigation) it would not make sense to include in the sample a large group of some 30 states which have almost all the same income. Such a group adds very little to the degree of significance of the correlation.

– South Korea is the third country that we examine. What was our motivation for including South Korea into this study? After observing that our theoretical argument was confirmed in the two cases of France and the US, we became convinced that similar confirmation would be observed in other developed countries. To consider an Asian country (rather than Germany or the UK) was an interesting challenge. Not only are there clear cultural differences, but there is also a much more rapid pattern of development. Will the same "law" hold despite these differences? That was the question and the answer was "yes".

[1]The fact that France is studied first is not because one of the authors is French. The criterion which was used is the number of regional subunits: about 90 in France, 50 in the US and 15 in South Korea. In principle more subunits should give more accurate results but one must also make sure that there are enough individuals in each subunit so as to keep statistical fluctuations under control.

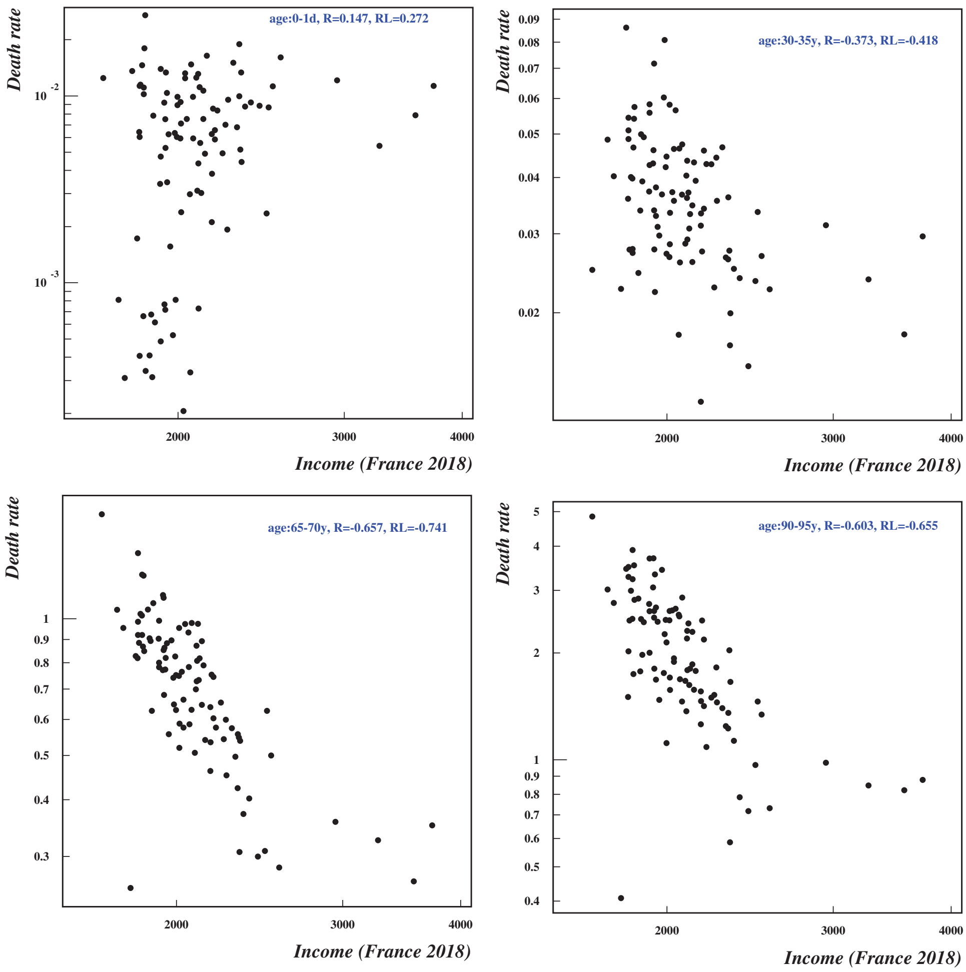

Fig. 1: Relationship between income and death rate in different age groups. (a) age 0–1 day, $R = 0.14$; (b) age 30–35 year, $R = -0.37$; (c) 65–70 year, $R = -0.65$; (d) 90–95 year, $R = -0.60$. R gives the coefficient of linear correlation for ($x =$ income, $y =$ death rate) whereas RL gives the correlation for the logarithms of the same variables. Panel (a) shows the data points for the first age interval whereas Panel (d) is for the last age interval. Panels (b), (c) which are for mid-age intervals, show higher correlations in conformity with life expectancy studies. Sources: INSEE (Institut National de la Statistique er des Etudes Economiques): DC1D, Décès selon le sexe, le groupe d'âges atteints dans l'année. Revenu mensuel déclaré par foyer fiscal.

