the review of income and wealth

Review of Income and Wealth
Series 57, Number 1, March 2011

LIFE SATISFACTION AND INCOME INEQUALITY

BY PAOLO VERME*

University of Torino

Do people care about income inequality and does income inequality affect subjective well-being? Welfare theories can predict either a positive or a negative impact of income inequality on subjective well-being and empirical research has found evidence of a positive, negative, or non-significant relation. This paper attempts to determine some of the possible causes of such empirical heterogeneity. Using a very large sample of world citizens we test the consistency of the effect of income inequality in predicting life satisfaction. We find that income inequality has a negative and significant effect on life satisfaction. This result is robust to changes of regressors and estimation choices and also persists across different income groups and across different types of countries. However, this relation is easily obscured or reversed by multicollinearity generated by the use of country and year fixed effects. This is particularly true if the number of data points for inequality is small, which is a common feature of cross-country or longitudinal studies.

1. INTRODUCTION

The role of income inequality in predicting subjective well-being is controversial.[1] Various theories put forward across the social sciences can predict either a positive or a negative impact of income inequality on subjective well-being. Empirical evidence which has emerged in studies carried out during the past few decades provides some support for both positions.

This paper returns to this question, proposes a number of possible hypotheses that could explain empirical heterogeneity of outcomes, and tests these hypotheses one by one. We find that, among the factors considered, multicollinearity is the most likely factor to explain empirical heterogeneity of results. In cross-country and longitudinal studies it is common to use country and year fixed effects to control for unobserved heterogeneity. This practice generates substantial collinearity between country and year dummies and variables estimated at the country/year level such as income inequality or GDP per capita. Such collinearity, in turn, can affect inference by changing sign and/or significance of the happiness–inequality relation. In cross-country or longitudinal happiness models, researchers face a real trade-off between addressing multicollinearity by dropping country and year fixed effects and addressing unobserved heterogeneity by keeping these variables into the model. Moreover, this trade-off increases in cost as the number of data points for inequality decreases.

Note: The author is very grateful to two anonymous referees who provided very useful comments. A special thank to the first referee who took the time to read the paper several times and provided extremely detailed and constructive comments.

*Correspondence to: Paolo Verme, Department of Economics "S. Cognetti de Martiis," University of Torino, Italy (paolo.verme@unito.it).

[1]For simplicity we consider well-being, utility, happiness, or life satisfaction as one and the same concept and measure it with a question on life satisfaction. This is a standard practice in happiness research (see, e.g. Easterlin, 2001; Alesina *et al.*, 2004).

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA, 02148, USA.

The paper is organized as follows. The next section discusses theory and practice of the study of income inequality and subjective well-being. Section 3 puts forward a number of hypotheses that could explain the different findings in the literature on income inequality and subjective well-being. Section 4 describes model, data, and variables, Section 5 presents the results, and Section 6 concludes.

2. Theory and Evidence

Studies on subjective well-being and income inequality have been partly inspired by the much larger literature on happiness and income. This literature has been rather consistent in finding that income is a good predictor of happiness across people and across countries but not over time and over the life-cycle. Individuals or countries with a higher income have been found to be happier (Inglehart, 1990; Diener *et al.*, 1995; Di Tella *et al.*, 2001; Blanchflower and Oswald, 2004) while longitudinal or life-cycle studies do not find a strong positive association between happiness and income (Easterlin, 1974, 1995, 2001; Veenhoven, 1993; Clark and Oswald, 1994; Mangahas, 1995; Diener *et al.*, 1999; Ravallion and Lokshin, 2000).

The search for an explanation of the paradox generated by these findings in longitudinal and life-cycle studies has led to the formulation of several theories, most of which focus on the role of the reference group and on the role of expectations. People consider their income relative to that of a reference group rather than absolute income and adjust expectations accordingly.

When applied to the context of income inequality, these theories can provide opposite predictions about the impact of inequality on subjective well-being. This is also the case for theories of revolutions, social justice, or relative deprivation which emerged during the second half of the twentieth century. As an example, take two of the most influential theories, the "tunnel" effect theory proposed by Hirschman and Rothschild (1973) and the relative deprivation theory proposed by Runciman (1966).[2]

Hirschman and Rothschild (1973) argued that people may appreciate inequality if this signals social mobility, a phenomenon dubbed by Hirschman as the "tunnel" effect. People who can observe others around them moving upwards in the income scale increase their expectations about their own social mobility and this makes them happier because it improves expectations about their own future.

This observation may be vulnerable to different criticisms. For example, an increase in others' mobility does not necessarily result in increased inequality if the upward "movers" are mostly poor people. Some people or income groups may be more sensitive than others to income mobility and some people may fear rather than appreciate mobility. And different people or groups of people may only be concerned with the mobility of a specific reference group rather than with the mobility of all others taken together.

[2]In this paper we do not provide a comprehensive review of the theoretical literature or offer an alternative theoretical model of the happiness–inequality relation. We simply provide one example of alternative theoretical views that could justify alternative empirical findings. For recent theoretical reviews and new models on the happiness–inequality relation, see Truglia (2007) and Hopkins (2008).

However, Hirschman and Rothschild referred to the population as a whole and did not discuss the implications for different tastes, income groups, or reference groups. They simply argued that increased social mobility for only part of a population leads to increased inequality, increased prospects for all, and increased individual and social welfare, at least in the short term.[3]

Runciman (1966) has instead devised a theory of social justice based on the notion that the individual sense of deprivation can be explained by the relative position that the individual occupies in relation to the self-selected reference group. Yitzhaki (1979) has formalized this concept applied to incomes, proposed to measure relative deprivation as the sum of the distances of a person's income from all incomes situated above in the income distribution, and showed how this measure is in fact equivalent to the absolute Gini index (the Gini multiplied by the mean). The prediction of the Runciman–Yitzhaki framework is that increasing income inequality increases relative deprivation and decreases subjective well-being.

Runciman theory implies that the poorest are the most deprived and those who appreciate the least income inequality. In this case, the reference group is always constituted by people with higher income, even if the reference group is restricted to sub-samples of the population. It does not matter whether the reference group is constituted by the poor, the rich, or both groups because individual satisfaction is only defined within the reference group.

Theory would therefore suggest at least two mechanisms through which income inequality may affect individual satisfaction. The first is that a rise in income inequality signals future mobility and increases present satisfaction. This implies a positive relation between income inequality and life satisfaction (the Hirschman–Rothschild mechanism). The second mechanism is that a rise in income inequality leads to an increase in relative deprivation and a decrease in life satisfaction (the Runciman–Yitzhaki mechanism). Moreover, while the Hirschman–Rothschild mechanism does not have clear predictions on which income group benefits the most from increased inequality, the Runciman–Yitzhaki mechanism indicates that the poor are more deprived and should be more inequality averse than the rich.

It is important to clarify at this point what we mean by inequality aversion and how we interpret the sign of the happiness–inequality relation. Economics and statistics offer different definitions of inequality aversion. One is the definition derived from risk theory, which describes inequality aversion as the concavity of the utility curve. A second is the inequality aversion parameter used in statistical indexes of inequality which attributes a different weight to incomes located in different parts of the income distribution. One example is the Atkinson inequality measure. A third is the inequality aversion measured with experimental questionnaires and games specifically designed to capture the taste for inequality; for example, the work conducted by Amiel and Cowell (1992). A fourth approach is to consider a negative relation between life satisfaction and income inequality as a

[3]In the long-run, if expectations for social mobility are not met, inequality can turn into an explosive social device. The Hirschman and Rothschild (1973) model predicts positive returns to increased inequality only if the benefits of expectations outweigh the cost of envy.

sign of inequality aversion. For example, Clark (2003) argues that workers may not be inequality averse because he finds a positive relation between happiness and income inequality, and Schwarze and Harpfer (2003) argue that Germans are only weakly inequality averse because a reduction in inequality does not increase well-being. In this paper we follow this last approach by interpreting a positive sign of the happiness–inequality relation as an indication that higher inequality is appreciated and provides a sense of satisfaction to individuals (the Hirschman–Rothschild mechanism), and a negative sign as an indication that higher inequality is not appreciated and provides a sense of dissatisfaction (the Runciman–Yitzhaki mechanism).

