Physica A 502 (2018) 1–13

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Exploration of the strength of family links

Peter Richmond [a], Bertrand M. Roehner [b,*]

[a] School of Physics, Trinity College Dublin, Ireland
[b] Institute for Theoretical and High Energy Physics (LPTHE), University Pierre and Marie Curie, Paris, France

HIGHLIGHTS

- Age-specific disability parallels strength of family ties.
- Husband–wife ties are weakened by a larger age-gap.
- Parent–children ties are strongest under the age of 5.

ARTICLE INFO

Article history:
Received 5 October 2017
Received in revised form 22 December 2017
Available online 26 February 2018

Keywords:
Disability
Death
Family
Children parents
Age

ABSTRACT

Ever since the studies of Louis-Adolphe Bertillon in the late 19th century it has been known that marital status and number of children markedly affect death and suicide rates. This led in 1898 Emile Durkheim to conjecture a connection between social isolation (especially at family level) and suicide. However, further progress was long hampered by the limited statistical data available from death certificates. Recently, it was shown by the present authors that disability data from census records can be used as a reliable substitute for mortality data. This opens a new route to investigations of family ties because census information goes much beyond the limited data reported on death certificates. It is shown that the disability rate of adults decreases when they have more family links. More precisely, the reduction of the parents' disability brought about by the presence of a child reveals that the strength of ties between parents and child is highest in the first year after birth and then weakens steadily as the child ages. It will also be seen that the strength of the bond between husband and wife is highest when they are of same age and decreases fairly steadily when the age gap increases.

1. Introduction

1.1. Influence of family links on death rates

This paper is a step forward in a story that started in the second half of the 19th century. At that time it was realized that in all countries where data were available the age-specific death rate of married people was 2 or 3 times lower than for non-married (i.e. single, widowed or divorced) people. This became known as the Farr–Bertillon law.[1] A high-accuracy

* Corresponding author.
E-mail addresses: peter_richmond@ymail.com (P. Richmond), roehner@lpthe.jussieu.fr (B.M. Roehner).

[1] William Farr (1807–1883) and Louis-Adolphe Bertillon (1821–1883) were epidemiologists regarded as two of the founders of medical statistics. We use the term "law" (rather than "effect" or "rule") to refer to its wide range of validity. Basically, it holds for all periods of time and all countries for which accurate data are available, including non western countries like China.

https://doi.org/10.1016/j.physa.2018.02.106
0378-4371/

study of this effect was conducted by the present authors in 2016 [1,2] which showed in particular that for widowed persons a majority of the deaths occurs within a few months after the death of the spouse.

Subsequently in 1879, Bertillon [3,4] extended his 1872 study to include the influence of the number of children on the suicide rate.[2]

1.2. Durkheim's conjecture

From this work emerged the conjecture that greater social isolation due to a lack of short-range husband–wife or parent–children interactions produced higher death rates. This became one of the strongest arguments in favor of Durkheim's thesis that the underlying cause of suicide is social isolation. Later on it was shown [1,6] that marital status is a significant factor not only in suicide rates but actually in age-specific death rates *separately* for all major causes of mortality, e.g. heart, pulmonary, cerebrovascular, cancer diseases, accidents.

Incidentally, this observation provides an answer to an objection which is sometimes raised that "it is difficult to distinguish the beneficial impact of marital status on health from the confounding effect of selection into the married state" [7]. In other words, this argument means that a married person does not live longer because of the beneficial effect of being married, but that the person got married because of some inherent initial characteristics such as better health and more stable psychological condition. In this conception marriage is seen as a kind of selective filter which rejects fragile characters. However, how can a marriage-based selection affect the death rate due (for instance) to cerebrovascular attacks, a disease which occur several decades later? The conclusion that marriage-based selection plays a very limited role is also confirmed by studies involving twins [8] and by a study of the suicide rate among Chinese Americans between 1930 and 1950, a time when there was a huge (but decreasing) male–female imbalance in this population [9, chapter 9].

One can also observe that the conception according to which "marriage has a positive effect on health by altering preferences for risky behavior" [10, p. 50] may be correct in some cases (e.g. diseases due to smoking or death by accidents) but has only a limited validity for it cannot account for mortality by diseases whose causative factors are not well known, e.g. cerebrovascular attacks or many forms of cancer.

1.3. Biodemographical experiments in animal populations

In order to test Durkheim's thesis in a broader and more systematic way, it is of course tempting to make observations on animal populations. This idea was tried in an experiment done with ants and fruit flies at the "South China Agricultural University" It lasted over two years from 2012 to 2014 [11]. Ants were extracted from the rest of their population in two ways.

- Single individuals
- Groups of 10 individuals.

In both cases the extracted insects were given appropriate and identical living conditions in terms of temperature, hygrometry, light and food. Two observations were of particular significance.

- Firstly, the two or three days immediately after extraction were marked by a surge in death rates.
- Secondly, in the course of time the average death rates were higher among the "singles" than in the groups of 10. This was observed for the ants as well as for the fruit flies.

1.4. Need of more detailed statistical sources

However, for a better understanding of the connection between social isolation and death rates there was a major obstacle in the sense that statistical information about deceased people was limited to the data recorded on death certificates, i.e. age and marital status of the deceased and cause of death (immediate as well as underlying cause) as determined by a physician.

Many other data would be of great interest from the perspective of social isolation, particularly data about the children of deceased persons (e.g. number and age) for, apart from the husband–wife interaction, the parent–child interaction is certainly the most important social link. Because of this lack of data the conjecture of a connection between mortality and social isolation could not be tested further.

1.5. Work disability as a substitute for mortality data

Progress became possible when it was found that disability data from census records could be used as a reliable substitute for mortality rates [12].

---

[2] Bertillon's results are summarized in [1, p. 751]. Table 1 of this paper tells us that in France the suicide rate of married persons with children is nearly one half of the rate of married persons without child. In his book of 1897 Durkheim [5] gives additional data for families according to their number of children which show that (at same age) the more children, the lower the suicide rate.

More specifically, the following question was asked in the US censuses of 1980 and 1990: [3]

> "Does this person have a physical, mental, or other health condition that has lasted for 6 or more months and which prevents this person from working at a job?"

Several tests [12] showed that the proportion of persons with complete work disability is strongly correlated with the annual mortality rate. One can mention the following facts.

- The fraction with disability increases exponentially with age (see Figs. 1, 2, 3) in a way which parallels the Gompertz law which describes the increase of the death rate.
- The doubling time of the exponential increase is about 10 years that is to say almost the same as for the death rate.
- At first sight it may seem surprising that disability rates which are bounded by 1 should follow the same increase rule as death rates which are unbounded. This is explained by the fact that the upper bound of the disability rate plays no real role because the rate remains low even in old age (at age 85 it is of the order of 0.5).

