Soc Indic Res (2014) 119:705–721
DOI 10.1007/s11205-013-0502-9

Resolution of the Happiness–Income Paradox

Y. Zee Ma · Ye Zhang

Accepted: 28 October 2013 / Published online: 8 November 2013

Abstract It is widely believed that happiness is strongly correlated with wealth and income, but according to the happiness–income paradox, this is not always true. The paradox predicates that there is a strong positive correlation between income and happiness nationally, but the correlation is essentially absent in international comparisons, or in a long-term longitudinal comparison. This paradox has been widely debated among economists and the controversy has persisted for several decades. In this article, the happiness–income paradox is explained in terms of ecological correlation due to spatial aggregation or data-grouping, change of reference classes, and confounding variables. The controversy is resolved when ecological correlations and third-variable effects are accounted for. At the individual level, happiness and income are correlated positively, but not as strongly as many believe. In international comparisons, happiness and income are, in general, quite strongly correlated as well, contrary to what Easterlin (Nations and households in economic growth: essays in honor of Moses Abramovitz, Academic Press, New York, 1974) claimed and similar to what others have found, but for different reasons. Long-term comparison is also related to ecological correlation, but it is related to the change of reference classes as well.

Keywords Happiness–income paradox · Ecological correlation · Spatial aggregation · Data-grouping · Simpson's paradox · Reference class

1 Introduction

Happiness and income levels are among the most important variables in society and human evolution, as both are directly related to people's basic well-being. Some countries have

Y. Z. Ma (✉)
Schlumberger, 1675 Broadway, Denver, CO 80202, USA
e-mail: yzeema3@yahoo.com

Y. Zhang
University of Wyoming, Laramie, WY, USA
e-mail: yzhang9@uwyo.edu

proposed the concept of Gross National Happiness (GNH) to counterbalance hyper-consumerism and excessive spending (Di Tella and MacCulloch 2008; Stiglitz et al. 2010). The economics of happiness has recently been considered to be an important academic discipline (Bernanke 2010).

It is generally agreed that attaining happiness is a key life goal while income is one of the most important means of attaining happiness. This implies that happiness is strongly correlated with wealth and income; but according to the happiness–income paradox, this is not always true. Easterlin (1974) postulated that there was a strong positive correlation between income and happiness within a country, but the correlation was very weak, or even nonexistent, in international comparisons, or in a long-term longitudinal comparison. This paradox is sometimes called the Easterlin paradox, and it has been widely debated by economists, psychologists, and social scientists (Easterlin 1995; Hagerty and Veenhoven 2003; Clark et al. 2008; Easterlin et al. 2010).

This paradox is, in fact, a result of using various ecological correlations (Robinson 1950; Grotenhuis et al. 2011) that are different from their counterpart at the individual level. Misunderstanding of the paradox is mainly due to the lack of awareness of ecological correlations in the economic community. Secondarily, many factors affect happiness, leading to a weaker correlation between happiness and income. The effect of these third variables is related to a counterintuitive statistical phenomenon known as the Simpson's paradox (Simpson 1951; Bickel et al. 1975). Below, these concepts and other related statistical problems are reviewed to facilitate the analysis of the happiness–income paradox.

1.1 Ecologic Correlation and Modifiable Areal Unit Problem

Correlation is traditionally calculated using observed elements—generally small samples in scientific data analysis and individuals in social data analysis. An ecological correlation is calculated from divisible-groups' means, such as state, region, or other units (Robinson 1950; Grotenhuis et al. 2011). Ecological correlation is common in both scientific and social data analysis because sample data and statistical inference are not always defined on the same unit (Gehlke and Biehl 1934). In spatial statistics, this is termed the change of support problem (Cressie 1996; Gotway and Young 2002). The difference between a correlation at the individual or elementary level and its ecological-correlation counterpart can be very large. Robinson (1950) noticed that many researchers have used an ecological correlation to substitute for the correlation at the individual level. He showed a general non-equivalency between the individual correlation and the related ecological correlation that is calculated with spatial aggregations. One of his examples is the correlation between illiteracy and nativity (both in percentage) as seen in the US census data from the early 1930s. The correlation is negative at the individual level, but positive at the state and division levels. This was one of the earliest real examples of correlation reversal for continuous variables, which is a manifestation of the Simpson's paradox (Blyth 1972).

Although the examples presented in Robinson (1950) deal only with spatial problems and are focused on correlation analysis, later investigators have extended the concept to ecological inferences beyond spatial data and correlation analysis (Wakefield 2004). For example, data groupings that are not necessarily based on spatial aggregation may also be considered to be ecological. A related concept, termed modifiable area unit problem (MAUP), deals with data grouping of both spatial and non-spatial aggregations (Openshaw and Taylor 1979). An example of voting-share versus income in the 2008 US presidential election is given here, in which the ecological correlation also leads to a manifestation of the Simpson's paradox.

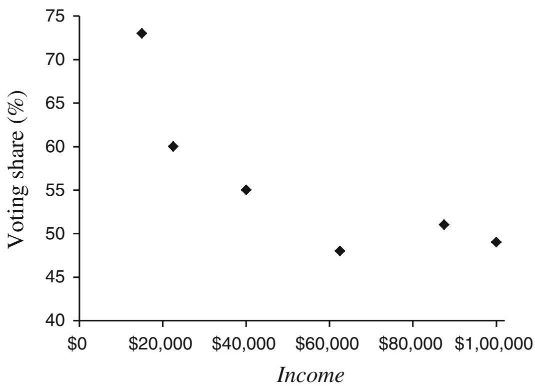
The voting share of the Democratic candidates, Obama–Biden, in the 2008 US presidential election has a negative correlation of −0.815 with the household income bracket (Fig. 1), but has a positive correlation of 0.534 with the median household income at the state level (Fig. 2). Excluding three outliers, Alaska, Utah, and Wyoming, the latter correlation becomes 0.713. The correlation between income and voting share at the state level is ecological because of spatial aggregation of the individuals within each state. Some consider the relationship in Fig. 1 to be a substitute for individual correlation, but it is actually an ecological correlation as a result of the aggregations of income (e.g., grouping of the voters with incomes between $35,000 and $45,000). Thus, the correlation from income grouping and the correlation from aggregation by states are both ecological, but in different ways. In this example, the correlation at the individual level will never be available because it would require surveying data for each voter's income and their vote, which is impossible to obtain. A similar reversal of correlation between income and voting share for the Republican candidates, Bush-Cheney, occurred in the 2004 US presidential election (Gelman et al. 2007).