Empirical evidence on the sign and significance of the happiness–inequality relation is controversial and heterogeneous. As described below, one can find positive, negative, or non-significant relations depending on the particular study considered.

Morawetz *et al*. (1977) have shown how two communities in Israel with different levels of income inequality differed in average happiness; where income inequality was found to be higher, average happiness was found to be lower. Schwarze and Harpfer (2003) find life satisfaction to be negatively correlated with inequality using the German socioeconomic panel over 14 waves, and Hagerty (2000), using aggregated data for eight countries, finds that average happiness levels are lower where income distributions are wider. On the contrary, Clark (2003), using the British Household Panel Survey, finds a positive correlation between happiness and inequality for the employed population. A study by Alesina *et al*. (2004) found that individuals tend to be less happy if inequality is high, but that this effect is stronger in the EU than in the U.S. Also, the poor and left-wing people in the EU are less happy if inequality is high, while this phenomenon is not visible in the U.S. Graham and Felton (2006) looked at Latin American countries and found that inequality (measured in terms of relative wealth) made people in upper quintiles happier and those in the poorest quintile less happy, but they also find that the Gini coefficient is non-significant in a happiness equation. Senik (2004) does not find a significant correlation between happiness and inequality for Russia using the Russian Longitudinal Monitoring Survey. A study by Helliwell (2003) finds no evidence that income inequality is correlated with happiness and, according to Veenhoven (1996), "Income inequality in nations appears almost unrelated to final quality of life as measured by average happiness . . ." (p. 34).

Table A3 in the Appendix provides more detailed information on the cited literature in chronological order. Leaving aside the first study by Morawetz *et al.* (1977), we can observe some similarities and dissimilarities. The datasets used in these studies are all different with the exception of two papers which both use the US-GSS study. Three studies use longitudinal panels of individual observations, four studies use cross-country studies with multiple years, and one study uses a cross-country study with one year. The estimation models used can be ordered logit, ordered probit, or OLS and this is a normative choice rather than a choice dictated by the data. The measure of inequality is the Gini for all studies except for part of the Hagerty (2000) study. Some papers estimate the Gini from the dataset used while others extract the Gini from other datasets. The Gini can also be estimated for countries, regions, primary sample units (PSU), or particular reference groups. All studies use, in conjunction with the inequality measure, one or more measures of income such as income (in continuous or categorical form), lagged income, relative income, or measures of countries' wealth. Most studies use country or regional fixed effects but two studies do not, while years fixed effects are used by all studies with longitudinal data except one. Finally, some papers report the use of robust standard errors and/or cluster estimations, while other papers do not report how the standard errors have been estimated. In the next section, we will put forward some hypotheses on how these diversities in choices may contribute to explain diversity in results.

3. Some Hypotheses

There are several factors that may lead to controversial empirical results on the correlation between happiness and income inequality. Some of these factors relate to the specific data available or to the choice of the inequality measure made by the researcher. Other factors relate to econometric choices that may or may not relate to the data at hand. We discuss these two groups of hypotheses in turn.

The *choice of the inequality measure* is a first critical choice. Some studies use the Gini exogenous to the survey used for the life satisfaction estimations, others use the Gini calculated from within the surveys used. For example, Alesina *et al.* (2004) use the Gini taken from the Deninger and Squire database[4] and Helliwell (2003) uses the Gini taken from a World Bank database, whereas Senik (2004) and Clark (2003) calculate the Gini from within their own surveys.

This choice is mostly dictated by the data. The first two studies are cross-country studies that make use of values surveys. Values surveys such as the World Values Surveys, the European Values Surveys, and the U.S. Social Survey do not hold information on individual incomes in continuous form. Income is typically reported in terms of income classes. When these surveys are used, researchers either transform income classes into comparable monetary values or they draw on external sources for measures of inequality. This explains the choice of "exogenous" inequality variables. The second set of studies uses instead longitudinal data on single countries such as Russia, the U.K., or Germany where individual income is typically available in continuous form. The shortcoming here is that only a few panel surveys have questions on life satisfaction and one also needs many years or to split the sample into sub-groups to make some inference on the role of inequality.

Combining longitudinal and cross-country data can also lead to different conclusions. Suppose that we could use an "endogenous" and an "exogenous" income Gini simultaneously. Suppose also that both samples on which the Ginis are estimated are representative of the population under study. The two Ginis may, in fact, be different in value either because the income distribution cannot be identical in the two samples or because the welfare measure is different (such as income as opposed to consumption). Moreover, when the two Ginis are compared

[4]For the Europeans countries considered.

across countries and time, the cross-section and longitudinal distributions of such Gini may also be very different, affecting the covariance between income inequality and subjective well-being.

Another factor may relate to *different tastes for inequality* across different population groups. This may relate to different income groups, to different groups partitioned on other criteria such as region, gender, ethnic group, or others, or to different group of countries. Some population groups are more sensitive to or have opposite tastes for inequality than other population groups and it may be difficult to isolate which groups behave homogeneously. When studies do not disaggregate by relevant group the net effect may be non-significant. Moreover, people in different countries may have very different tastes for inequality due to cultural and other factors and this effect may overlap with the effect due to the different wealth of countries. Poor and rich countries may have different tastes for inequality.

A different set of explanations for the empirical heterogeneity relates to econometric factors. The *choice of key regressors* is a first critical choice. Combining different sets of regressors can lead to different results especially if these regressors include other measures of income or relative income which are likely to be correlated with the inequality measure under study. For example, the Gini index can be expressed as a function of income, income relative to the mean, distances from the mean, or distances from the median (see, e.g. Xu, 2004). Combining the Gini with other income measures is a rather common approach in happiness research because one of the recurrent themes is to test how relative income rather than income affects happiness. However, this has non-negligible statistical implications. Using in the same equation an income Gini and the income variable on which the Gini is calculated or another relative income measure can lead to multicollinearity and to unpredictable coefficients and standard errors. This is a point hardly considered in happiness studies but very relevant if we wish to explain the empirical heterogeneity in outcomes of these studies.

Second, the *use of country and year fixed effects* in cross-country or longitudinal studies may generate substantial collinearity with the inequality measure. By fixed effects, we mean including dummies for countries or regions in a cross-country study or dummies for years in longitudinal studies. These dummies are useful to account for unobserved country heterogeneity and time dependence and they are routinely included in empirical models. However, inequality measures are estimated at the country/year level and the use of country and year fixed effects leads to increased multicollinearity. Multicollinearity, in turn, can make parameter estimates sensitive to small changes in the data, can inflate standard errors and coefficients, and can also change the sign of predictors (Greene, 1997).

Multicollinearity also relates to the *number of data points available.* One may have hundred of thousands of individual observations, but what really matters for the relation happiness–inequality is the number of data points for the measure of inequality. When inequality is measured at the country/year level, the number of data points available in cross-country or longitudinal studies is limited.

An additional factor may be the *estimation of the standard error*. In particular, using a robust form of estimator and regional clusters may alter significantly the results in cross-country studies for a variable calculated on aggregated units such as inequality. The Gini coefficient is forcibly calculated on groups of individuals and this restricts the degrees of freedom. A robust estimation of the standard error provided by standard statistical packages makes use of estimators such as the Huber–White Sandwich estimator of variance which, by definition, changes the estimation algorithm of the standard error. And introducing clusters, such as regional clusters, relaxes the assumption that observations are independent and adjusts standard errors for intra-region correlation accordingly. Estimating standard errors with a Huber–White Sandwich estimator and regional clusters do not affect coefficients but affect inferences about coefficients and significance levels. The choice of estimation procedure for the standard error should normally be dictated by the underlying structure of the data, but researchers may have incomplete information on the original data structure or simply overlook some important aspects.[5]

In the rest of the paper we test how these different factors affect inference about the relation between subjective well-being and income inequality. The list of factors is non-exhaustive and we do not pretend to cover in this paper all possible causes of empirical heterogeneity. However, if the factors listed above contribute to explain such heterogeneity, then any inference from any study on the relation between happiness and inequality is context specific and cannot be generalized to other contexts. On the contrary, if life satisfaction and income inequality are strongly correlated, then the significance of this relation should persist under different specifications of the life satisfaction equation and the sign of this relation should be consistent irrespective of the factors listed.