The paper proceeds as follows.

- First, we briefly explain the procedure (more detailed explanations can be found in Appendix A) and we give an overview in the form of three graphs which illustrate two possible presentations of the results. With respect to married persons, a point of particular importance is the fact (shown in Fig. 1) that it is not the legal status of marriage which matters but the interdependence aspect, that is to say the fact that husband and wife really live together.
- Then we investigate in more detail the effect of the age of the child on the disability of the parents.
- Finally, we investigate the effect on disability of an age gap between husband and wife.

2. Overview of age-specific disabilities

2.1. Procedure

The procedure is fairly straightforward. Nevertheless, for the benefit of readers who would like to repeat the calculations or do similar investigations, the practical details of the procedure are explained in Appendix A.

2.2. Introduction of two metrics for the analysis of disability data

The key-variable is the work disability. We limit ourselves to two cases: no disability versus disability which prevents the person from working; in other words, we discard intermediate cases in which working is made difficult but is still possible.

The two metrics that we shall use parallels the metrics commonly used for the analysis of age-specific death rates.

- For any group, A, defined by a specific marital situation and a given age group the *disability likelihood* is obtained by dividing the number of persons in A who have a disability by the total population of A. We denote this probability by $P1$. It was shown elsewhere [12] that $P1$ is closely related to the death rate.[4]
- Similarly to the death rate ratios and for the same reason, namely to get rid of the exponential growth along with age, one can define disability ratios. A disability ratio represents the disability of persons in a given situation to the disability of the persons of a reference group (for instance married persons).

The following subsections review these variables for several situations of interest.

2.3. Social status versus interaction between spouses

With respect to marital situations a question comes to mind immediately which is of both practical and conceptual interest: for married persons is it the *social status* which matters (in terms of disability) or the actual *interaction* between spouses? What we mean is the following. There is a pressure to get married exercised on young adults by families as well as by society (more so in former times than nowadays) which can lead to a substantial amount of stress. Getting married is a way to remove that pressure and feel better. In this status explanation the interaction between spouses is either downplayed or ignored.

Fig. 1 allows a clearer insight into this question. Among the groups considered, there are two, namely "married, spouse absent" and "separated" which provide the status of being married without however any real presence of a spouse. It can be seen that the disability likelihood of these groups is much closer to the case of divorced persons than to the case of married persons with spouse present. So, this graph strongly suggests that the actual interaction between spouses matters much more than the social status alone.

---

[3] Similar questions were asked in other years but which did not have the same accuracy. For instance, in all "American Community Surveys" (ACS) from 2000 to 2007 the following question was asked:

> "Because of a physical, mental, or emotional condition lasting 6 months or more, does this person have any difficulty in working at a job or business?".

Clearly there is more subjectivity in assessing a difficulty than an impossibility.

[4] We use the term "rate" because it is commonly used but it should be recalled that in fact the death rate represents a probability, namely the likelihood of dying in the age interval under consideration.

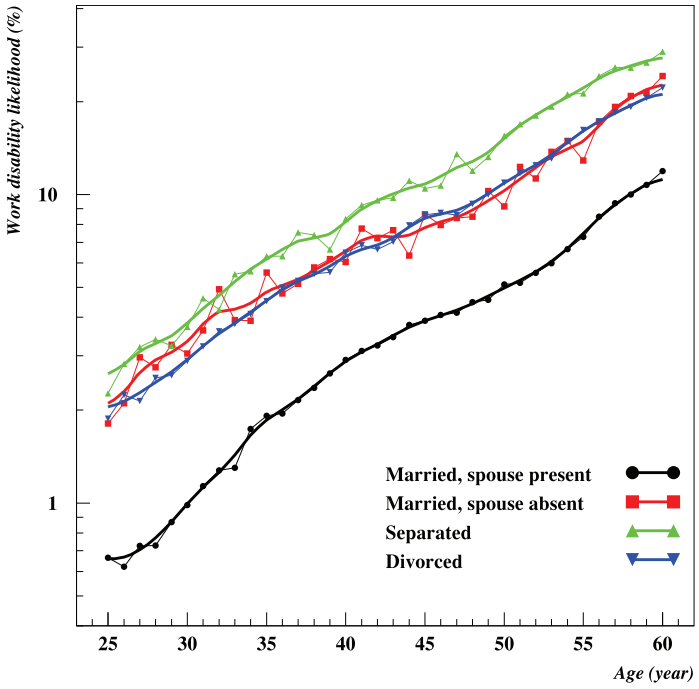

Fig. 1. Likelihood of disability in various marital situations (for males and females together in 1990). In view of their disability levels the four curves fall into two classes: "married with spouse present" on the one had and on the other hand there are the three curves "married with spouse absent" + "separated" + "divorced". In other words, it is the real presence of the spouse which matters, rather than the social status of being married.
Source: 5% sample of the US census of 1990 downloaded from the IPUMS database [16].

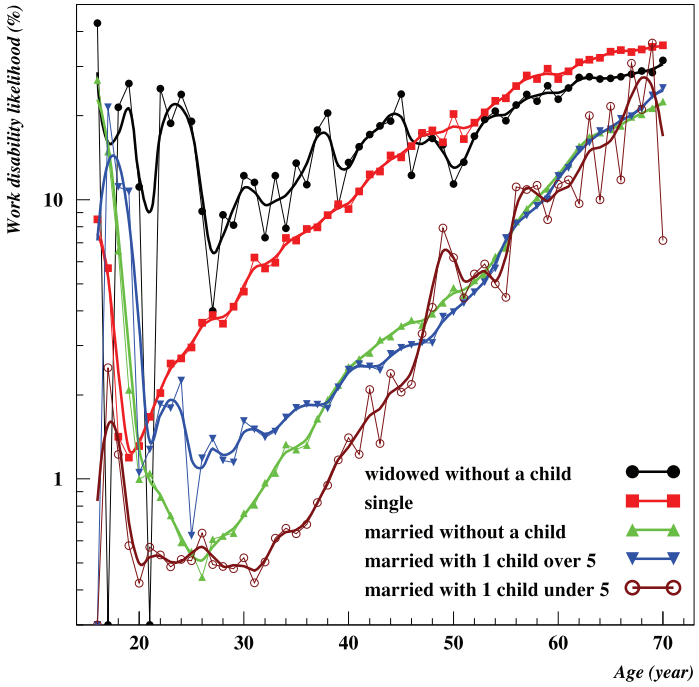

Fig. 2. Likelihood of disability in various family configurations (for males in 1990). The amplitude of the fluctuations (as shown by the yearly data points) is mostly determined by the number of persons in each age group. The well-known "young widower effect" has an amplitude of more than 10. While for some reason not yet well understood this is specific to males, the other results display the same pattern for men and for women. Moreover the fact that one gets a very similar graph for the census data of 1980 shows that the observed effects are fairly stable.
Source: 5% sample of the US census of 1980 and 1990; downloaded from the IPUMS database [16].