1.2 Reference Class, Ecological Correlation, and Simpson's Paradox

The reference-class problem arises when an event is incorporated into a class, while the event possesses a number of attributes. Given a different attribute, the event may be assigned into a different class. Since the reference-class problem was recognized (Venn 1876), it has received rather limited attention, except by a small number of science philosophers and social analysts (Gillies 2000; Hajek 2007). Recently, the concept has been extended to applications in project planning and risk management (Lovallo and Kahneman 2003; Fryvbjerg 2006) as well as in spatial data analysis (Ma et al. 2008; Ma 2009). These extensions emphasize the construction or definition of reference classes, which mitigates the ambiguity of assigning individual events to a reference class. For example, a heterogeneous spatial or temporal phenomenon can be divided into reference classes, within which the phenomenon is relatively homogeneous (Ma et al. 2008). A nonstationary stochastic process can often be disaggregated into locally stationary processes (Matheron 1989; Ma 2010).

Reference classes can be defined solely based on the attributes that are related to the reference classes, without resorting to a specific spatial or temporal aggregation (Lovallo and Kahneman 2003; Fryvbjerg 2006). The example shown in Fig. 3 illustrates the Yule-

Fig. 1 Voting share (in percentage) of the Democratic candidates, Obama–Biden, as a function of household income level (grouping based on $10,000 range) in the 2008 US presidential election



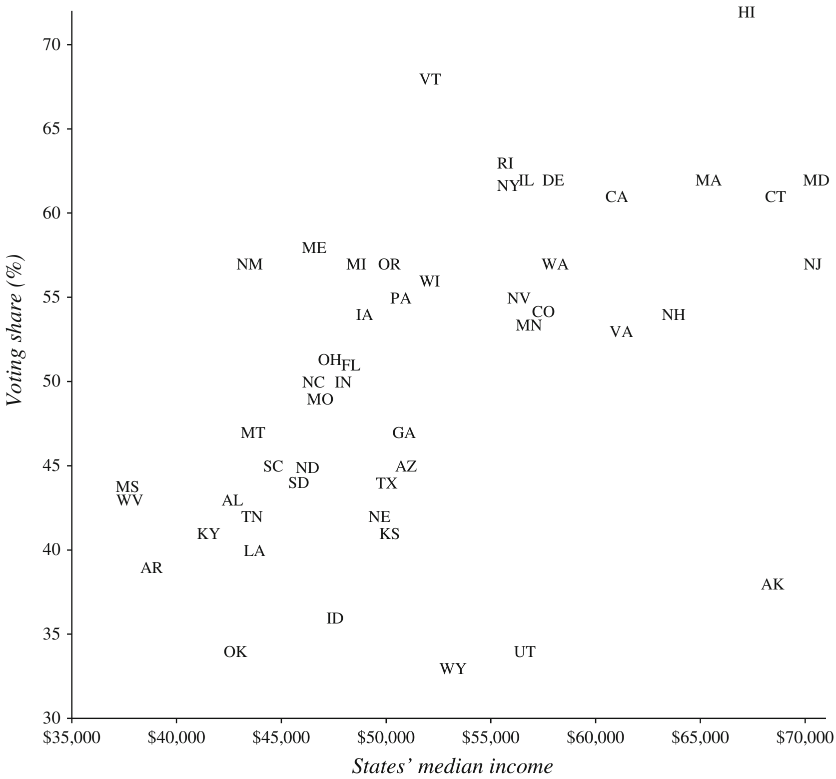


Fig. 2 Voting share (in percentage) of the Democratic candidates, Obama–Biden, as a function of the states' median income in the 2008 US presidential election. The overlain texts are the US state symbols

Simpson's Effect (Yule 1903; Blyth 1972) in an analysis of continuous variables when two different reference classes are mixed. Result of a survey of 63 graduate students indicates that the correlation between their own height and the ideal height of their future spouses is −0.414. At first glance, it appears that the greater the height of a student, the shorter he/she looks for the spouse's height. However, two distinct clusters exist due to the gender difference. The upper-left cluster in Fig. 3 represents female students seeking prospective male spouses with a certain height, and the lower-right cluster represents male students seeking prospective female spouses with a certain height. The two clusters are simply two different reference classes. The within-class correlation for the female reference class is 0.763; it is 0.649 for the male reference class. Clearly, the overall negative correlation that is reversed from the within-class correlations is spurious, which is related to the group means of the two reference classes. Imagine that with more reference classes, it will be possible to calculate the correlation in the group means, which is an ecological correlation, as discussed earlier. On the other hand, the positive correlations that are conditioned to the two gender-based reference classes are intrinsic correlations.

The more common manifestation of the Simpson's paradox for categorical variables has been discussed elsewhere (Bickel et al. 1975; Meng 2009; Ma and Ma 2011), as it occurs frequently in social- and medical data analysis. For example, a medical treatment, compared to placebo, appears to be "bad for men and bad for women, but good for people" (Ma 2009) or "good for men and good for women, but bad for people" as shown in

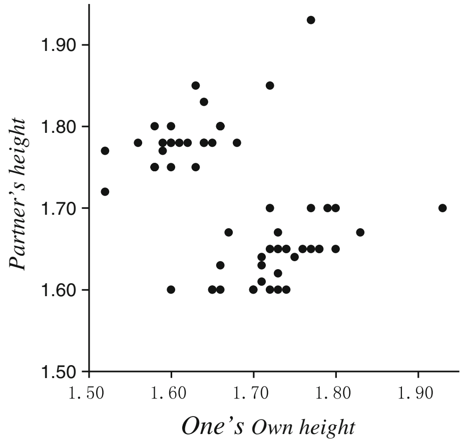


Fig. 3 Scatter plot between one's height (in meter, X-axis) and the ideal height of one's spouse (in meter, Y-axis). *Source* a survey conducted in the international classes for applied spatial statistics for natural resource evaluation at China University of Petroleum (Beijing) in June 2012. Some data are *overlapped* because some people have the same height and seek for a spouse of the same height

Table 1. Examples related to happiness are discussed later to illustrate the effects of the third variables on the correlation between happiness and income.