4. Data, Model, and Variables

The *dataset* used has been compiled aggregating all rounds of the European and the World Values surveys carried out between 1981 and 2004.[6] These surveys question individuals worldwide on happiness, personal values, social attitudes, and individual attributes and include questions on income and inequality. The version of the dataset we use is a 2006 version which contains a total of 267,870 individuals, 1,349 regions, and 84 countries where each country has been

[5]Economically and statistically speaking, robust estimations are indicated when we expect heteroskedasticity or have outliers, while cluster analysis is indicated if we expect individuals to be very similar within sub-country clusters of observations (such as regions). While the use of robust estimations is mostly a statistical issue that can be decided looking at data distribution, the use of clusters requires some information on sampling and on the population at hand that may or may not be available to the researcher. In the case of welfare studies, if information on household welfare is very homogeneous within clusters, this is essential information to decide on the use of clusters. Therefore, in the absence of complete and reliable data information, the least risky choice would be to use both robust and cluster options, while the most transparent choice would be to compare and discuss results with and without robust and cluster options.

[6]Data can be freely downloaded from: http://www.jdsurvey.net. We are grateful to the Values Study Group and World Values Survey Association for creating and making accessible the European and World Values Surveys Four-Wave Integrated Data File, 1981–2004 (v.20060423, 2006). Aggregate File Producers: Análisis Sociológicos Económicos y Polticos (ASEP) and JD Systems (JDS), Madrid, Spain/Tilburg University, Tilburg, The Netherlands. Data Files Suppliers: Análisis Sociológicos Económicos y Políticos (ASEP) and JD Systems (JDS), Madrid, Spain/Tilburg University, Tilburg, The Netherlands/Zentralarchiv fur Empirische Sozialforschung (ZA), Cologne, Germany. Aggregate File Distributors: Análisis Sociol ógicos Económicos y Políticos (ASEP) and JD Systems (JDS), Madrid, Spain/Tilburg University, Tilburg, The Netherlands/Zentralarchiv fur Empirische Sozialforschung (ZA) Cologne, Germany.

surveyed from a minimum of one to a maximum of four times. Table A2 in the Appendix provides details on countries, years, and number of observations.

We also merged this dataset with two other variables: GDP per capita at purchasing power parity (PPP) extracted from the IMF world economic outlook database[7] and the Gini coefficient extracted from the United Nations University, World Institute for Development Economics Research (UNU–WIDER) database on inequality.[8] We use GDP per capita to control for countries' wealth and the UNU–WIDER Gini to adopt an alternative measure of income inequality independent of the database we use.

As a benchmark for our analysis, we use what we could call a "standard" model in happiness studies that combines cross-country and longitudinal data (see, e.g. Alesina *et al.*, 2004). Let H = subjective well-being; X = income; I = income inequality; R = relative income; W = a measure of countries' wealth; C = a vector of control variables for individual characteristics; T = a vector of country dummies; Y = a vector of year dummies; α, β, γ, δ, η = parameters to be estimated; ε = error term; i = individuals; c = countries; and y = years. We estimate the life satisfaction equation cross-section on a pooled sample of world citizens as described below:

$$H_i = \alpha X_i + \beta I_{cy} + \gamma R_i + \delta W_{cy} + \eta C_i + T_c + Y_y + \varepsilon_i. \tag{1}$$

A wide range of reduced specifications will be considered as well as alternative estimations of the standard error. We use the robust Huber–White sandwich estimator and regional clusters for a robust estimation of the standard error. As shown below, the dependent variable is categorical and all estimations are made with an order logit model.

As a measure of subjective well-being (H), we use *life satisfaction*. The question asked is: "All things considered, how satisfied are you with your life as a whole these days?" Answers include a ten-steps ladder where "1" stands for "Dissatisfied" and "10" stands for "satisfied." This is a common question used in happiness research; validation studies conducted by psychologists and social scientists show that answers to such questions are reliable (Fordyce, 1988; Inglehart, 1990; Sandvik *et al.*, 1993; Saris *et al.*, 1996; Lepper, 1998).

Income (X) is measured as self-positioning in a ten-steps income scale, where the income brackets have been measured in local currency in each country.[9] This is not self-declared income but the positioning of individuals into income brackets. In some sense, this is a more accurate indicator than self-reported income, which is known to be underreported in household surveys worldwide. That is because people are not asked to tell how much they earn but simply to say to which income bracket they belong to.

For cross-country comparability purposes, the income variable has been further transformed into mid-class values, real terms, USD, and PPP. In the World and European Values surveys, each country uses a ten-steps income scale where each step is reported in local currency. For each country/year, we first calculated

[7]Wired at www.imf.org/data.
[8]Version "WIID2C" wired at: http://www.wider.unu.edu.
[9]This variable is the only income variable present in the database.

the mid-class values in local currencies. For the lower class, we used the average between zero and the lower bound of the second class. For the upper class, we used the lower bound inflated by 20 percent. This is evidently a normative choice based on the notion that the distribution of incomes in the top decile is typically right-skewed, with most of the observations concentrated near the lower bound. The top class has a relatively small number of observations, and changing the inflation factor from 20 percent to, say, 30 percent has a very marginal impact on results. However, the upper class contains outliers, and if we had used higher inflation factors it is as if we were trying to better represent these outliers rather than the median value of the top class.

Mid-class values were then transformed into constant, USD and PPP values using the IMF GDP and PPP data published in the IMF economic outlook report. The IMF GDP data are reported in nominal values (local currency and USD), constant values with base 2000 (local currency), and PPP values (constant USD equivalent) providing all ingredients necessary for the transformations.[10] In order to check on the results of this work, we compared the resulting values from our database with the IMF GDP per capita PPP data and we also verified the consistency of results across countries and years.[11]

We use two different Ginis as alternative measures of *income inequality* (I). The first is calculated by country and year using the income variable already described, present in the database we use (Gini WVS for short). The second is the Gini coefficient taken from the UNU–WIDER database on inequality (Gini WIDER for short).

The Gini WIDER puts together country estimates of the Gini coefficient calculated from a variety of income and consumption measures. For this reason, we opted to use two forms of the Gini WIDER. The first form is constructed with different types of income or consumption measures giving priority to disposable income, other forms of income, and consumption in this order.[12] This allows us to cover all country/year data points available in the World and European Values surveys. The second form is the Gini WIDER estimated using disposable income only. This restricts the usable sample to two-thirds of the original size but provides more precise estimates for income inequality. Given that we use two different samples for the Gini WIDER, we also present results for the Gini WVS for both samples. Thus, all tables will report four sets of results (two Ginis for each of the two samples).

Relative income (R) is measured as income divided by mean income within each country and year. In happiness research this variable is often used in

[10]Note that for countries that changed currency during the period considered (adoption of the Euro or USD or introduction of a new local currency) the IMF data use only the latest currency. This meant that we had to transform first the income values from our database into the same currencies used by the IMF using the appropriate exchange rates for each currency and each year and only then apply the constant, USD and PPP transformations.

[11]The final conversion sheet is available from the authors on request in Excel or STATA format.

[12]For the selection of the most appropriate Gini, we followed indications provided by Gruen and Klasen (2008). According to tests conducted by these authors on the Gini WIDER database: "Gini coefficients based on expenditures or consumption are significantly lower than based on incomes, and those based on disposable incomes are also significantly lower than those based on gross incomes, particularly in OECD countries" (p. 219).

conjunction with income and/or income inequality to test the importance of the relative income position as opposed to absolute income in explaining satisfaction. Relative income has been found to have a significant and positive effect on satisfaction, but the sign and significance of this variable may be affected by collinearity with other variables such as income or income inequality. It is important therefore to test how the inclusion and exclusion of this variable may alter results for income inequality.

Countries' wealth (W) is measured with GDP per capita for each country and year. As already mentioned, this variable has been extracted from the IMF database and is used in real terms, USD and PPP values.

We also use a number of control variables (C) as follows. A first set of variables measures *individual and family attributes* which are possible predictors of life-satisfaction. These are being unemployed (dummy), sex (female), age (continuous with the addition of age squared), and a dummy for tertiary education and marriage status (dummy where one includes: "*married*" and "*living together as married*"). These are all variables which have been found in the past to explain life satisfaction well.