Incidentally, one may find surprising that the curve of "separated" is slightly above "divorced" for the reason that "divorced" seems a more serious and definitive situation. A possible explanation may present itself when one refers to what "separated" really means in the census. It applies (i) to persons who have parted by themselves because of marital discord

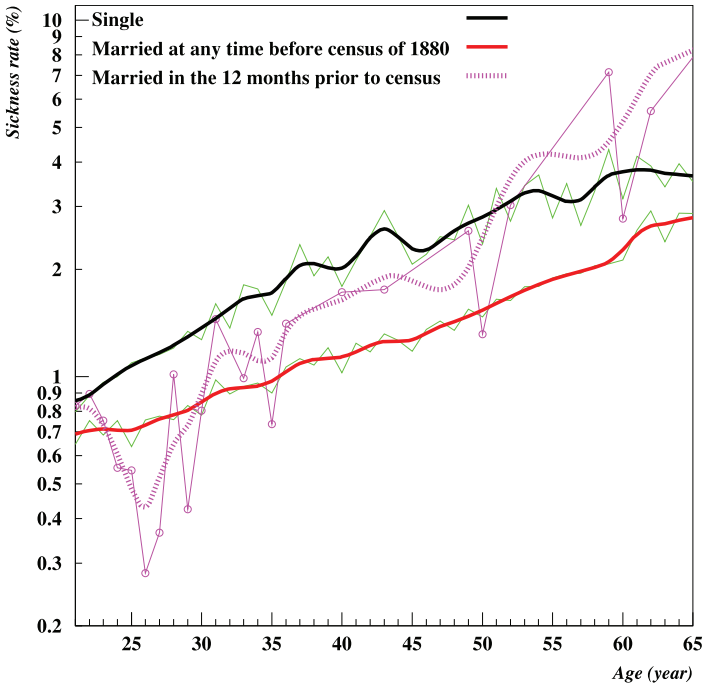

Fig. 3. Sickness rate in the transition from single to married state. The sickness rate gives the proportion of sick persons at the moment of the census. The dotted curve corresponds to couples for whom the time between marriage and census is less than 12 months; in other words, it corresponds to newly married couples. The sickness rate of these persons is substantially higher than the rate of couples who have been married for a long time; this is especially clear in the age interval 30–55; the dip between 20 and 30 may be due to a different effect to which we come back later; the fact that there are large fluctuations over the age of 55 does not come as a surprise because in this age group there are very few newly married persons.
Source: 10% sample of the US census of 1880; downloaded from the IPUMS database [16].

(ii) to couples who have already a legal separation prior to an upcoming divorce. We see that both cases refer to the early stage of a separation. It is known (see [2][5]) that for events such as divorce or widowhood there is a death rate spike which lasts about 6 months and surges above the stationary death rate of the initial (married) situation. Obviously, the weight of such spikes will be more diluted in longer time intervals, that is to say for the "divorced" category.

2.4. Age-specific disability in various family situations

Fig. 2 shows curves of disability likelihood for various family types. Three observations can be made.

- All curves display an upward exponential trend (albeit with different doubling times) which is in line with Gompertz's law which says that death rates increase exponentially with age.
- Although the curves display sizable differences particularly in mid-age, they all converge toward a same limit in old age. Here we restricted the analysis to the age interval (16, 70). In subsequent graphs it will be seen that this convergence continues in the age interval (70, 90). It has a simple interpretation; it means that in old age purely biological factors of aging become predominant. As a result, the differences due to marital circumstances become imperceptible.
- Finally, an important observation is the fact that for any age-group the disability likelihood decreases as the number of family bonds increases. From widowers or unmarried persons to married persons with a child under 5 years the likelihood is divided by factors which range between 5 and 20.

An additional remark is in order. At first sight it may seem strange that between the age of 20 and 40 the disability rate of couples with a child over 5 is higher than the disability rate of married couples without children. There can be several explanations.

- For instance, it may be that in this case for some yet unknown reason, the connection between "more disability" on one hand and "less interaction" on the other hand may not hold.
- It may also be that the interaction is more complicated than just the sum of the number of links; in short, our common sense understanding may be inappropriate. Thus, this piece of evidence raises a stimulating question which may allow further progress.[6]

[5] A graph (Fig. 2) in [13] gives the same order of magnitude of about 6 months for the death rate spike following widowhood.
[6] To rely on observation rather than on anthropomorphic reasoning is the approach strongly recommended by Durkheim [14] in his landmark book.