1.3 Scope and Results of the Study

This article examines and explains the happiness–income paradox using ecological correlation, change of reference classes, and confounding variables. The controversies related to the paradox are resolved when ecological correlations and third-variable effects are accounted for. Specifically, due to the third-variable effects, happiness and income at the individual level within a given country are correlated positively, but not as strongly as many believe. In international comparisons, they are generally quite strongly correlated as well, contrary to what Easterlin (1974) predicated and similar to what other researchers have found, but for different reasons, namely the ecological correlation by geographic aggregations. The long-term comparison is related to ecological correlation as well, but it is also related to the change of the reference classes.

2 Happiness–Income Relationship within a Country

2.1 Review of Previous Works

Cross-sectional studies within various countries have confirmed that personal income is the foremost factor that affects the human's subjective sense of well-being, and happiness is positively correlated to income at a given point in time (Easterlin 1974, 2001; Stevenson and Wolfers 2008). As an example, based on result of a survey, the percentage of people who expressed their level of happiness as a function of six income brackets is shown (Fig. 4). The correlation between income and the percentage of people expressing

Table 1 Comparison of the recovery rates between a treatment and the placebo

	Placebo	Treatment
Men	34 % (12/35)	36 % (32/90)
Women	76 % (72/95)	80 % (32/40)
All patients	65 % (84/130)	49 % (64/130)

happiness is 0.864. Other surveys of happiness as a function of income level have generally given similar results (Easterlin 2001; Stevenson and Wolfers 2008).

Happiness increases more significantly with increased income for people with low-to-moderate income levels, but increases more slowly for the individuals with high-income levels. This phenomenon, sometimes referred to as the diminishing-marginal utility, is illustrated in Fig. 5 (see also Easterlin and Angelescu 2009; Kahneman and Deaton 2010). This nonlinearity in happiness–income relationship affects the correlation between happiness and income, since the Pearson's correlation measures the linear relationship between two variables.

More importantly, all the variables involved in these cross-sectional studies are aggregated quantities with a non-negligible "sample-support", as the incomes are bracketed into levels and happiness is judged by either the percentage of, or the mean value of happiness of the respondents who expressed being happy within each income bracket. Therefore, results of these cross-sectional studies must be re-interpreted using ecological correlation based on the grouped properties (Robinson 1950; Openshaw and Taylor 1979). Some researchers pay attention to spatial aggregations, but not to other aggregations, such as income level or time duration. As a result, they unconsciously ignore the difference between income-aggregated ecological correlations and the individual-level correlation. As Robinson (1950) demonstrated, the individual correlation is generally not the same as its counterpart at the ecological level. The correlation between happiness and income at the individual level should be much lower than the ecological correlation, which involves aggregations of income and/or happiness in a cross-sectional study. Causally, the lower correlation between happiness and income at the individual level is due to the third-variable effects, which are discussed below.

2.2 Factors That Affect Happiness and Counteract the Correlation between Happiness and Income

Although income appears to be the most important factor that affects human well-being, there are many other factors that affect happiness, including health, family, entertainment, job satisfaction, social networking, and personal value (see e.g., Stiglitz et al. 2010; Kahneman and Deaton 2010; Toepoel 2013). All of these variables impact the correlation between happiness and income. Below, three cases are presented to illustrate the third-variable effects that may counteract the correlation between happiness and income.

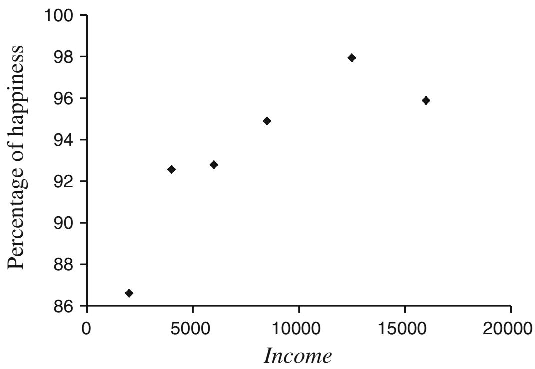


Fig. 4 Percentage of people who expressed happiness in a poll by AIPO in December 1970. *Source* Easterlin (1974, Table 2); 100 people were surveyed for each income bracket

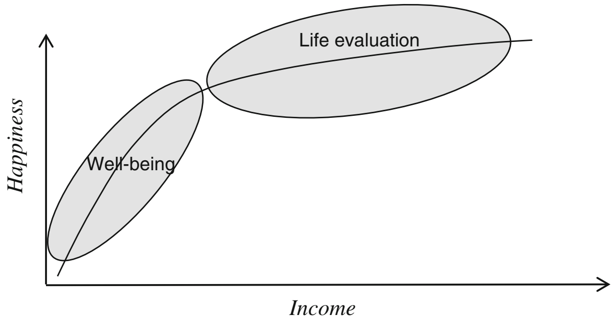


Fig. 5 Happiness–income relationship with two reference classes: emotional well-being is strongly correlated with income, but life evaluation correlates with income more moderately. The schema is synthesized based on several publications (Kahneman and Deaton 2010; Easterlin and Angelescu 2009; and Inglehart et al. 2008)

2.2.1 *Health and Health Care*

Health and health care are important factors that affect happiness. In fact, being healthy and being happy are linked, as are being unhealthy and being unhappy. In particular, maintaining good health, while controlling medical costs, is one of the top priorities for many elderly people. It is quite common that people sacrifice their health during their youth while gaining wealth, but they spend their wealth to regain their health in later life. Variables related to the elderly health issues include possible utilization of a nursing home, medical reimbursement policy, primary care, outpatient services, and in-hospital services, among others.

Several demographic trends are important in understanding the happiness surveys. One such trend is that women, on average, are happier than men, even though women have less average income and wealth than men. Some of the reasons are that women tend to be better at controlling costs, have better budget planning, and live longer. Sometimes, for certain individual categories of spending, women may outspend men, but they typically spend less in terms of the total budget. Table 2 presents an example of the Simpson's paradox, whereby elderly women outspent their male counterparts in nursing-home and all other medical cares, but they under-spent their male counterparts in the overall health-care cost. This is because, compared to elderly men, a larger proportion of the elderly women go to nursing homes, where health care costs are much lower than the average costs of outpatient and in-hospital services. This happens despite the fact that women live longer, represent a larger proportion of the elderly population, and have more chronic illnesses (Kronman et al. 2010).