Age-specific income/mortality correlation in France. –

Data. For our purpose this country has some commendable aspects. i) Metropolitan France is divided into 96 administrative units called "départements". Given the relatively limited size of the country as compared to countries like the United States or China these units are fairly homogeneous areas. ii) French demographic statistics give detailed infancy mortality data (by *département*), starting in the first days after birth and on the old-age side there are data for all 5-year age groups until the age of 100.

Graphs for individual regional units. In fig. 1, each dot represents a regional unit and each graph corresponds to an age group. Two comments are in order.

i) Not surprisingly in most cases the mortality rate is a decreasing function of income. Whenever the correlation is positive, it is rather low and in fact not significant (in the sense that 0 belongs to the confidence interval of the correlation). Because we are chiefly interested in the magnitude (*i.e.*, the absolute value) of the correlation, it is the opposite of the correlation which will be represented in subsequent graphs.

ii) The relationship is not linear. When one draws the graphs of fig. 1 in linear rather than in log-log axes one sees (not shown here) clearly a steep fall for low and mid income which is followed by a part with a much flatter slope. On the log-log plots of fig. 1 the relationship becomes almost linear with the second part being limited to a few data points.

Graph of the correlation as a function of age. Correlations as a function of age are summarized in fig. 2. The term "infancy" which appears in this graph refers usually to the period of a child's life before it can talk. Here we

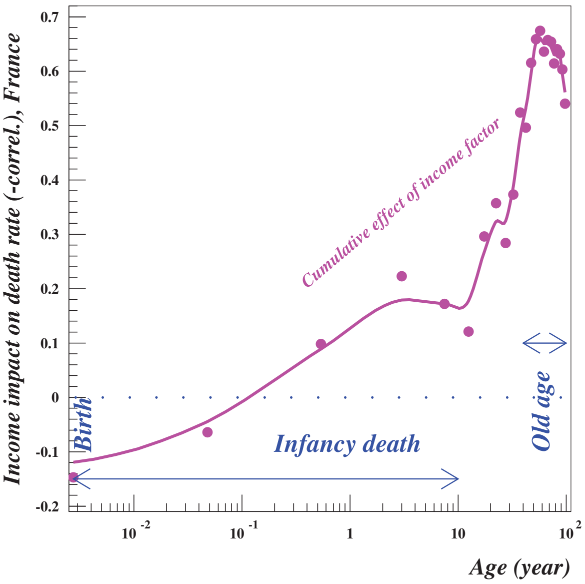

Fig. 2: Income/mortality correlations as a function of age in France, 2018. For convenience of the representation, the graph shows the opposite of the correlations. The parts of the curve which are below the zero line correspond basically to non-significant correlations. The red solid line is a centered moving average over 3 successive data points, *i.e.*: $m(i) = [y(i-1)+2y(i)+y(i+1)]/4$. Sources: INSEE (Institut National de la Statistique er des Etudes Economiques): DC1D, Décès selon le sexe, le groupe d'âges atteints dans l'année. Revenu mensuel déclaré par foyer fiscal.

give the word an extended meaning so that it covers the age interval during which the age-specific death rate is decreasing. This decrease is due to the progressive elimination of new-borns affected by congenital anomalies.

Figure 2. shows the gradual build up of the impact of economic factors. This part of the graph answers the question raised in the introduction about the transition between the biological and socio-economic regimes. It is a gradual, not an abrupt, transition.

Comment. For elderly the decrease of the correlation starts only after the age of 90 and even then it is a fairly small fall. In the United States we will see a more substantial fall. Before one tries to find a definite reason one should observe that the mortality rates in age groups above 90 are beset with large statistical fluctuations because of the small sizes of such age groups.

Summary of successive steps. Before we consider other countries, let us summarize the successive steps of the investigation.

1) The first, and possibly the most time-consuming step, is to obtain the data. This is highly country dependent. The data may or may not be publicly available on the Internet but one can be sure that they have been recorded. This is fairly evident for income because it is a key macroeconomic variable. The same also holds, though for a different reason, for age-specific mortality; the reason here is that there are international agreements for the publication of death data by causes of death.

   For the present investigation such data are needed, not only nationally, but for each of n regional subunits, *e.g.*, French *départements* or US states.

2) Then, for each of the age intervals, one plots the n data points (x = income, y = death rate) as shown in fig. 1. At the same time one computes the correlation between income and death rate.