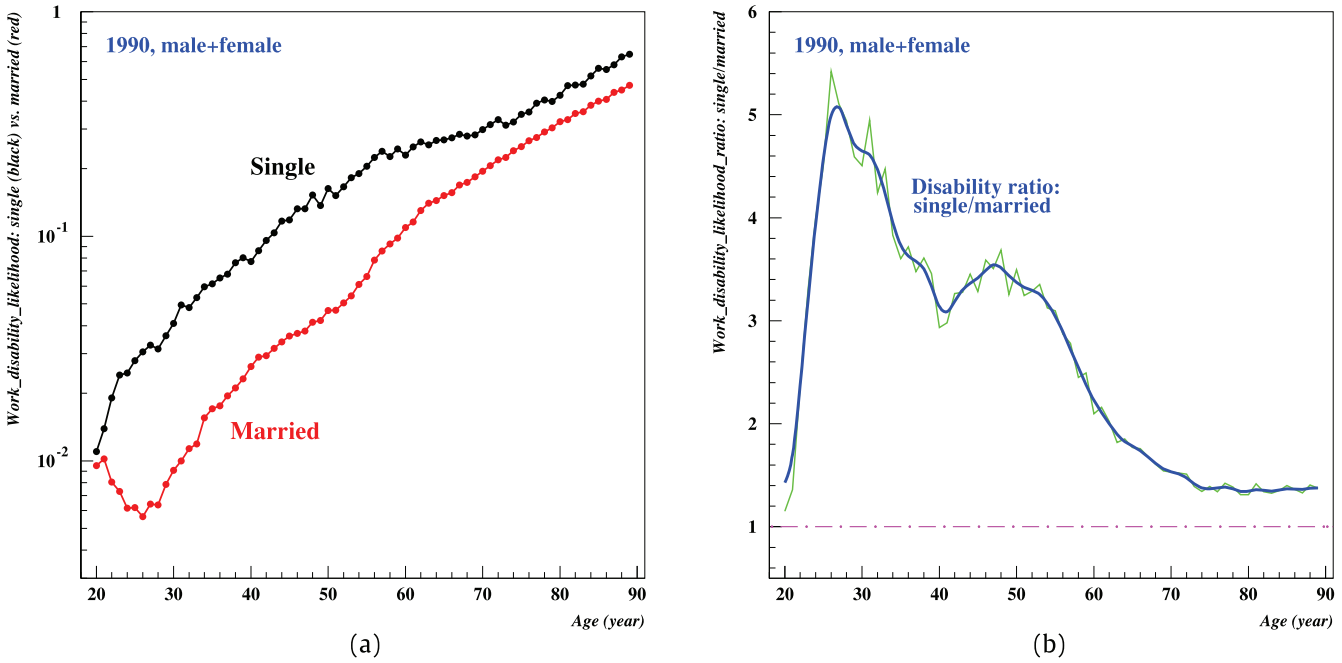

Fig. 4a,b. Disability for single and married persons (in both cases there is nobody else in the household). (a) Disability likelihood for single and married persons. In the (20, 32) age interval, it is remarkable that the benefit of being married is strong enough to overcome the upward trend of disability with age. It is true that disability may hinder marriage but this effect is of significance mostly for elderly people. The bulk of marriages occurs between 20 and 35 that is to say at an age when the disability rate is very small. Therefore, as already explained in the introduction, the selection process based on disability should affect the set of married persons in a limited way. This selection argument is also contradicted by Fig. 3 which shows that even for persons who get married in old age their disability rate is actually (and surprisingly) higher than the rate of single persons. (b) Disability ratio R = single/married.
Source: 5% sample of the US census of 1990 downloaded from the IPUMS database [16].

2.5. Time lag of the transition from "single" to "married"

The census of 1880 of which IPUMS gives a 10% sample offers a rare opportunity to explore the transition from the state of being single to the state of being married. This census tells us whether or not the respondents got married in the 12 months which preceded the taking of the census. Whilst in a legal sense getting married is instantaneous, the physical and mental state does not change immediately but with a time delay that we denote by θ. Is θ of the order of a few weeks, a few months or a few years?

If one assumes a uniform distribution of marriage dates, the dotted curve of Fig. 3 corresponds to an average time lag of 6 months between marriage and census time. For ages between 30 and 55 this curve of newly married persons is mid-way between the curves for "single" and "married". This shows that θ is of the order of a few months. The dip the 20–30 age interval is probably related to the one observed in Fig. 4a.

2.6. Disability ratio

The likelihood curves in Fig. 2 give an overall picture but they are not well suited for accurate measurements. In studying the effect of marital status on death rates it is appropriate to define a reference case C_r (for instance the situation of being married) and then, for any other marital situation, to compute the ratio with respect to C_r. This is the procedure followed by the present authors in [1]. It is illustrated in Fig. 4a, b. By discarding the exponential growth trend which is common to all cases, this procedure allows higher accuracy.

One drawback, however, is that with this representation one is limited to pairs of variables. In order to cover the 5 curves of Fig. 2 one would need to consider 10 pairs. Another drawback is the fact that when the ratio displays a peak one does not know if it is due to a peak in the numerator or a trough in the denominator.

Notwithstanding these limitations, the disability ratio will provide accurate assessments of two important features:

- Is the ratio larger or lower than one?
- What is the maximum amplitude of the ratio?

In Fig. 4a, b the procedure is illustrated for the pair (single, married). In Fig. 4b we see that R is above 1 for all ages. In other words, with respect to non-married persons, the marital bond provides a protection against disabilities. This effect is quite considerable in the age range between 25 and 35.

3. Connection between disability and social isolation

3.1. Why a disability ratio larger than one implies some form of interaction

In order to get an intuitive perception it is useful to consider two simplified cases. They are simplified in the sense that we assume that the disability ratio R is the same for all ages.

• First, suppose that $R = 1$. This means that from the perspective of personal health, it does not make any difference whether the persons are single or married. Such an outcome is consistent with a complete lack of interaction between spouses. In this case each spouse would be just like a non-married person.

• Now, suppose that $R = 5$. Unless, marriage has pre-selected the spouses in some way[7] the state of being married must have a beneficial effect on the spouses. We have shown above that this is not merely a "status effect" but is conditioned by the actual presence of the spouse.

Clearly the interaction between spouses can take many forms, some of which may have no health effects. For instance, spouses may cooperate in running a business together; such a cooperation certainly implies a form of interaction. That specific interaction would have an effect on *economic* variables. Thus, if the average income (i_2) of couples is higher than twice the income (i_1) of single persons involved in the same activity, then the excess-income $e = i_2 - 2i_1$ may be used as a measure of the strength and usefulness of such a business interaction.

Economic cooperation, as considered above, is easier to conceive intuitively than an interaction resulting in health effects. At this point one can only say that there *must* be some form of interaction that is beneficial to both individuals.

3.2. Connection between social isolation and high suicide rates

The connection between high suicide rates and social isolation has been documented in many situations ([9], section entitled "Effect of social isolation on suicide" p. 209–220). Let us recall some of them.

(1) The disruption of former links experienced by immigrants when they leave their country of origin for a new place results in a multiplication of suicide rates by a factor of 1.4.

(2) The suicide rate of prison inmates is between 2 and 5 times higher (depending on the specific type of prison) than the rate for the general population of same age and gender.

(3) The suicide rates of prisoners in solitary confinement, in remand centers or in lockup (i.e. immediately after arrest) are at least three times higher than that of the general prison population.

(4) According to the "Schizophrenia Society of Ontario", the suicide rate of schizophrenics is about 10 times higher than that of the general population.

Most of the cases mentioned above concern young or middle-aged persons which means that their death rates for other causes than suicide are rather low. In other words, for such people it is not only the suicide rate which is inflated but also the "all causes" death rate.

What is the mathematical relationship between disability and bond strength?

As they vary in opposite direction one can posit a relationship between bond strength and the inverse of the disability. For the moment as we limit ourselves to *relative* estimates of bond strength, we cannot state (and in fact do not need to know) what is the precise mathematical form of the relationship.[8]

4. Impact of the presence of children according to their age

4.1. Presence of a child of any age

Fig. 5a shows that for almost all ages until 60, the disability is higher for couples without children than for couples with one child. It is true that there is a small interval between the ages of 20 and 28 where the disability of mothers with a child is slightly higher than the disability of childless mothers. This may be due to child-bearing health problems. It is not uncommon that work must be discontinued in order to prevent a miscarriage or a premature birth.

4.2. Comparison with the observations made by Modig et al. [15]

In an interesting paper [15] give estimates for the health benefits of having children for persons aged 60 and over. Their study is based on the computerized "Register of the Total Population" (TPR) that exists in Scandinavian countries.[9] As summarized in graphical form in our conclusion their results are quite compatible with ours. It would be interesting to be able to compare similar results for ages under 60 for which the disability and death ratios are higher and therefore easier to compare. We hope this will be possible in the future.