2.2.2 *Entertainment*

Entertainment has a significant impact on happiness. High-income individuals tend to have less time for leisure and entertainment, which can negatively affect their happiness. On the other hand, low-to-moderate income individuals have more time for entertainment, which generally increases their sense of well-being. As an example, the "hot-hand" phenomenon in the game of basketball illustrates how people enjoy entertainment, thus increasing their happiness, without necessarily having a high-level understanding of the game.

Table 2 Simpson's paradox as exhibited in the annual health care costs for elderly (>66 years old) men and women

	Nursing home residents	Others	All (66+ years old)
Women	$16,000	$26,000	$22,000
Men	$15,500	$25,000	$23,000

This table was compiled, with rounded numbers, based on the data in Kronman et al. (2010)

Several studies (Tversky and Gilovich 1989; Koehler and Conley 2003; Sundali and Croson 2006) have shown that players do not exhibit "hot-hand" individually, but rather display a "regression towards the mean." Psychologically, however, fans are not excited by the dull concept of "regression towards the mean" or the randomness in players' shooting sequences. They wish to believe in the hot-hand phenomenon because it makes the game more exciting. Table 3 shows the aggregation effect using the statistics of free throws made by two players during a NBA basketball season. Individually, neither Bird nor Robery of the Boston Celtics showed the hot hand since their second free throw was worse if their first one was a hit than if their first one was a miss. But in the aggregated marginal table, the hot-hand phenomenon is present, i.e., their second free throws were much more accurate if their first ones were a hit. Therefore, the difference between the perception of the fans and what actually happens in the individual player's performance is due to the aggregation effect. In this case, the occurrence of Simpson's reversal is explainable by the confounding effect due to the large difference in the free-throw shooting percentage and the disproportion in the number of attempts between the two players. The objective interpretation of this example is that individual players do not show hot hand. Fans see patterns in a random sequence, but that does not prevent them from being happy to see the "streak shooting" phenomenon.

2.2.3 *Family Situation*

Family situation can affect an individual's sense of well-being in many ways. For instance, the world population has grown fast during the past several decades, but population growth is quite slow in most of the developed countries. The fast population growth in the developing countries has contributed to their lower per-capita income, and the overall lower per-capita GDP of the world. In many developing countries, rapid population growth is partly due to improved health care, but also due to the fact that people want to have more children, which has the dual effect of reducing the average per-capita GDP while increasing the individual's subjective well-being, and thus reducing the correlation between happiness and income. This dual effect contributes to a manifestation of the

Table 3 Comparison of two Boston Celtics players' free throws during the 1981 NBA basketball season

	L. Bird		R. Robery		Aggregated	
First shot	Second	Hit/total	Second	Hit/total	Second	Hit/total
Hit	0.881	251/285	0.593	54/91	0.811	305/376
Miss	0.906	48/53	0.613	49/80	0.729	97/133

Source extracted from Vokey (1997)

Table 4 Per-capita GDPs of the developed and the developing countries in 1995 and 2001

	1995		2001		Change	
	Per-capita GDP (USD)	Population (millions)	Per-capita GDP (USD)	Population (millions)	Per-capita GDP	Population (%)
Developed countries	24,315	1,017	24,720	1,057	1.67 %	3.93
Developing countries	1,065	4,666	1,162	5,101	9.11 %	9.32
World	5,225	5,683	5,205	6,158	−0.38%	8.36

Source World Bank (2011). The developed countries are labeled high-income by the World Bank while the developing countries are those defined by low and middle incomes. The per-capita GDPs are in 2010 US dollar (USD)

Simpson's reversal when aggregating the per-capita GDPs of the developing and the developed countries (Table 4).

2.3 Simulation of Happiness–Income Relationship at the Individual Level

Because most of the cross-sectional studies of happiness–income relationships use aggregated group-level data, the inferences about the correlation at the individual level based on the aggregated data are thus affected by an ecological inference bias. This section discusses the simulations of the correlation between happiness and income at the individual level.

Openshaw and Taylor (1979) showed an example in which different types of aggregations caused the correlations to range from −0.99 to 0.99. The correlation from an aggregation by income, such as that shown in Fig. 4, is higher than the correlation at the individual level within a given country. While the individual correlation cannot be easily estimated for the total population of a country or of the world, simulating the impact of other influential variables can help assess the correlation at the individual level.

To accurately simulate the correlation between happiness and income at the individual level would require frequency distributions for income, happiness, and all the third-variable effects, which are inaccessible. It is, nonetheless, reasonable to assume an overall variation of happiness for a given income level. Extending the example shown in Fig. 4, a simulation using a normal distribution with a standard deviation of 0.4 while preserving the mean happiness for each income bracket, gives a correlation of 0.535 at the individual level. This is illustrated by the scatter plot in Fig. 6a. A simulation with a standard deviation of 0.8 gives a correlation of 0.298 (Fig. 6b). When the variance is a little lower in the simulation, then the correlation is higher than 0.535, but reasonable values of standard deviation used in the simulation generally results in a correlation ranging from 0.30 to 0.75.