3) Finally, the correlations for each age group are plotted in a summary graph whose x-axis is age. This gives a graph similar to fig. 2 on which one can read how the connection between income and death changes with age.

The age-specific income/mortality relationship in the US. – In this section we examine the same phenomenon in the United States.

Data. In the US the income data can be found on the website of the "Bureau of Economic Analysis" whereas the mortality data are available on the "CDC-WONDER" website for detailed underlying mortality.

Alternative procedure based on a selection of states. Let us for a moment assume that one can choose at will the average income level of the states. Then, the most accurate procedure would be to take as many incomes as possible in extreme (*i.e.*, very low and very high) income levels and to cover the mid-income interval with a number of cases providing uniform covering density. This makes sense because the confidence interval of a correlation is determined by two factors: i) the level of the correlation and ii) the number of data points. As one expects the correlations on both ends to be lower than in the middle, a high density of data points will be necessary to define accurately the left- and right-ends of the curve.

How can one use this argument in the real situation where income levels cannot be attributed at will? While it is of course impossible to generate more data points than those provided, one can limit the number of data points falling in the mid-income range. The main advantage of this procedure is that one can get good correlation estimates just by visual inspection of the position of the lowest and highest income cases. This procedure gives a clearer insight into why correlations are low or high. The comparison of fig. 3(a) and (b) shows that with only 10 states one gets a picture very similar to the one based on 51 states.

Comment. The US graph is the only one of the three whose concavity is downward. However, as in all three cases the curvature is small the difference may not be very significant. One would need results for a broader set of countries in order to provide a reliable interpretation.

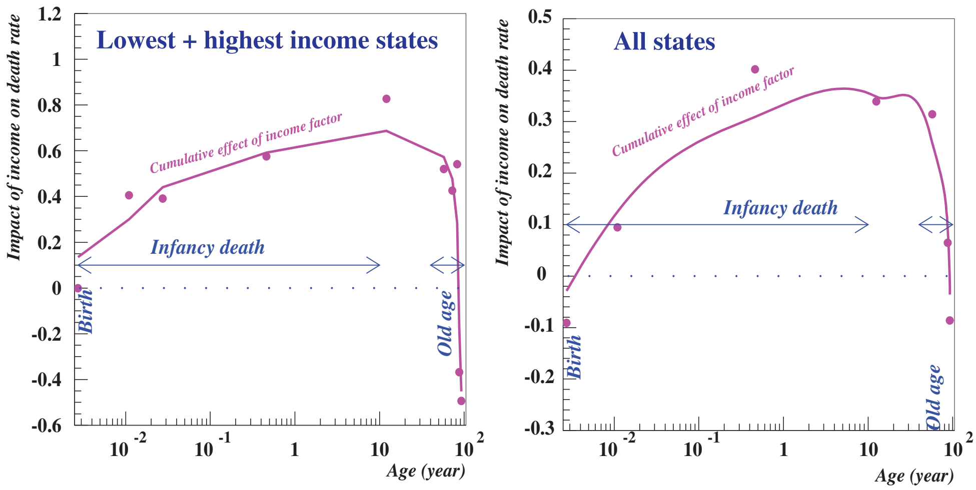

Fig. 3: Income/mortality correlation as a function of age, United States, 2010. Left-hand side: graph in which the focus is on the 5 states with lowest and highest incomes, respectively. Right-hand side: graph for all states. Sources: Personal income by state on the website of the "Bureau of Economic Analysis" (BEA); CDC (Centers of Disease Control) WONDER website for detailed mortality by underlying causes of death.

The age-specific income/mortality relationship in South Korea. –

Data. In the case of South Korea, the data of regional income and mortality rate by province can be obtained from the Korea National Statistical Office (KNSO). However, there is no publicly open source for detailed infant mortality data by province in Korea. Hence, we acquired the raw data of "The Complementary Survey for Causes of Death" from "Microdata Integrated Service" (MDIS) and processed the data to calculate the infant mortality of different age groups by province.

Results. As expected, the mortality rate in Korea (fig. 4) shows little correlation with income at both ends of the life span, where biological factors are more dominant than socio-economic ones. Specifically, the effect of income factors gradually accumulates with age until it reaches the peak at the age of 60–64 ($r = -0.89$), and a sharp fall occurs between the groups of age 70–74 ($r = -0.61$) and age 75–79 ($r = -0.03$).

Comment. This observation confirms that our theoretical argument holds not only in the cases of the Western countries but also in an Asian country despite cultural differences.

At this point we refrain from making more comments for two reasons:

– As this is a letter space is limited.