A second set of variables is used as control variables for *personal values*. This includes the importance attributed by individuals to family and friends (average of these two variables), the importance attributed to work relative to leisure (importance of work/importance of leisure), the importance of politics, and the importance of religion (categorical form).[13] All these variables are measured on a scale from one to four. The original variables assigned to one the value "very important" and to four the value "not important at all." We reversed this order to make the variable increasing in life satisfaction.

A last set of variables measures *trust*. One is individual trust in people which is measured with a dummy variable, where one is "most people can be trusted" and zero is "can't be too careful." A second variable measures individual trust in institutions, also reported as a reversed one to four scale. This variable is the average trust that individuals reported to have *vis-à-vis* a number of institutions including the army, police, justice system, parliament, civil service, press, companies, and trade unions. Trust in people and institutions can be understood as measures of social capital as in Helliwell (2003).[14]

5. Tests

In this section, we propose a systematic approach to test the consistency of the Gini coefficient as a possible predictor of life satisfaction comparing sign and significance of the Gini coefficient across different specifications of the life satisfaction equation and different samples.

[13]Note that it makes little difference whether these last two variables are split into dummies. We re-estimated the first equation of Table 2, splitting importance of religion and importance of politics into dummies. The coefficient of the Gini changed from –0.0288 to –0.0285 and the z-stat from 5.91 to 5.87. The other variables in the equation had similar marginal changes and none of the variables changed sign or significance level.

[14]As noted by one of the referees of this paper, if inequality reduces individuals' trust, then controlling for trust may underestimate the effect of inequality on well-being. This is also an issue related to multicollinearity.

The original database we use is unbalanced, meaning that not all variables are observed in all countries and for all years. This posed a problem when comparing different sets of reduced equations. We therefore opted to balance the sample for all variables we considered, which reduced the number of observations from 263,097 to 95,612 for sample "1" (where the Gini WIDER is constructed on income and consumption measures) and to 66,630 for sample "2" (where the Gini WIDER is constructed with only disposable income). Also important is the fact that the number of country/year points is reduced from 173 to 77 for sample "1" and to 56 points for sample "2." The variance of the Gini and of GDP per capita depends on the number of country/year points present in the data, and changes in this number affect inference.

Table A1 in the Appendix compares means, standard deviations, and maximum and minimum values for all variables in the full sample and the two reduced samples. As can be seen from the table, differences between the full sample and the two reduced samples are small for the dependent and control variables. There are instead noticeable differences for the two Ginis, GDP per capita, and income between sample "1" and sample "2." This is due to the fact that some countries and years are lost with the smaller sample and this has an impact on the mean of key variables, particularly those variables estimated at the country/year level. For the discussion that follows, it is important to keep in mind that sample "2" is a sub-sample of sample "1" and represents a sub-set of countries and years. Table A2 in the Appendix provides details on the samples considered in this paper.

We start by estimating the full model as described in equation (1) and including robust and regional clusters and country and year fixed effects. A robust estimator allows relaxation of the assumption that regressors and error term are identically distributed, whereas the regional cluster option let us relax the assumption that individual observations within regions are independent. Country and year dummies control for country heterogeneity and time dependence. This is what we could call a standard approach when working with a pooled sample of world citizens. It is also the approach followed by Alesina *et al.* (2004) that we said we use as a benchmark for our tests. The exercise is repeated for the Gini WVS and Gini WIDER and for the two samples considered.[15]

Results are shown in Table 1. The coefficients for both Ginis are negative and significant in both samples, indicating that higher income inequality is associated with lower life satisfaction. The fact that this result is consistent across different Ginis and different samples shows that results are robust. Sign and significance concord whether we use the Gini WVS (constructed with mid-class values from the ten-steps income variable contained in the World and European Values surveys) or the Gini WIDER (imported from the UNU–WIDER database). They also concord if we use the Gini WIDER constructed with different income measures or the Gini WIDER constructed only with disposable income and they concord if we use the larger or smaller samples. The standard model provides consistent evidence of a negative association between life satisfaction and income inequality.

[15]Note that we are not trying to replicate the results of Alesina *et al.*; we simply use the same form of equation as also used in other contributions and with different data. Our purpose is to test this general form of equation under different specifications.

TABLE 1

Predictors of Life Satisfaction

		Sample 1				Sample 2			
		Gini WVS		Gini WIDER		Gini WVS		Gini WIDER	
Symbols	Variables	Coeff.	z-stat.	Coeff.	z-stat.	Coeff.	z-stat.	Coeff.	z-stat.
I	Gini	–0.029	(5.91)**	–0.045	(2.81)**	–0.034	(6.04)**	–0.047	(3.23)**
X	Income (000, USD, PPP)	–0.052	(3.02)**	–0.068	(3.99)**	–0.111	(5.50)**	–0.134	(6.55)**
R	Relative income (income/mean income)	0.308	(9.94)**	0.325	(9.74)**	0.418	(11.92)**	0.447	(13.20)**
W	GDP per capita (000, USD, PPP)	–0.173	–0.91	0.073	–0.3	0.361	–1.76	–0.112	–0.38
C	Unemployed	–0.592	(12.27)**	–0.594	(12.58)**	–0.664	(16.15)**	–0.664	(16.18)**
C	Female	0.020	–1.36	0.022	–1.47	0.009	–0.37	0.010	–0.44
C	Age	–0.045	(14.54)**	–0.045	(14.58)**	–0.050	(12.84)**	–0.050	(12.93)**
C	Age squared (/1000)	0.446	(13.57)**	0.445	(13.58)**	0.486	(11.96)**	0.487	(12.01)**
C	Tertiary education	0.128	(5.41)**	0.142	(5.88)**	0.130	(4.40)**	0.145	(4.99)**
C	Married	0.371	(18.93)**	0.374	(19.18)**	0.405	(19.68)**	0.405	(19.68)**
C	Trust in people	0.213	(11.21)**	0.212	(11.16)**	0.249	(12.27)**	0.247	(12.15)**
C	Trust in institutions	0.210	(11.10)**	0.209	(10.82)**	0.270	(8.71)**	0.271	(8.71)**
C	Importance of family and friends	0.306	(14.08)**	0.304	(14.00)**	0.340	(13.99)**	0.339	(13.79)**
C	Importance of work	–0.123	(5.22)**	–0.119	(5.15)**	–0.126	(4.98)**	–0.122	(4.79)**
C	Importance of politics	–0.031	(2.84)**	–0.031	(2.86)**	–0.033	(2.74)**	–0.033	(2.77)**
C	Importance of religion	0.108	(10.12)**	0.110	(10.34)**	0.108	(10.47)**	0.110	(10.77)**
	Collinearity								
	Countries (dropped/total)	3/56		3/56		3/42		4/42	
	Years (dropped/total)	0/10		0/10		1/10		0/42	
	GINI VIF	17.4		79.6		12.9		131.0	

Notes: Ordered logit, robust standard errors, regional clusters, country and year fixed effects.

Units = Individuals. Dep. Var. = Life satisfaction.

Where significant, z-stat is reported in absolute terms and in parentheses. Where the coefficient is rounded up to zero, the original sign of the coefficient can be positive or negative.

*Significant at 5%; **significant at 1%.

Sample 1: 95,612 observations, 77 country/year points; Sample 2: 66,630 observations, 56 country/year points.

It is important to note, however, that the Gini is highly collinear with other independent variables used in the model. This is visible in all four equations considered as indicated by the high levels of the variance inflation factor (VIF) reported at the bottom of Table 1.[16] When we tested for collinearity of the Gini with other variables, we found that this is due to GDP per capita and to most countries and years dummies included into the model. We found a large and significant correlation between the two Ginis and GDP per capita (Pearson correlation coefficient of +0.6 for the Gini WVS and +0.5 for the Gini WIDER) and we also found these correlations to be high, with most country and year dummies retained by the model.

These correlations are not surprising and due to the fact that the Gini, GDP per capita, and country and year dummies are all country and/or year variables and count on a very restricted number of data points as compared to individual variables. This collinearity affects the reliability of the coefficients, often leads the software to drop selected countries and years, and increases in importance with smaller samples. In Table 1, some countries and years have been dropped by the software for multicollinearity, and the share of country and year dummies dropped increases as the number of country/year points decreases (see bottom of Table 1).