3 Happiness–Income Relationship across Countries

Based on the relationship between mean happiness and per-capita GDP of 14 countries, Easterlin (1974) argued that the correlation between happiness and per-capita GDP by country was weak. It is noteworthy that Easterlin's interpretation was focused on the graph's appearance without quantifying the relationship. Using the same data from

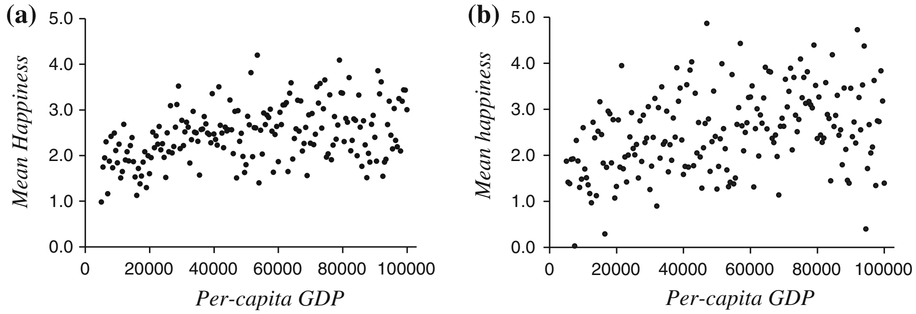


Fig. 6 Simulations of the happiness–income relationship at the individual level, based on the example in Fig. 4. A normal distribution is used while the mean happiness is not changed, (**a**) with a standard deviation of 0.4, and (**b**) with a standard deviation of 0.8

Easterlin (1974), the correlation between these two variables is actually quite significant, at 0.487, as illustrated by the scatter plot (Fig. 7). Even after excluding four outliers, India, the Unite states, Cuba, and the Dominican Republic, as suggested by Easterlin, the correlation is still quite significant, at 0.328. Other works (Hagerty and Veenhoven 2003, Stevenson and Wolfers 2008) have also confirmed a significant correlation between mean happiness and per-capita income in international comparisons. Figure 8 presents a scatter plot between mean life satisfaction and per-capita GDP using the World Value Survey and World Bank's (2011) data, respectively. The correlation between the two variables is significant, at 0.603.

When comparing different countries, more confounding variables are at play (Veenhoven and Hagerty 2006; Sarracino 2013), and these confounding variables can change, or even reverse, the correlation. In the two preceding examples (Figs. 7 and 8), many Latin American countries are off the trend line. Although the six Latin American countries had less than one-fifth the per-capita GDP of the developed countries, their mean happiness is equally high. In particular, Colombia and Mexico had the highest happiness levels among all the 50 countries surveyed. Excluding six Latin American countries in the second example (Fig. 8), the correlation between life satisfaction and income is even higher, at 0.774.

Another way of looking at the data is to exclude the 16 eastern European countries because they were considered to be in transition after regime changes (Inglehart et al. 2008; Easterlin and Angelescu 2009). The correlation without these countries is reduced to 0.380. Recall that in Easterlin (1974)'s example, the correlation coefficient was 0.487. This is because Easterlin's example included two eastern European countries that were not in transition during the time of the survey. Therefore, even though the two examples present a difference in correlation, they are actually quite consistent.

Because of cultural and other country- or region-specific effects, each country or region can be considered as a reference class and sensitivity of the variables that impact happiness may be different for different countries or regions (Sarracino 2013). Within each reference class, happiness level may be different and the happiness–income relationship may differ. Other arguments that favor distinguishing the within- and between-reference-class correlations include the so-called "neighborhood" effect. Keeping up with others in the same reference class is commonly an important consideration in determining an individual's subjective well-being. According to Easterlin (1995), after comparing the developed and developing countries' happiness in the 1960s, the investigator concluded that people in the developing countries had more modest aspirations because people were unaware of the life

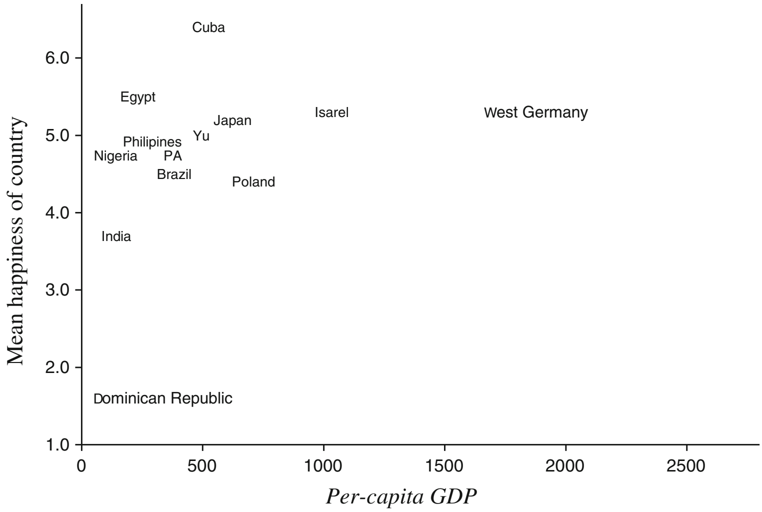


Fig. 7 Scatter plot between average happiness per country and the per capita GDP in the late 1950s to the early 1960s (data from Easterlin 1974, Table 6). *Note PA* Panama, *Yu* Yugoslavia, and *US* United States

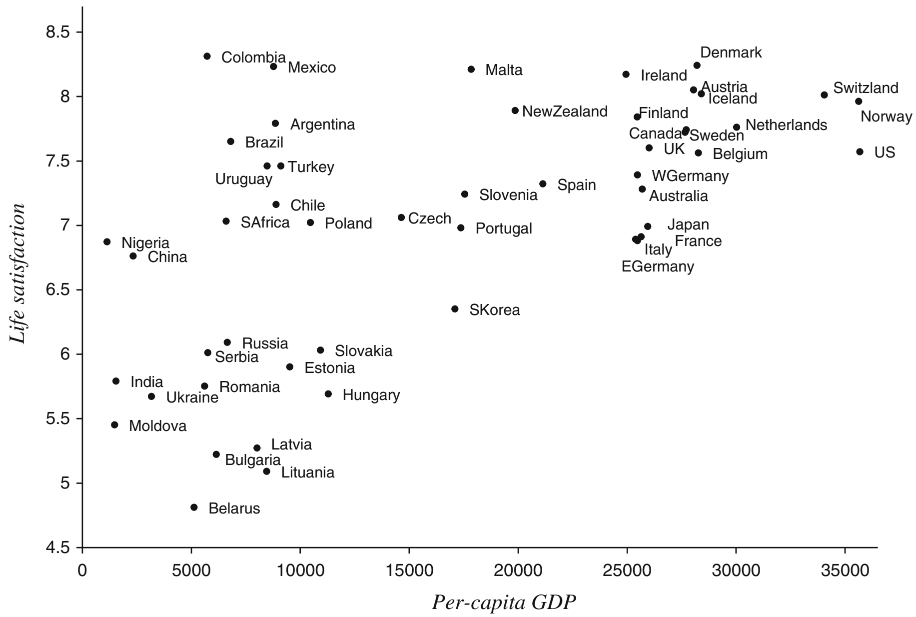


Fig. 8 Scatter plot between average life satisfaction per country and the per capita GDP (using Purchasing Power Parity or PPP, in international dollars) of the country between 1999 and 2007 (the latest available for each country). *Source* Per capita GDP data (PPP) are from the World Bank (2011) and life-satisfaction data from a worldwide value survey by Inglehart et al. (2008, table in Appendix A)

potential that was enjoyed by people in the developed countries. As the income level increases, the aspiration increases as well. This reflects an adaptation to income that rescales the happiness level.