– In order to make significant comments about the slight differences between the graphs one would need more cases. Here the point we wish to emphasize is the common pattern.

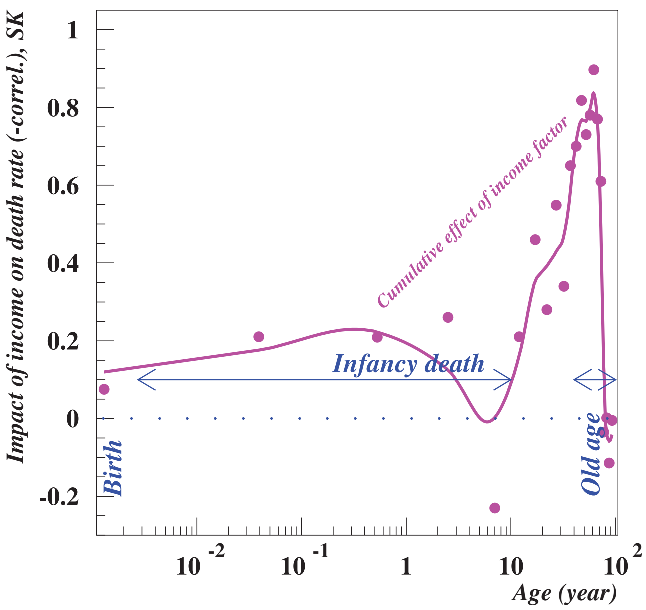

Fig. 4: Income/mortality correlation as a function of age, South Korea, 2017. The data consist of 16 provincial level divisions of Korea. Sources: Regional income and mortality data obtained from the Korea National Statistical Office (KNSO). Infant mortality data are produced based on the survey data acquired from Microdata Integrated Service (MDIS).

Conclusion. –

Interest of an analysis based on regional units. The methodology used in the present paper differs from the more standard approach mainly in two ways.

– Instead of using a sample comprising different countries we analyzed regional units within the same country. In terms of health and mortality the main difference between developing and developed countries is the fact that infectious diseases are still important causes of death in the first whereas they have been almost completely eliminated in the second. For that reason one is not surprised to find a high correlation between income and life expectancy in samples containing a mixture of developing and developed countries. The correlations analyzed here cover effects which are more subtle.

– Most publications mentioned in the introduction took life expectancy as their target variable but this does not allow to study separately the effect of income level on diverse age groups[2]. That is why we took the mortality rate as our main variable.

What should be the next step?. In a physics-like perspective the next step would be to check to what extent a similar pattern holds in other developed countries, for instance Australia, Canada, China, Germany, Italy, Japan or Spain. As in physics, cases which do *not* follow the law will be of greater interest than those which are in agreement with it for then one must find out for what reason the law is violated. This should give a better understanding.

* * *

The authors did not receive any specific funding for this study and do not have any conflict of interest. This study does not involve any experiment with animals that would require ethical approval. Moreover, it does not involve any participants that would have to give their informed consent. WC and BJK were supported by the National Research Foundation of Korea funded by the Korean government (MSIT) Grant No. 2019R1A2C2089463.

Data availability statement: The data that support the findings of this study are available upon reasonable request from the authors.

REFERENCES

[1] Berrut S., Pouillard V., Richmond P. and Roehner B. M., *Phys. A*, **463** (2016) 400.
[2] Richmond P. and Roehner B. M., *Phys. A*, **447** (2016) 446.
[3] Preston S. H., *Popul. Stud.*, **29** (1975) 231.
[4] Marmot M., *Lancet*, **365** (2005) 1099.
[5] Cutler D., Deaton A. and Lleras-Muney A., *J. Econ. Perspect.*, **20** (2006) 97.
[6] Berkman L., Kawachi I. and Glymour M., *Social Epidemiology* (Oxford University Press, Oxford, UK) 2014.
[7] Chetty R., Stepner M., Abraham S., Lin S., Scuderi B., Turner N., Bergeron A. and Cutler D., *J. Am. Med. Assoc.*, **315** (2016) 1750.
[8] Frederiksen H., *Econ. Develop. Cult. Change*, **14** (1966) 316.
[9] Frederiksen H., *Pub. Health Rep.*, **81** (1966) 715.
[10] Preston S. H., *Int. J. Epidemiol.*, **36** (2007) 484.
[11] Richmond P. and Roehner B. M., *Physica A*, **450** (2016) 748.

[2]Whereas it is possible to explore the old-age interval by computing the life expectancy for people having reached 60, 70 or 80 years, it is impossible to explore the infancy age intervals in the same way.