A typical result of their study is the following.

---

[7] An explanation already discussed above and which is furthermore contradicted by the fact that the ratio shows strong variations with age.

[8] The relationship could be something of the form: $b \sim 1/d^{\alpha}$, where b is the bond strength, d the disability likelihood and α a positive parameter.

[9] For obvious confidentiality reasons this register is not publicly available.

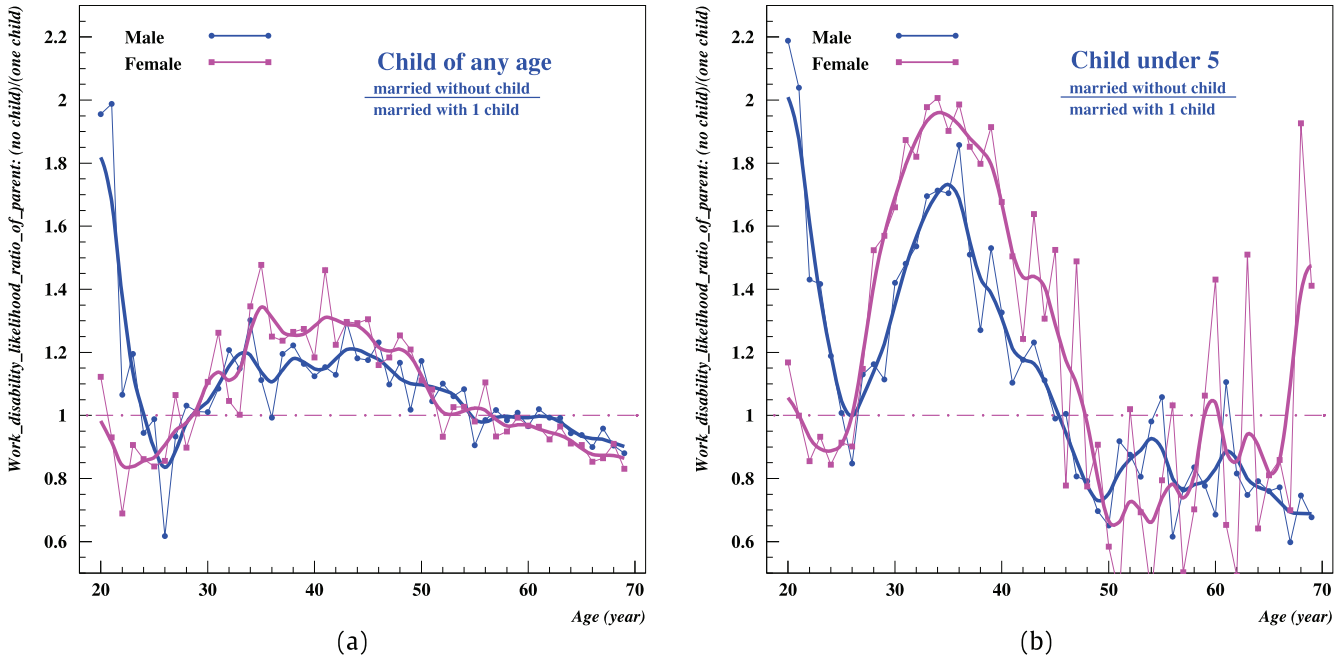

Fig. 5a,b. Changes in the disability ratio due to the presence of a child (1990). The fact that the curves are mostly above 1 shows that the presence of a child reduces the disability. (a) All children are considered whatever their age. (b) Only children less than 5 year old are considered.
Source: 5% sample of the US census of 1990 downloaded from the IPUMS database [16].

For men and women at age 60 (whether married or not) the presence of (at least) one child resulted in a life expectancy higher by 7.4% than for childless men and women.[10]

4.3. Presence of a child under 5

Fig. 5b analyzes the same situation but limited to children under five. The most obvious observation is that the peak values are notably higher for both fathers and mothers, something which seems to agree with our intuitive perception.

Needless to say, children under 5 are very uncommon for women over 45. They may of course have adopted or step children but such cases are rare. Consequently this age interval is marked by large fluctuations.

The fact that the "under age 5" condition has a significant impact encourages us to consider more closely the effect of the age of the child. It turns out that the IPUMS database has two variables that allow us to know the child's age when there is only one. Denoted YNGCH and ELDCH, these IPUMS variables give the ages of the youngest and eldest child. When there is only one child, YNGCH and ELDCH are equal and give the relevant age.

In Fig. 6 the reference is the case of children of less than one year of age (referred to as age 0). For the three parent age groups shown, the disability ratio increases with child's age at least until the age of 10.

5. Impact of an age gap between husband and wife

In Fig. 7 we investigate the influence of an age gap between husband and wife. For that purpose the first requirement is to be able to identify the pairs of record lines corresponding to married couples. This can be done thanks to the variable SERIAL which is the identification number of each household. The whole procedure can be summarized as follows.

First, we select the records for which the variable NUMPREC (number of persons in the household) is equal to two. Most, but not all of them, are of course married couples. The latter can be extracted thanks to the marital status variable. In this way, one gets all married couples who are without children. The children must of course be excluded for otherwise they would affect the disability of the parents. Finally, we select the couples with the SERIAL variable and measure their disabilities.

Fig. 7 suggests that a non-zero age gap increases the disability ratio particularly for young partners. Under our conjecture of an inverse connection between disability and interaction, it means a reduction in interaction. More precisely, one should say that there is a reduction in the health-linked interaction for, as already observed, there can be other forms of interaction that do not translate into health benefits.

[10] This percentage was obtained by dividing the life expectancy benefit given in the paper by the life expectancy at 60 in Sweden, namely 23.6 years in 2010.