The issue of multicollinearity can also be relevant in the standard model for individual variables, where the number of data points is much greater than for the Gini or GDP per capita. Table 1 shows that the coefficient for income is always negative and significant while relative income is always positive and significant. This would suggest that the two variables have opposite effects on life satisfaction. However, the two variables are correlated by construction (Pearson of +0.66 for sample "1" and +0.67 for sample "2") and excluding one of the two variables from the equation changes results for the other variable. For example, when income is used without relative income, this variable is always positive and significant in all models considered in Tables 1–4. In this paper, we are mostly concerned with the Gini and we will test the impact on the Gini of including and removing other income variables from the model. However, in studies on income and life satisfaction, it is recurrent to use as regressors income together with other income related measures such as relative income, income classes, or income rank (see, e.g. Clark, 2003; Schwarze and Harpfer, 2003; Senik, 2004; Ball and Chernova, 2008; Graham and Felton, 2005, 2006) because several welfare theories underline the importance of relative income in addition to absolute income. We find that this practice may pose non-negligible problems in terms of collinearity and interpretation of the coefficients.

Concerning the control variables and as compared with empirical results in previous studies on happiness, our results largely confirm known correlations with life satisfaction.[17] With a positive and significant sign we find age squared, tertiary

[16]The VIF is estimated as 1/(1-R2) from an OLS regression where the dependent variable is the Gini and the independent variables are all other regressors used in the equations. This is perhaps the most popular test for collinearity. A VIF equal to one indicates no collinearity while values higher than one indicate higher degrees of collinearity. Values of five or more are generally considered as indicators of high levels of multicollinearity.

[17]See among others: Wilson (1967), Veenhoven (1996), Diener *et al.* (1997), Clark and Oswald (1994), Blanchflower and Oswald (1997), Winkelmann and Winkelmann 1998, and Alesina *et al.* (2004).

education, being married, trust in people and institutions, the importance of family and friends, and the importance of religion. These are the factors that are associated with increased life satisfaction. Regressors with a negative sign are being unemployed, age, importance of work, and importance of politics. These are the factors associated with lower life satisfaction. All these findings are consistent across the four equations in Table 1, indicating that our standard model replicates previous results well.

In Table 2, we test the consistency of the Gini coefficient by changing the set of regressors (income, relative income, GDP per capita, and controls). As in Table 1, in Table 2 we use robust standard errors, the regional cluster option, and country and year fixed effects. For simplicity, the inclusion or exclusion of the different regressors has been marked with a 1/0 code, where "1" stands for inclusion and "0" stands for exclusion. We also report only the coefficients of the Gini and, as before, we repeat the exercise for the two Ginis and for the two samples.

The two Ginis maintain a negative and significant sign in both samples and with no exceptions. The inclusion or exclusion from the model of other variables that make use of the same income measure on which the Gini WVS is constructed, such as income and relative income, do not alter the sign or significance of the Gini. In all estimations carried out in this paper we find a strong collinearity between income and relative income, and this collinearity changes the sign and significance of these variables when used separately or in conjunction. This phenomenon does not seem to affect the Gini coefficient. In fact, the Pearson correlation coefficient between the Gini WVS and income is significant but small (+0.2 for the Gini WVS and +015 for the Gini WIDER), while the same coefficient between the two Ginis and relative income is non-significant. Similarly, the inclusion or exclusion of GDP per capita does not seem to affect inference on the Gini coefficient with any of the Gini or samples used despite the relevant correlation found between the Gini and GDP per capita.

As in Table 1, in Table 2 the VIF values for the Gini are all very high, especially for the Gini WIDER, and multicollinearity of the Gini persists when we remove other income variables from the model, including GDP per capita. This suggests that the high levels of multicollinearity observed for the Gini are not generated by other income variables present in the model or by control variables. This is an important finding for empirical research. We have some evidence that the Gini can be safely used in conjunction with other income variables and that correlation between income variables does not necessarily lead to fragile inference about the happiness–inequality relation.

In the following exercise we keep all key regressors and all control variables in the equation while we test the Gini coefficient with alternative estimation choices of the life satisfaction equation including and excluding the robust standard error, regional clusters, country, and year fixed effects. Despite the popularity of the standard model, different authors make different choices. Such choices may depend on the particular sample used or on the particular economic model that one has in mind, but all these choices carry a certain amount of uncertainty about the underlying assumptions that justify the choice. We expected a strong predictor of life satisfaction to be consistent irrespective of the choice made, and testing results under different choices can be regarded as a validation exercise.

TABLE 2

Test Gini with Alternative Regressors

					Sample 1						Sample 2					
					Gini WVS			Gini WIDER			Gini WVS			Gini WIDER		
Eq.	X Inc.	R Rel.Inc.	W GDP/cap	C Controls	coeff.	z-stat	vif	coeff.	z-stat	vif	coeff.	z-stat	vif	coeff.	z-stat	vif
1	0	0	0	0	–0.028	(4.54)**	16.8	–0.045	(3.38)**	78.2	–0.041	(7.65)**	11.5	–0.049	(3.57)**	129.8
2	1	0	0	0	–0.034	(4.60)**	16.9	–0.048	(3.36)**	78.2	–0.051	(8.42)**	11.7	–0.052	(3.41)**	129.9
3	0	1	0	0	–0.028	(4.53)**	16.8	–0.042	(2.80)**	78.2	–0.041	(7.51)**	11.5	–0.05	(3.73)**	129.8
4	0	0	1	0	–0.031	(6.34)**	16.8	–0.046	(3.09)**	78.7	–0.042	(7.81)**	12.2	–0.057	(3.39)**	130.4
5	0	0	0	1	–0.03	(4.95)**	16.8	–0.052	(4.11)**	78.7	–0.038	(6.79)**	11.6	–0.046	(3.48)**	130.3
6	1	1	0	0	–0.027	(4.36)**	17.3	–0.041	(2.70)**	78.4	–0.036	(6.97)**	12.3	–0.048	(3.93)**	130.1
7	0	0	1	1	–0.032	(6.54)**	16.9	–0.052	(3.79)**	79.2	–0.039	(7.05)**	12.2	–0.052	(3.19)**	130.9
8	0	1	1	0	–0.031	(6.34)**	16.8	–0.043	(2.53)*	78.8	–0.042	(7.60)**	12.2	–0.058	(3.50)**	130.4
9	1	0	0	1	–0.034	(4.95)**	17.0	–0.053	(3.88)**	78.7	–0.046	(7.49)**	11.8	–0.048	(3.36)**	130.4
10	1	0	1	0	–0.039	(6.86)**	17.0	–0.05	(3.10)**	78.7	–0.052	(8.69)**	12.3	–0.064	(3.36)**	130.4
11	0	1	0	1	–0.029	(4.95)**	16.8	–0.048	(3.25)**	78.9	–0.038	(6.77)**	11.6	–0.047	(3.57)**	130.3
12	1	1	1	0	–0.029	(5.98)**	17.3	–0.041	(2.43)*	78.9	–0.037	(6.88)**	12.9	–0.054	(3.56)**	130.5
13	0	1	1	1	–0.031	(6.52)**	16.9	–0.048	(2.94)**	79.5	–0.04	(6.94)**	12.2	–0.052	(3.18)**	130.9
14	1	0	1	1	–0.037	(6.87)**	17.0	–0.054	(3.58)**	79.3	–0.047	(7.70)**	12.4	–0.057	(3.13)**	130.9
15	1	1	0	1	–0.027	(4.65)**	17.4	–0.046	(3.13)**	79.1	–0.031	(5.93)**	12.4	–0.045	(3.76)**	130.6
16	1	1	1	1	–0.029	(5.91)**	17.4	–0.045	(2.81)**	79.6	–0.034	(6.04)**	12.9	–0.047	(3.23)**	131.0

Notes: Each Gini coefficient in the table is estimated with a different equation and set of regressors. The regressors are indicated with "1" if included and with "0" otherwise. All equations are estimated with ordered logit, robust standard errors, regional clusters, country and year fixed effects.