Figure 9 shows a 2-level model that describes higher within-country correlations, and a weaker, but still significant, across-country correlation. Note that both correlations are ecological as a result of aggregations. The model is an example of multilevel analysis and modeling (Hox 2002; Welzel and Inglehart 2010) and it is also consistent with the relative-wealth hypothesis (Bernanke 2010). The fact that people in wealthy countries are, on average, not significantly happier than people in middle-income countries can be explained by the diminishing marginal utility, as discussed earlier for the individual's subjective well-being (Fig. 5). For countries with income that sufficiently meets their citizens' basic needs, income becomes less important because of the considerations of other factors, such as health, leisure time, work-life balance, entertainment, and social networking.

Another third-variable effect may also explain the difference in correlation that was observed in the two surveys (Figs. 7, 8). During the early survey in the late 1950s and early 1960s, there were more boundaries between countries, and the countries in the survey differed to a greater degree from one another. Thus the difference between the reference classes was strong. The second survey, on the other hand, has been highly influenced by globalization since the early 1990s. Globalization has made the world "flatter" or a global "village", and the country-level reference class is less important. The comparison is more global, frequently extending beyond the country boundaries.

Besides the third-variable effects discussed above, the aggregation effect is also important. Since Robinson's article (1950), researchers have emphasized the difference between individual correlation and ecological correlation; but much less attention has been paid to the problem when two different types of ecological correlations are mixed, especially for cases with non-spatial aggregation of the data. The correlation between happiness and income across countries is lower than its counterpart within a given country because of the difference in aggregation and the third-variable effects; but such correlation is still quite high, contrary to what Easterlin (1974) predicated, even using his original data. This is not necessarily for the same reasons that others have thought, but because there is an aggregation effect in using the average income or the per-capita GDP.

4 Happiness–Income Relationship in a Long-Term Time Series

The correlation between happiness and income in a long-term time series often appears weak. For example, although per-capita GDP rose steadily in the United States between

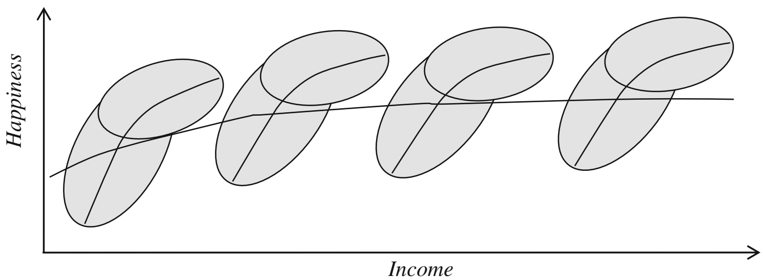


Fig. 9 Two levels of the happiness–income relationship: within a given country and across countries. Each level includes two trends within a given country: a strong correlation between well-being and income, and a weaker one between life evaluation and income. Across the countries, a moderate to strong correlation exists for low-to-middle-income countries, while high-income countries exhibit a weaker correlation

1946 and 1970, the average reported happiness did not exhibit any obvious long-term trend, with increases in some periods and declines in other periods (Fig. 10). While the per-capita GDP has increased more than eight times from 1972 to 2006, happiness increased only marginally during these 34 years. The period between 1972 and 1974 corresponds to the late stage of the Vietnam War and the general public confidence and their expression of happiness were lower. Excluding that period, there was essentially no overall increase in the reported happiness (Fig. 10).

Similarly, happiness in Great Britain between 1995 and 2006 did not change much while the per-capita GDP more than doubled (Fig. 11a); the correlation between happiness and per-capita GDP is only 0.10. Happiness barely changed in France between 1975 and 1986 while the per capita GDP more than doubled (Fig. 11b). The correlation between the two variables is slightly higher, at 0.29, but is still much lower than the correlation in a typical cross-sectional study. Most western European countries had a similar experience in this period, with some countries exhibiting even a negative correlation. This was the case in Belgium, where the per-capita GDP and happiness had a negative correlation of −0.17 (Fig. 11c). Incidentally, this example shows that it is imprudent to equate correlation to

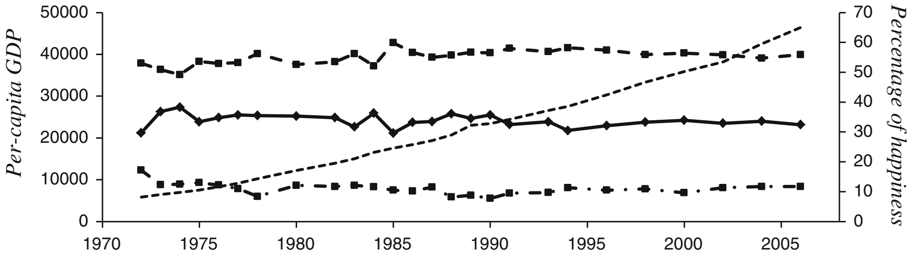


Fig. 10 Percentage of happiness (*right scale*) and per-capita GDP in the US (*dashed line*) from 1972 to 2006 (*left scale*). *Bottom dotted line*—not happy; *top dashed line*—happy; and *middle solid*—very happy. *Source* Happiness data from General Social Survey (GSS) and the per-capita GDP data from the World Bank (2011)