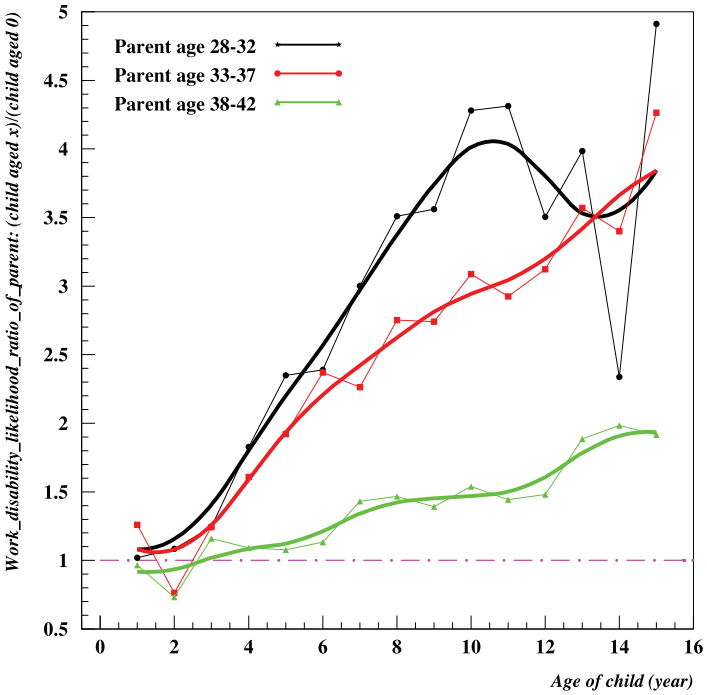

Fig. 6. Disability likelihood of parents according to the age x of their child; the likelihood was normalized to make it equal to 1 for $x = 0$. The fact that all curves are above 1 and going up shows that the disability of the parents is smallest with young babies and that it increases with child age.
Source: 5% sample of the US census of 1990 downloaded from the IPUMS database [16].

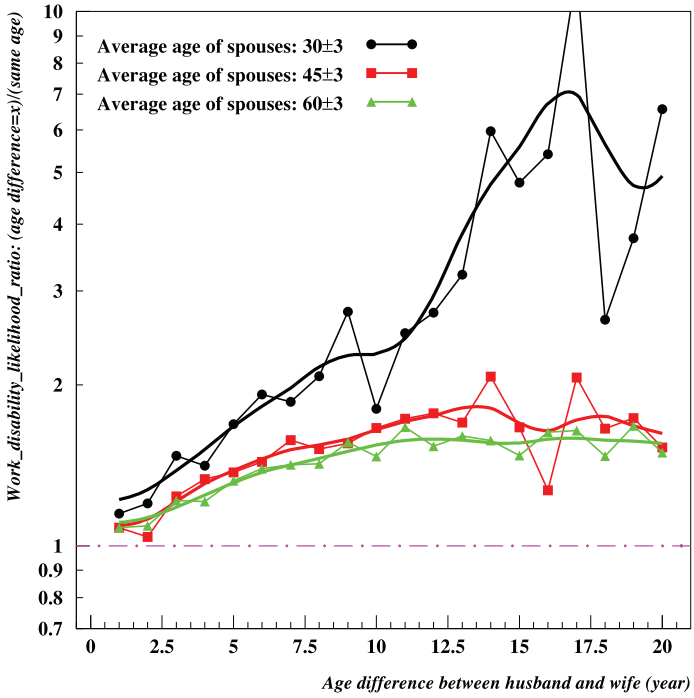

Fig. 7. Disability likelihood of spouses as a function of their age difference x; the likelihood was normalized to make it equal to 1 for $x = 0$. The fact that all curves are above 1 and are upgoing shows that the disability is smallest for $x = 0$ and increases with the age gap.
Source: 5% sample of the US census of 1990 downloaded from the IPUMS database [16].

6. Conclusion

There can be two readings for the present paper. The first and most direct is in terms of disabilities and health. The second one which is in terms of bond strength is based on what we called Durkheim's conjecture. Here we will summarize our results separately for the two interpretations.

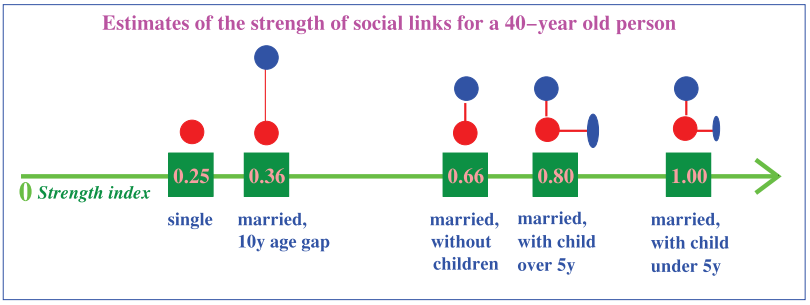

Fig. 8a. Summary of bond strengths b between 40-year old persons and their family members in various situations. The estimates for b were derived from the disability likelihood d read in Figs. 1b, 2, 4 and 6, through the relationship $b \sim 1/d$. The situation "married with a child under 5" was taken as the reference level and assigned the value 1.00. The strength numbers are as given by our estimates. However, it is clear that as they should be seen as rough indications one digit would be enough. We kept two digits to make contact with the data of Modig et al. [15].
Source: Figs. 1b, 2, 4 and 6.

6.1. Health effects seen in an evolutionary perspective

Can the outcomes observed in the paper be viewed in a unified manner? In other words can we divine a thread which connects them? The following characteristics seem to derive from what may be called an evolutionary principle.

- Married people have better health than unmarried and this effect is at a maximum for young adults between 20 and 30 (Fig. 4b),
- In addition such young couples have better health when husband and wife are of same age (Fig. 7) and their health is even better when they have already a young child (Fig. 5b).

Clearly these successive characteristics all favor people who are in optimum condition to have children. In other words, they may be viewed as important contributors to the sustainability and optimization of the human reproduction process.

It is true that when one compares different species it does not always make sense to suppose reproduction optimization. For instance, some birds will have to fly 10,000 km to their nesting area, a feat that will cause many losses, while a closely related species may be able to reproduce without migrating at all. However, within one and the same species there must be a form of optimization at work in order to keep the species viable. This should be particularly true for species whose reproduction margin is fairly narrow.

6.2. Basic questions about bond strength

The results presented in this paper answer some basic questions which we have had in mind for many years.

- How does the strength of the interaction between parent and children compare with the interaction between husband and wife? We have seen that the husband–wife link is stronger than the parent–child link except when the child is under 5 in which case the links are of similar strength.
- How does the parent–child interaction change as the child becomes a teenager? We have seen that the link becomes weaker.
- How does the strength of the marital bond depend upon the partners' age gap? We have seen that the bond becomes weaker.

Needless to say, on all such questions everybody may have an opinion based on personal experience. The key-point was how to address these questions in a *scientific way*. The answers given in this paper were based on Durkheim's famous conjecture of a connection between social isolation and suicide rates or more generally death rates. By using the quasi-equivalence established previously between death rates and work disability frequencies we can find relative estimates for bond strengths.

The main objective of Figs. 8a, 8b is to make the results obtained previously more readable. One difficulty in the interpretation of the graphs is the fact that the disability level does not measure the strength of the interaction but its weakness. The higher the disability rate the lower the strength of the interaction. We can build an index which truly reflects strength by replacing the disability rate by its inverse. This leads to the diagram of Fig. 8a. Ever since Durkheim it is well known that family links represent the main component of the social links of a person. In Fig. 8a we can see that for a person who is single its remaining social links represent 25% of the strength ascribed to the situation with the strongest links. The estimates given in Fig. 8b correspond to the results of an independent study. The two sets of estimates appear to be consistent with one another.