Units = Individuals. Dep. Var. = Life satisfaction.

Where significant, z-stat is reported in absolute terms and in parentheses. Where the coefficient is rounded up to zero, the original sign of the coefficient can be positive or negative.

*Significant at 5%; **significant at 1%.

Sample 1: 95,612 observations, 77 country/year points; Sample 2: 66,630 observations, 56 country/year points.

In Table 3, we find no consistency in sign or significance of the coefficient for both Ginis and in both samples. Both Ginis can be negative, positive, significant, and non-significant with either sample "1" or sample "2."

The robust and cluster choices can make a difference to inference about the Gini, especially if the sample is small and the number of countries and years is reduced. Different choices do not have an impact on the sign or size of the coefficients but can have an impact on standard errors and significance levels. In Table 3 we see that when the robust and cluster estimations are introduced (equations 2 and 3), the z-statistics can visibly change. This is particularly true for the Gini WIDER and for the smaller sample that considers a restricted number of countries and years (sample "2").

More importantly, introducing or removing country and year fixed effects can alter inference on inequality remarkably. When country fixed effects or both country and year fixed effects are introduced, all Ginis are negative and significant (equations 4, 7, 8, 10, 12, 13, 15, and 16). This is what we found in Tables 1 and 2 where we used country and year fixed effects for all equations. When country and year fixed effects are removed or year fixed effects are used alone, the Gini turns positive and significant or non-significant (equations 1, 2, 3, 5, 6, 9, 11, and 14). In particular, it would seem that country fixed effects have an important influence on multicollinearity and significance levels while year fixed effects have a relevant role in changing the sign of the Gini. Indeed, the VIF values for the Gini are small only when country fixed effects are removed and the Gini turns positive only when year fixed effects are used alone. This phenomenon applies equally to both Ginis and both samples considered, indicating that this is not a phenomenon dependent on the choice of these factors.

The sensitivity of the Gini coefficient to country and year fixed effects may relate to various factors. It may be, for example, that there is moderate within-countries variation of the Gini over time or, if there is variation, time trends are similar across countries. The World Values survey is characterized by many countries, few years for each country (from one to four years depending on the country), and several years in between any two consecutive observations within countries. Changes in the Gini can be significant but, among the countries with more than one observation, about half have decreasing Ginis, one has a Gini that goes up and down, and the rest of the countries have increasing Ginis. Therefore, the Ginis do not move together over time across countries, whereas the within countries variation is limited by the number of years available for each country.

Having more data points for the Gini may help to better capture the relation between happiness and inequality, but this is not always the case. For example, Senik (2004) finds non-significant coefficients for the Gini when she estimates the Gini at the national, regional, or PSU levels. Also, in a previous version of the paper, we estimated the Gini WVS at the regional level and compared the coefficient of this variable with that of the Gini estimated at the country level. We found that the Gini region was even more sensitive than the Gini country to the use of country and year fixed effects.

In substance, there is a trade-off between the inclusion of country and year dummies, which allows us to control for unobserved factors but generates collinearity, and the exclusion of these dummies, which fixes collinearity but increases

TABLE 3

Test Gini with Alternative Options

	Standard Error		Fixed Effects		Sample 1						Sample 2					
					Gini WVS			Gini WIDER			Gini WVS			Gini WIDER		
Eq.	Robust	Cluster	Country	Year	coeff.	z-stat	vif	coeff.	z-stat	vif	coeff.	z-stat	vif	coeff.	z-stat	vif
1	0	0	0	0	0.000	–0.49	1.7	0.016	(26.17)**	1.5	0.002	–1.73	1.6	0.011	(12.42)**	1.6
2	1	0	0	0	0.000	–0.47	1.7	0.016	(24.66)**	1.5	0.002	–1.64	1.6	0.011	(11.82)**	1.6
3	0	1	0	0	0.000	–0.09	1.7	0.016	(3.40)**	1.4	0.002	–0.24	1.6	0.011	–1.66	1.5
4	0	0	1	0	–0.035	(20.97)**	7.3	–0.037	(9.69)**	60.0	–0.030	(10.57)**	10.6	–0.030	(4.74)**	76.7
5	0	0	0	1	0.010	(10.80)**	2.1	0.011	(15.31)**	1.9	0.009	(7.11)**	2.0	0.013	(11.27)**	2.5
6	1	1	0	0	0.000	–0.09	1.7	0.016	(3.40)**	1.4	0.002	–0.24	1.6	0.011	–1.66	1.5
7	0	0	1	1	–0.029	(13.19)**	12.5	–0.045	(11.00)**	67.8	–0.034	(10.09)**	14.0	–0.047	(6.12)**	114.0
8	0	1	1	0	–0.035	(4.38)**	7.7	–0.037	(2.84)**	76.6	–0.030	(5.69)**	9.8	–0.030	(3.57)**	108.1
9	1	0	0	1	0.010	(10.37)**	2.1	0.011	(14.19)**	1.9	0.009	(6.69)**	2.0	0.013	(10.90)**	2.5
10	1	0	1	0	–0.035	(20.62)**	7.3	–0.037	(9.71)**	60.0	–0.030	(9.81)**	10.6	–0.030	(4.83)**	76.7
11	0	1	0	1	0.010	(2.67)**	2.2	0.011	(2.83)**	1.9	0.009	–1.65	2.0	0.013	(2.63)**	2.6
12	0	1	1	1	–0.029	(5.91)**	17.4	–0.045	(2.81)**	79.6	–0.034	(6.04)**	12.9	–0.047	(3.23)**	131.0
13	1	0	1	1	–0.029	(12.60)**	12.5	–0.045	(11.14)**	67.8	–0.034	(9.07)**	14.0	–0.047	(6.24)**	114.0
14	1	1	0	1	0.010	(2.67)**	2.2	0.011	(2.83)**	1.9	0.009	–1.65	2.0	0.013	(2.63)**	2.6
15	1	1	1	0	–0.035	(4.38)**	7.7	–0.037	(2.84)**	76.6	–0.030	(5.69)**	9.8	–0.030	(3.57)**	108.1
16	1	1	1	1	–0.029	(5.91)**	17.4	–0.045	(2.81)**	79.6	–0.034	(6.04)**	12.9	–0.047	(3.23)**	131.0

Notes: Each Gini coefficient in the table is estimated with a different equation and set of options. The options are indicated with “1” if included and with “0” otherwise. All equations are estimated with ordered logit and include the full set of key variables and controls used in Table 1.

Units = Individuals. Dep. Var. = Life satisfaction.

Where significant, z-stat is reported in absolute terms and in parentheses. Where the coefficient is rounded up to zero, the original sign of the coefficient can be positive or negative.

*Significant at 5%; **significant at 1%.

Sample 1: 95,612 observations, 77 country/year points; Sample 2: 66,630 observations, 56 country/year points.

the problem of unobserved heterogeneity. Moreover, with smaller samples, increased standard errors can be generated by both the use of robust and cluster estimations and by increased collinearity between the Gini and country and year fixed effects. This combination of factors can make inference on inequality very fragile, and the use of data with very different structures such as cross-country, longitudinal, or panel data can lead to different results because the structure of the data can tip the balance of the Gini coefficient towards negative or positive values. These factors help to explain the existing heterogeneity in empirical results.[18]

One alternative hypothesis that we put forward in previous sections is that people located in different parts of the income distribution may have a different appreciation of inequality. It seemed therefore important to test alternative specifications, dividing observations into income groups. For this purpose, we split the sample into rich and poor individuals using as a poverty line median income within each country/year point.

We also split the sample into Western and non-Western nations, dividing in this way rich and poor countries and also countries that may differ in state institutions. It is entirely possible that people living in countries at different levels of economic and institutional development may have a different appreciation of inequality, which is another important question to address. Evidently, poor and rich individuals and poor and rich nations are not overlapping definitions. Poor and rich individuals are defined within each country and year and relative to median income, whereas poor and rich nations are split according to national wealth (an individual may be poor but live in a rich nation).[19]

Table 4 shows the results.[20] There is no difference in sign between poor and non-poor people and between Western and non-Western nations for both Ginis and both samples considered. With the gross distinctions we made in terms of income and countries, we seem to find that higher income inequality is invariably associated with lower life satisfaction.