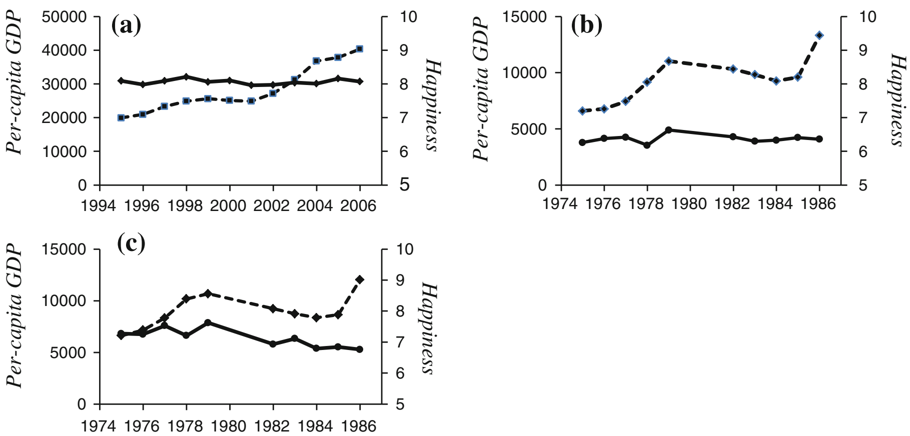


Fig. 11 Per-capita GDP (*dashed line*) and happiness (*solid line* 0–10 scale) in time series. **a** UK from 1995 to 2006. **b** France from 1975 to 1986. **c** Belgium from 1975 to 1986. *Source* GDP per capita data from the World Bank (2011) and happiness data from World Database of Happiness (2012)

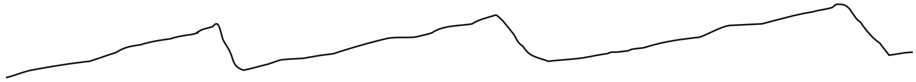


Fig. 12 A schematic diagram illustrates the short-time increases and long-term stagnation of happiness as a result of rescaling, particularly after big negative events

causation, because of the third variables' effect. If this were the case, Belgium could give money to a country with a positive correlation so that people in both countries would be happier. The more they give, the happier people of both countries would be. Alternatively, the Belgium government could make an effort to reduce the per-capita GDP in order to make the citizens happier!

Why are the short-term and long-term relationships between happiness and income so different? Happiness is rated on a bounded scale, for example, between 0 and 10, but the income is unlimited. Figure 12 illustrates the short-time increases and the long-term stagnation of happiness as a result of rescaling. Big events, especially negative events, rescale the happiness, and a new temporal reference class is created. While income increases without a mathematical bound, mean happiness changes little over long periods of time because of the rescaling that forms a new reference class.

5 Conclusion and Discussions

The happiness–income relationship is analyzed using income-grouping ecological correlation (cross-sectional studies within a country), spatial aggregation ecological correlation (across countries), change of reference class (across countries, as well as long-term time series), and the third-variable effects. This analysis suggests the high impact of the types and levels of grouping, such as income or geographic region, on the happiness–income relationship. Thus, the happiness–income paradox is related to an ecological inference bias. The correlation between happiness and income at the individual level within a country has been perceived to be very strong, largely due to the income-grouping ecological inference. In reality, the correlation is significant, but is significantly lower at the individual level. Causally, the lower correlation is due to the third-variable effect, which is shown by several presented examples of the Simpson's reversal. Simulations have shown the correlation to be in the range of 0.30–0.75 at the individual level, versus a correlation above 0.85 when an income grouping is used. As we have suggested earlier, correlation is not necessarily causation; there is certainly a causal relation between happiness and income, but it is almost impossible to quantify it accurately. It is likely that the causal relation between happiness and income at the individual level is generally a little higher in the developing countries than in the developed countries, conforming to the concept of diminishing marginal utility (Fig. 5).

The happiness–income correlation across countries is generally quite high as a result of geographic grouping. In fact, it is sometimes stronger than the correlation at the individual level within a country. This is contrary to what Easterlin (1974) initially predicated because his estimate of the within-country correlation was based on income-grouping that had a stronger effect than the geographic country-level grouping.

Recent studies have shown that judging happiness solely based on income is a focusing illusion because many other variables are also very important to individual's subjective well-being (Kahneman et al. 2006; Bernanke 2010; Stiglitz et al. 2010). The relatively moderate correlation between happiness and income at the individual level causally can be explained by the third-variables' effects. Blaise Pascal once remarked that *"All men seek happiness. This is without exception."* Yet, there are many ways to attain happiness, and many variables impact happiness. Multiple sources of causation to happiness lead to a moderate correlation between happiness and income at the individual level. Correlation between a cause and the effect tends to be moderate, rather than strong, when there are multiple causes, especially when different causes are counteracting one another. The aggregation effect on the happiness–income relationship leads to an ecological inference bias, while the impact of third-variables on the happiness–income relationship is a manifestation of the Simpson's paradox or the Yule-Simpson' effect.

Although we did not study the correlation at the individual level using worldwide data, that correlation should be weaker than the within-country correlation. This apparently is similar to the Easterlin's original claim, but we have demonstrated that Easterlin made the claim using ecological correlations that suffer from an inference bias; so do the arguments against his claim. While Easterlin now has shied away from his original claim on the difference in correlation across and within countries (Easterlin et al., 2009, 2010), we state a different interpretation of this paradox at the individual level – moderate correlation within a country and weaker correlation worldwide. We argue that the "paradox" exists when the correlations at the individual level are used for both within a country and across countries. This hypothesis is formed based on our discussions in Section 3; in particular, mixing heterogeneous reference classes (different countries have large differences in their income distribution and happiness attainment) has an effect of reducing the aggregated across-country correlation from the within-country correlations. This is a manifestation of the Simpson's paradox. The 2-level model (Fig. 9) also supports this hypothesis.

Acknowledgments The authors thank Samantha Perkins, Jason Sitchler and Ernest Gomez for reading and commenting the manuscript. The opinions expressed in this article are those of the authors only; they do not necessarily reflect the opinions of their respective institutions.

References

Bernanke, B. S. (2010). The economics of happiness. Federal Reserve Chairman's Speech: http://www.federalreserve.gov/newsevents/speech/bernanke20100508a.htm#f7.

Bickel, P. J., Hammel, E. A., & O'Connell, J. W. (1975). Sex bias in graduate admissions: Data from Berkeley. *Science, 187*, 398–404.

Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association, 67*(338), 364–366.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal Economic Literature, 46*, 95–144.