6.3. Other perspectives

Before closing this paper, it must be mentioned that Durkheim's perspective based on interaction strength is not the only one that has been considered. As a matter of fact, nowadays the dominant perspective seems to be one based on *individual*

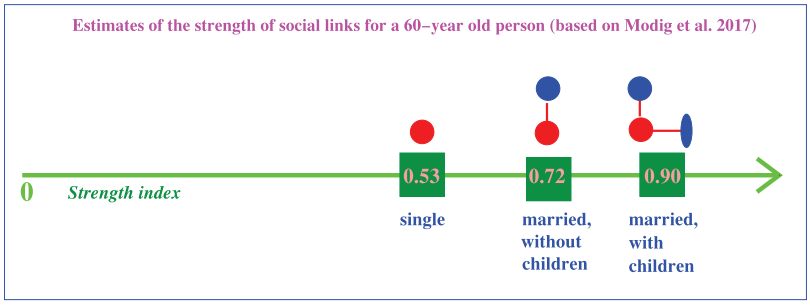

Fig. 8b. Summary of bond strengths b between 60-year old persons and their family members in various situations. The estimates for b were derived from the death rates computed in the study by Modig et al. [15] in the same way as in Fig. 8a. The situation "married with children under 5" was taken as the reference level and the value 0.9 was attributed by comparison with Fig. 6.
Source: Modig et al. [15] and complementary data from the same study (personal communication).

psychological characteristics. As typical of this perspective one can mention a paper by Walter Gove [6]. In a sense this conception marks a return to the anthropomorphic conception prevailing before Durkheim. Its main drawback is that it is difficult to test it in any meaningful way for it incorporates a variety of *ad hoc* psychological explanations.

Let us illustrate this point by an example from [6]. Observing in his Table 8 that for death by tuberculosis widowed males in the age-group 25–34 (and in the time interval 1959–1961) have a death rate 11.5 higher than the rate of married males of same age, the author suggests (p. 47) that being "more aggressive and willing to take risk", men would be "less apt to enter treatment and persist in the prolonged and careful treatment that is necessary for tuberculosis". From a scientific perspective recourse to specific arguments (in this case special features of tuberculosis treatment) for different causes of death is hardly satisfactory in so far as the high death rate ratio for young males is by no means limited to tuberculosis. Only *testable predictions* would be convincing.

6.4. Prospects

What should be done next?

In physics there are several ways to measure inter-molecular interactions. One is to measure the amount of energy necessary to break the molecular bonds. For a liquid, simply heating it to its boiling point does this. Another approach is to estimate the parameters of the Lennard-Jones potential on the one hand and the average inter-molecular distance on the other hand. It is the fact that all such methods lead to consistent results that makes the methodology truly convincing.

Ideally, we should follow the same path to uncover the nature and strength of social interactions. We have earlier suggested that other forms of interaction such as business relations might also be explored in a similar quantitative manner. A study of the correlation between the duration of bonds and their strength may also be of significance; it can be illustrated by the difference between family bonds and temporary links with colleagues or friends. Exploring these and other forms of interaction should be enlightening.

Appendix A. Data selection and analysis

As an example, we wish to compare non-married persons who live alone to married couples without children.

A.1. Data selection and extraction from the IPUMS data base

On the IPUMS database [16] there are two kinds of variables (i) those which refer to the household and (ii) those which refer to individuals. Here we will use one household variable, namely the number of persons who compose the household (noted NUMPREC in the IPUMS coding system) and the following 5 variables pertaining to individuals.

- The number of children of each individual (NCHILD)
- The sex and age of each individual (SEX, AGE)
- The marital status of each individual: noted MARST it offers 6 options but here we will use only 2, namely 1 = married (spouse present), 6 = never married.
- A variable (noted DISABWRK) which describes the disability status of the person. IPUMS offers 4 options but here we will use only two, namely: 1 = No disability affecting work, 3 = Disability prevents work.

Once these 6 variables have been selected, specific limits will be chosen within IPUMS for 3 of them.

- Only persons without children (NCHILD = 0) will be retained. This eliminates all couples with children and markedly reduces the size of the file.
- As the disability variable is defined only over the age-range (16,89) we will limit our age interval to (20,89).
- Moreover, as we consider only two marital situations. we eliminate all records for which MARST is not equal to 1 or 6.

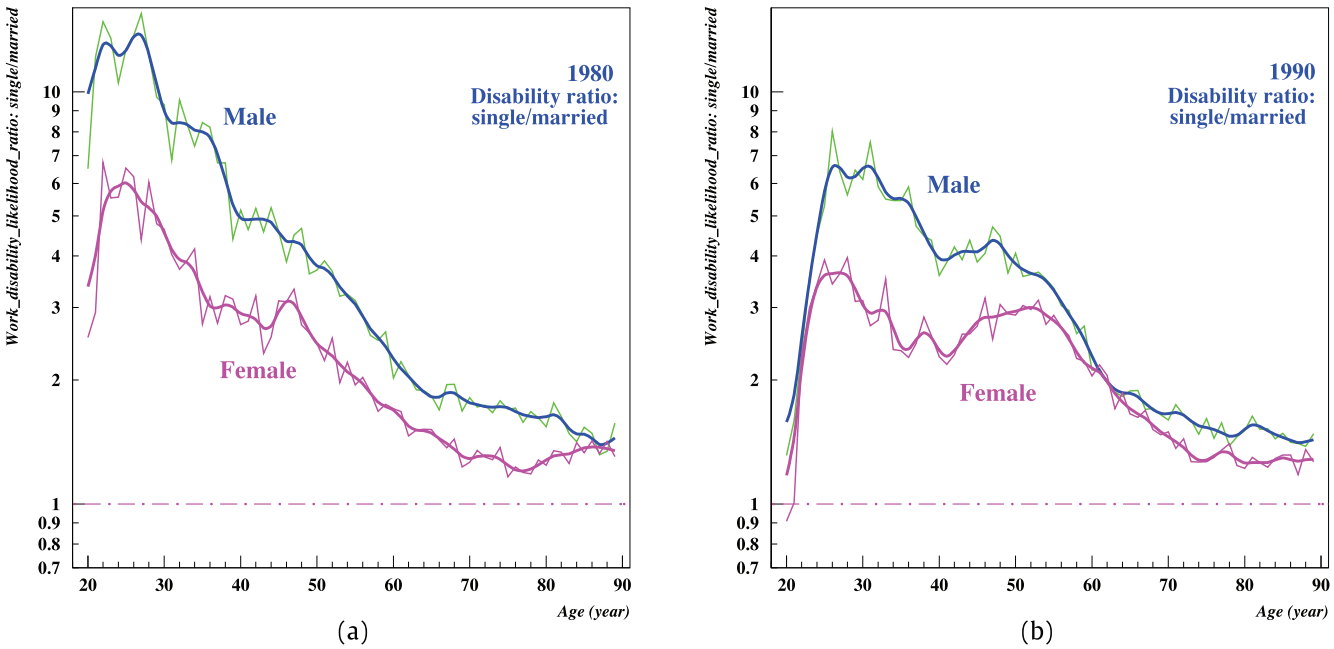

Fig. B.1a,b. Variability assessment: single/married disability ratio in 1980 and 1990. Here "single" means that the persons are not living with other persons in the same household; "married" means without children and without anyone else in the same household. The fact that all curves are above 1 is consistent with the result already known from death rate data. (a) The 1980 sample consisting of single persons and married couples without children comprises 3,150,056 individuals. (b) In 1990 it comprises 3,695,444 persons.
Source: US censuses of 1980 and 1990 from the IPUMS data base [16].