However, this relation is not always significant. In sample "1" the Gini WIDER is non-significant for poor individuals and non-Western (poorer) countries, although it becomes significant for poor individuals with the use of the more precise Gini WIDER in sample "2." In the smaller sample "2" the Gini WVS becomes non-significant for poor individuals and Western countries.

As we discussed in Section 2 of this paper, there are various reasons why the poor and the non-poor may or may not be inequality averse. The consistent negative sign that we find for poor and non-poor alike indicates that both groups are inequality averse but does not tell us anything about the reasons that may explain such aversion. It is natural for the poor to be inequality averse because lower inequality could imply better distribution of resources and improved

[18]Note that when we tested whether the control variables used in Table 3 could be positively correlated with both life satisfaction and the Gini, we found that only one variable (trust in institutions) was positively correlated with life satisfaction and both Ginis, while one variable was positively correlated with life satisfaction and the Gini WIDER (importance of politics).

[19]Note that splitting the sample into smaller income classes or greater regional detail made the samples too small.

[20]As in Table 1, we estimated the model with the two Ginis and the two samples and included all regressors (key and controls) and all estimation choices (robust standard errors, regional cluster, country and year fixed effects). For simplicity, only the coefficients of the Gini are shown in the table.

TABLE 4

Test Gini with Alternative Sub-Samples

	Gini WVS					Gini WIDER					Observations			
	coeff.	z-stat	vif	Dropped Countries	Dropped Years	coeff.	z-stat	vif	Dropped Countries	Dropped Years	Individual	Countries	Years	Country/ Year
Sample 1														
Poor individuals	–0.023	(4.27)**	18.2	2	1	–0.019	–1.35	79.4	2	1	56451	56	10	77
Non-poor individuals	–0.031	(4.24)**	20.1	3	0	–0.086	(3.32)**	95.1	2	1	39161	56	10	77
Western countries	–0.035	(5.59)**	9.4	0	2	–0.057	(3.44)**	15.8	0	2	39641	22	7	35
Non-Western countries	–0.016	(2.46)*	13.0	3	1	–0.023	–1.73	122.0	3	1	55971	34	9	42
Sample 2														
Poor individuals	–0.01	–0.93	13.8	1	3	–0.016	(2.21)*	137.4	1	3	38685	42	10	56
Non-poor individuals	–0.039	(7.24)**	15.3	1	3	–0.074	(2.85)**	124.6	1	3	27945	42	10	56
Western countries	–0.003	–0.22	15.1	1	1	–0.006	–0.42	13.2	0	2	36667	22	6	32
Non-Western countries	–0.021	(4.99)**	na	4	1	–0.091	(4.99)**	na	4	1	29963	20	8	24

Notes: The model used in this table is the same as in Table 1 with ordered logit, robust standard errors, regional clusters, country and year fixed effects. Only the coefficients of the Gini are reported.

Units = Individuals. Dep. Var. = Life satisfaction.

Where significant, z-stat is reported in absolute terms and in parentheses. Where the coefficient is rounded up to zero, the original sign of the coefficient can be positive or negative.

*Significant at 5%; **significant at 1%.

Sample 1: 95,612 observations, 77 country/year points; Sample 2: 66,630 observations, 56 country/year points.

welfare, but Hirschman and Rothschild (1973) suggested that even the poor could favor inequality. On the other hand, inequality aversion of the non-poor is less intuitive although scholars across the social sciences have sometimes explained such aversion with sentiments of guilt, regret, or compassion or with a preference for more stable and less conflictual societies. This paper did not investigate the motives that may explain such attitudes on the part of the poor and non-poor, but our findings clearly speak in favor of Runciman's view that more inequality generates a greater sense of dissatisfaction.

Despite the consistency of the negative sign in Table 4, inference on inequality is less robust than in Table 1 where we used the same standard model. The difference in Table 4 is that we use reduced sample sizes, having split the sample into different groups. This reduces the number of observations and the number of countries, years, and country/year points available, and increases the likelihood of multicollinearity between the Gini and other variables.

When multicollinearity with countries and years dummies increases, the software is also more likely to drop some of these dummies. If we compare the number of countries and years dummies dropped by the software with the total number of countries and years available for each equation in Table 4, we can see that the share of countries and years dummies dropped by the software is larger for smaller samples. For example, in the two equations on poor and non-poor individuals in sample "2," the software drops three of the ten years dummies because of multicollinearity. It is evident that by excluding a third of the years fixed effects we are estimating a different model and we could reach rather different conclusions on the Gini. This is an issue hardly discussed in the empirical literature.

In conclusion, the central issue in studying the happiness–inequality relation is the interplay between multicollinearity, data structure, and sample size. The combinations of different sets of key regressors such as income and relative income does not affect the sign and significance of the inequality measure, but may generate collinearity between income and other individual measures constructed with income. Robust and cluster estimations do not have an impact on sign and size of coefficients but can contribute to change significance levels, especially in small samples. More importantly, the inclusion and exclusion of country and year fixed effects represents a real trade-off between addressing issues of collinearity and issues of unobserved heterogeneity, and the cost of this trade-off can change with different data structures and sample sizes.

6. Conclusion

Both theory and empirics can provide alternative views on how income inequality may affect subjective well-being. We discussed how, for some scholars, an increase in inequality may lead to improved happiness while, for other scholars, an increase in inequality should lead to decreased happiness. We have also discussed and shown in Table A3 how empirical contributions have reached rather different conclusions about the covariance of happiness and inequality. Some papers find a positive association, some a negative association, and others no association at all.

We put forward a number of hypotheses that could explain the existing empirical heterogeneity, and tested these hypotheses one by one making use of a standard happiness model and of a large sample of world citizens. These hypotheses relate to the choice of Gini, tastes for inequality across population subgroups, choice of key regressors, use of country and year fixed effects, number of data points available, and estimation of the standard error.

Overall, we found income inequality to have a consistent, negative, and significant effect on life satisfaction worldwide when a standard happiness model is used. However, this relation can be sensitive to different factors. The use of Ginis estimated from within the sample used or imported from other datasets can make a difference in estimating coefficients, although sign and significance of the happiness–inequality relation is preserved (Table 1). The choice of key regressors can be important if we use variables that use the same income variable used to estimate the Gini, such as income or relative income. However, we found the sign and significance of the Gini to be robust to such changes (Table 2). The use of subsamples, such as poor and rich individuals or Western and non-Western countries, also preserves the negative and significant sign of the happiness–inequality relation although this relation is more difficult to detect as we use smaller samples of countries and years (Table 4).

Instead, we found very high levels of collinearity between inequality and country and year fixed effects, and we found this multicollinearity to have the potential to change size, sign, and significance of the happiness–inequality relation (Table 3). We argued that a real trade-off between addressing issues of multicollinearity and issues of unobserved heterogeneity may exist. In particular, such collinearity may be more or less relevant, depending on how the standard error is estimated and the structure of the dataset. Robust and cluster estimations of the standard error can make the happiness–inequality relation more difficult to detect, particularly when country and year fixed effects are used. And similar specifications of the happiness equation that use different datasets, number of countries, years, or observations can lead to different results because the role of multicollinearity can change when the structure of the data changes.

These last factors are the most likely to explain the heterogeneity found in empirical studies. All studies considered used different datasets. Some studies worked cross-country, others cross-country and longitudinally, and others are panel studies. Not all studies use robust and cluster options, and when cluster options are used these can be at different levels (country, region, or smaller units). Most studies use country and/or year fixed effects, but the collinearity that these fixed effects can generate with the Gini can be very different depending on the structure of the data and on the estimation procedure for the standard error.

In order to compare results across studies, readers should have full information on the number of countries and years, the number of observations within each country/year data point, the exact procedure used for the estimation of the standard error, the use of country, year, or other fixed effects, the number of country or year dummies dropped by the software during estimations, and more generally the full estimation model. When some of this information is missing, it becomes very hard to replicate and compare results.