Cressie, N. (1996). Change of support and the modifiable areal unit problem. *Geographical Systems, 3*(2–3), 159–180.

Di Tella, R., & MacCulloch, R. (2008). Gross national happiness as an answer to the Easterlin paradox? *Journal of Development Economics, 86*, 22–42.

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz*. New York: Academic Press.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization, 27*, 35–47.

Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The Economic Journal, 111*, 465–484.

Easterlin, RA., and Angelescu L. (2009) *Happiness and growth the world over: Time series evidence on the happiness–income paradox*, IZA Discussion paper, No. 4060.
Easterlin, R. A., Angelescu, Mc Vey L., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences of the United States of America, 107*(52), 22463–22468. doi:10.1073/pnas.1015962107.
Fryvbjerg, B. (2006). From Nobel Prize to project management: Getting risk right. *Project Management Journal, 37*(3), 5–15.
Gehlke, C. E., & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association, 29*(185), 169–170.
Gelman, A., Park, D., Shor, B., Bafumi, J., & Cortina, J. (2007). Rich state, poor state, red state, blue state: What's the matter with Connecticut? *Quarterly Journal of Political Science, 2*, 345–367.
Gillies, D. (2000). *Philosophical theories of probability* (p. 223). London: Routledge.
Gotway, C. A., & Young, L. J. (2002). Combining Incompatible spatial data. *Journal of American Statistical Association, 97*(458), 632–648.
Grotenhuis et al. (2011) Robinson's ecological correlations and the behavior of individuals: Methodological corrections, *International Journal of Epidemiology*, 1–3.
Hagerty, M. R., & Veenhoven, R. (2003). Wealth and happiness revisited—Growing national income does go with greater happiness. *Social Indicator Research, 64*, 1–27.
Hajek, A. (2007). The reference class problem is your problem too. *Synthese, 156*(3), 563–585.
Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.
Inglehart, R., Foa, R., Peterson, C., & Welzel, C. (2008). Development, freedom, and rising happiness. *Perspectives on Psychological Science, 3*(4), 264–285.
Kahneman, D., & Deaton A. (2010) High income improves evaluation of life, but not emotional well-being. In: *Proceeding of the National Academy of Science of the USA*, August 2010, doi: 10.1073/pnas.1011492107.
Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. A. (2006). Would you be happier if you were richer? A focusing illusion. *Science, 312*, 1908–1910.
Koehler, C., & Conley, A. (2003). The "hot hand" myth in professional basketball. *Jounal of Sports and Exercise Psychology, 25*, 253–259.
Kronman, A. C., Freund, K. M., Hanchate, A., Emanual, E. J., & Ash, A. S. (2010). Nursing home residence confounds gender differences in Medicare utilization. *Women's Health Issues, 20*, 105–113.
Lovallo, D., & Kahneman, D. (2003). *Delusions of success: How optimism undermines executives' decisions*, Harvard Business Review, July, pp. 56–63.
Ma, Y. Z. (2009). Simpson's paradox in natural resource evaluation. *Mathematical Geosciences, 41*, 193–213.
Ma, Y. Z. (2010). Error types in reservoir characterization and management. *Journal of Petroleum Science and Engineering, 72*(3–4), 290–301. doi:10.1016/j.petrol.2010.03.030.
Ma, Y. Z., & Ma, A. M. (2011). Simpson's paradox and other reversals in basketball: Examples from 2011 NBA playoffs, Intl. *Journal of Sports Science and Engineering, 5*(3), 145–154.
Ma, Y. Z., Seto, A., & Gomez, E., 2008, Frequentist meets spatialist: A marriage made in reservoir characterization and modleing. SPE 115836, SPE ATCE, Denver, CO.
Matheron, G. (1989). *Estimating and choosing: An essay on probability in practice*. Berlin: Springer. 141p.
Meng, X. L. (2009). Statistics: Your chance for happiness (or misery). *The Harvard University Undergraduate Research Journal, 2*(1), 21–27.
Openshaw, S., & Taylor, P. J. (1979). *A million or so correlation coefficients: Three experiments on the modifiable areal unit problem: In statistical applications in the spatial sciences* (pp. 127–134). London: N. Wrigley.
Robinson, W. (1950). Ecological correlation and behaviors of individuals. *American Sociological Review, 15*(3), 351–357. doi:10.2307/2087176.
Sarracino, F. (2013). Determinants of subjective well-being in high and low income countries: Do happiness equations differ across countries? *The Journal of Socio-economics, 42*, 51–66.
Simpson, E. H. (1951). The interpretation of interaction in contingency tables: *Journal of the Royal Statistical Society: Series B, 13*, 238–241.
Stevenson B. and Wolfers J. (2008) *Economic growth and subjective well-being: Reassessing the Easterlin paradox. Brookings Paper Economic Activity,* Spring pp. 1–87.
Stiglitz, J., Sen, A., & Fitoussi, J. (2010). *Why GDP does not add up: Mismeasuring our lives*. New York: New Press.
Sundali, J., & Croson, R. (2006). Biases in casino betting: The hot hand and the gambler's fallacy. *Judgment and Decision Making, 1*(1), 1–12.
Toepoel, V. (2013). Ageing, leisure, and social connectedness. *Social Indicators Research, 113*(1), 355–372.

Tversky, A., & Gilovich, T. (1989). The cold facts about the 'Hot Hand' in basketball. *Chance, 2*(1), 16–21.

Veenhoven, R., & Hagerty, M. (2006). Rising happiness in nations 1946–2004. *Social Indicators Research, 79*, 421–436.

Venn, J. (1876). *The logic of chance* (2nd ed.). MacMillan, 488 pp.

Vokey, J. R. (1997). Collapsing multiway contingency tables: Simpson's paradox and homogenization. *Behavior Research Methods, Instruments and Computers, 29*(2), 210–215.

Wakefield, J. (2004). Ecological inference for 2 × 2 tables. *Journal of the Royal Statistical Society: Series A, 167*(3), 385–445.

Welzel, C., & Inglehart, R. (2010). Agency, values, and well-being. *Social Indicators Research, 97*(1), 43–63.

World Database of Happiness (2012). http://worlddatabaseofhappiness.eur.nl.

Yule, G. H. (1903). Notes on the theory of association of attributes in statistics. *Biometrika, 2*, 121–134.