Finally we select the 5% sample of the census of 1990; remember that the work disability variable exists only for the censuses of 1980 and 1990.

This selection procedure yields a file of 3,150,056 lines (31 M once decompressed) where each line corresponds to one person.

By way of illustration, the first data line reads:

→ 050102361.

If one introduces spaces between the 6 successive variables, one gets a more readable data line:

→ 05 0 1 023 6 1.

These variables have the following meaning:

→ 05 persons in household, 0 child, 1 = male, 023 = age, 6 = single, 1 = no disability.

One should make a distinction between 05, 0, 023 which are "real" data and 1, 6, 1 which are code numbers that take significance only through the conversion tables given in IPUMS.

A.2. Analysis of the data

The file is read line by line and all non-married persons who live *alone* (i.e. NUMPREC = 1 and MARST = 6) are collected. Let HZ denote the number of all persons who correspond to this specification Among them, a number HZd have a disability (i.e. DISABWRK = 3). Thus, non-married persons will be disabled with a probability $PZ = HZd/HZ$.

In the same way, we collect all couples who live without any other person in the same household by selecting the lines with NUMPREC = 2 and MARST = 1; a fraction $P1 = H1d/H1$ of them will have a disability.

Appendix B. Confidence interval vs. social variability

In this paper we are interested in stable *structural* properties. A relationship between marital status and disability which would be valid for only one (or a few) years would be of little interest. In other words, beyond the purely statistical notion of confidence intervals there is the question of the variability of the phenomenon under consideration. In our perspective the later is of greater significance than the former. Below these two aspects are discussed successively.

B.1. Error bars

In [12] error bars (or rather error bands) showing confidence intervals for every age were drawn. However, such error bands can impair the readability of the graphs. Here we have no need for accurate error bars for we are mostly interested in

the overall shape of the disability curves. That is why we will use an alternative representation taking advantage of the fact that the magnitude of the confidence intervals is well represented by the amplitude of yearly fluctuations. These fluctuations around the 5-year centered moving average define the uncertainty margin almost quite as well as the error bands used previously.

B.2. Variability

With respect to error bars, the comparison between Fig. B.1a and b tells us something else of interest. It shows that from one decennial census to the next, the disability ratio changes far more than what the confidence interval within one and the same sample would suggest. This is of course hardly surprising for in an interval of 10 years, many conditions may change which may to some extent affect the disability ratio. This shows that confidence intervals within a given census are in fact of limited interest for the purpose of assessing the variability of the phenomenon under consideration.

References

[1] P. Richmond, B.M. Roehner, Effect of marital status on death rates. Part 1: High accuracy exploration of the Farr-Bertillon effect, Physica A 450 (2016) 748–767.
[2] P. Richmond, B.M. Roehner, Effect of marital status on death rates. Part 2: Transient mortality spikes, Physica A 450 (2016) 768–784.
[3] L.-A. Bertillon, Article "mariage", in: Dictionnaire Encyclopédique des Sciences Médicales, in: Encyclopedic Dictionary of the Medical Sciences, vol. 5, 1872, p. 752 2nd series, [Available on "Gallica", the website of digitized publications of the French National Library].
[4] L.-A. Bertillon, Article "France", in: Dictionnaire Encyclopédique des Sciences Médicales, in: Encyclopedic Dictionary of the Medical Sciences, vol. 5, 1879, pp. 403–584 4th series, [Available on "Gallica", the website of digitized publications of the French National Library.].
[5] E. Durkheim, Le suicide. Etude de sociologie. F. Alcan, Paris. [A recent English translation is: "On Suicide" (2006), Penguin Books, London], 1897.
[6] W.R. Gove, Sex, marital status and mortality, Am. J. Sociol. 79 (1) (1973) 45–67 [This study relies on the data given in "US Public Health Service" 1970 cited below.].
[7] K. Christenson, J.W. Vaupel, Genetic factors in adult mortality, in: R.G. Rogers, E.M. Crimmins (Eds.), International Handbook of Adult Mortality, Springer, Heidelberg, 2011, pp. 399–410.
[8] M. Osler, M. McGue, R. Lund, K. Christensen, Marital status and twins' health and behavior: an analysis of middle-aged Danish twins, Psychosom. Med. 70 (4) (2008) 482–487.
[9] B.M. Roehner, Driving forces in physical, biological and socio-economic phenomena, in: A Network Science Investigation of Social Bonds and Interactions, Cambridge University Press, Cambridge, 2007.
[10] R.E. Santerre, S.P. Neun, Health Economics, sixth ed., in: Theory, Insight, and Industry Studies, South Western Cengage Learning, 2013.
[11] L. Wang, Y. Xu, Z. Di, B.M. Roehner, Effect of isolation on life expectancy of red imported fire ants *Solenopsis invicta* and tephridid fruit fly *Bactrocera dorsalis*, Acta Ecol. Sin. 36 (2016) 252–255.
[12] P. Richmond, B.M. Roehner, Impact of marital status on health, Physica A 486 (2017) 542.–553 Preprint available on the arXiv website (arXiv: 1704.00752, 30 March 2017).
[13] F. Elwert, N.A. Christakis, Widowhood and race, Amer. Sociol. Rev. 71 (2006) 16–41.
[14] E. Durkheim, Les règles de la méthode sociologique. Payot, Paris. English translation: Rules of sociological method. Free Press, New York, 1894.
[15] K. Modig, M. Talbäck, J. Torssander, A. Ahlbom, Payback time? Influence of having children on mortality in old age, J. Epidemiol. Community Health 71 (2017) 424–430.
[16] S. Ruggles, K. Genadek, R. Goeken, J. Grover, M. Sobek, Integrated Public Use Microdata Series (IPUMS), University of Minnesota, Minneapolis (Minnesota), 2017.