Appendix

TABLE A1

Full and Reduced Samples Compared

Variable	Observations			Mean			Standard Deviation			Minimum			Maximum		
	Full Sample	Sample 1	Sample 2	Full Sample	Sample 1	Sample 2	Full Sample	Sample 1	Sample 2	Full Sample	Sample 1	Sample 2	Full Sample	Sample 1	Sample 2
Life satisfaction	263097	95612	66630	6.62	6.78	6.94	2.49	2.44	2.34	1.00	1.00	1.00	10.00	10.00	10.00
Unemployed	267870	95612	66630	0.08	0.08	0.07	0.27	0.27	0.25	0.00	0.00	0.00	1.00	1.00	1.00
Female	267870	95612	66630	0.52	0.51	0.52	0.50	0.50	0.50	0.00	0.00	0.00	1.00	1.00	1.00
Age	264839	95612	66630	41.24	41.77	43.34	16.33	16.09	16.54	15.00	16.00	16.00	101.00	99.00	98.00
Tertiary education	267870	95612	66630	0.15	0.17	0.15	0.35	0.38	0.35	0.00	0.00	0.00	1.00	1.00	1.00
Married	267870	95612	66630	0.63	0.64	0.63	0.48	0.48	0.48	0.00	0.00	0.00	1.00	1.00	1.00
Trust in people	267870	95612	66630	0.28	0.28	0.30	0.45	0.45	0.46	0.00	0.00	0.00	1.00	1.00	1.00
Trust in institutions	260301	95612	66630	2.42	2.39	2.34	0.59	0.57	0.54	1.00	1.00	1.00	4.00	4.00	4.00
Importance of family and friends	238856	95612	66630	3.56	3.56	3.58	0.45	0.45	0.44	1.00	1.00	1.00	4.00	4.00	4.00
Importance of work	233484	95612	66630	1.28	1.27	1.20	0.61	0.58	0.51	0.25	0.25	0.25	4.00	4.00	4.00
Importance of politics	234025	95612	66630	2.27	2.24	2.21	0.96	0.96	0.94	1.00	1.00	1.00	4.00	4.00	4.00
Importance of religion	234563	95612	66630	2.90	2.90	2.71	1.08	1.07	1.07	1.00	1.00	1.00	4.00	4.00	4.00
Gini WVS	na	95612	66630	na	37.85	35.53	na	9.34	7.98	na	23.67	23.67	na	63.82	61.84
Gini Wider	na	95612	66630	na	37.75	34.60	na	11.35	9.50	na	21.45	21.45	na	73.20	59.50
Income (000, USD, PPP)	na	95612	66630	na	1.34	1.47	na	1.42	1.38	na	0.00	0.01	na	36.49	21.46
Relative Income	na	95612	66630	na	1.01	1.01	na	0.83	0.76	na	0.01	0.01	na	16.50	13.19
GDP capita (000, USD, PPP)	na	95612	66630	na	1.03	1.25	na	0.69	0.63	na	0.02	0.15	na	2.77	2.41

TABLE A2
Number of Observations by Country, Year and Sample

No.	Country/Year	Sample 1	Sample 2	Poor	Non-Poor	Non-Western	Western
1	albania2002	947		317	630	947	
2	algeria2002	963		267	696	963	
3	argentina1999	1,220		494	726	1,220	
4	austria1990	1,326	1,326	545	781		1,326
5	austria1999	1,185	1,185	553	632		1,185
6	belgium1990	1,613	1,613	739	874		1,613
7	belgium1999	1,473	1,473	703	770		1,473
8	bosnia and herzegovina2001	1,118		525	593	1,118	
9	bulgaria1999	847	847	386	461	847	
10	canada1990	1,441	1,441	668	773		1,441
11	canada2000	1,688	1,688	692	996		1,688
12	chile1990	1,424	1,424	637	787	1,424	
13	chile1996	895	895	421	474	895	
14	chile2000	1,096	1,096	491	605	1,096	
15	china2001	831		371	460	831	
16	colombia1998	2,960		1,000	1,960	2,960	
17	croatia1999	904	904	373	531	904	
18	czech republic1991	1,944	1,944	874	1,070	1,944	
19	czech republic1999	1,670	1,670	699	971	1,670	
20	denmark1999	796	796	361	435		796
21	egypt2000	2,597		1,017	1,580	2,597	
22	el salvador1999	975	975	462	513	975	
23	estonia1999	818	818	345	473	818	
24	finland1990	555	555	177	378		555
25	france1999	1,265	1,265	528	737		1,265
26	germany1999	1,490	1,490	423	1,067		1,490
27	great britain1990	1,053	1,053	487	566		1,053
28	greece1999	910	910	292	618		910
29	hungary1991	951	951	346	605	951	
30	iceland1999	884	884	390	494		884
31	india1990	2,323		805	1,518	2,323	
32	india2001	1,721		730	991	1,721	
33	ireland1990	880	880	427	453		880
34	ireland1999	812	812	291	521		812
35	italy1990	1,391	1,391	652	739		1,391
36	italy1999	1,465	1,465	646	819		1,465
37	japan1990	687		321	366		687
38	japan2000	987	987	407	580		987
39	jordan2001	1,081		506	575	1,081	
40	latvia1999	888	888	271	617	888	
41	lithuania1999	745	745	363	382	745	
42	macedonia, republic of2001	998	998	431	567	998	
43	malta1999	696	696	339	357		696
44	mexico1990	1,367	1,367	475	892	1,367	
45	mexico2000	1,153	1,153	430	723	1,153	
46	morocco2001	1,247		566	681	1,247	
47	netherlands1990	782	782	323	459		782
48	netherlands1999	928	928	457	471		928
49	new zealand1998	955	955	457	498		955
50	peru1996	919	919	342	577	919	
51	peru2001	1,455		528	927	1,455	
52	portugal1990	1,055	1,055	351	704		1,055
53	portugal1999	653	653	233	420		653
54	republic of korea1990	1,147	1,147	268	879		1,147
55	republic of korea2001	1,167		414	753		1,167
56	republic of moldova2002	783	783	288	495	783	

TABLE A2 (*continued*)

No.	Country/Year	Sample 1	Sample 2	Poor	Non-Poor	Non-Western	Western
57	russian federation1999	2,130	2,130	964	1,166	2,130	
58	serbia and montenegro2001	1,744	1,744	832	912	1,744	
59	slovakia1999	1,175	1,175	447	728	1,175	
60	slovenia1999	641	641	314	327	641	
61	south africa1990	1,870		545	1,325	1,870	
62	south africa1996	1,485		602	883	1,485	
63	south africa2001	2,239		973	1,266	2,239	
64	spain1990	3,279	3,279	1,390	1,889		3,279
65	spain1995	849	849	253	596		849
66	spain1999	775	775	281	494		775
67	spain2000	839	839	413	426		839
68	sweden1999	956	956	391	565		956
69	switzerland1996	919	919	416	503		919
70	taiwan province of china1994	670	670	295	375	670	
71	turkey2001	4,228	4,228	1,653	2,575	4,228	
72	uganda2001	439		202	237	439	
73	united states1990	1,620	1,620	796	824		1,620
74	united states1999	1,120		437	683		1,120
75	uruguay1996	898		322	576	898	
76	venezuela2000	998	998	457	541	998	
77	zimbabwe2001	614		274	340	614	
	TOTAL	95,612	66,630	39,161	56,451	55,971	39,641

TABLE A3

Summary of Data, Equation Specifications, and Results for Selected Studies on Happiness and Inequality

Study	Data	Model	Happiness Variable	Inequality Variable	Happiness–Inequality Relation	Other Income Variables	Country/ Region fe	Year/ Wave fe	Robust	Cluster
Morawetz *et al.*, 1977	Ad-hoc questionnaire in two villages in Israel	na	Happiness and Life satisfaction	Comparison of an equal and unequal society	Society with lower income inequality has higher happiness	Absolute income, relative income	na	no	no	no
Hagerty, 2000	US-GSS 1989–1996	OLS	Happiness	Max and min income, skewness, 20th and 80th percentiles	Max negative and signif, skew positive and signif. Min non signif. 20th pc positive and signif., 80th pc negative and signif. (Table 2)	Household income category	no	no	na	na
Hagerty, 2000	8 countries study 1972–1994	OLS	Life satisfaction	Gini	Positive and significant in 6 countries, negative and significant in 1 country and non significant in 1 country (Table 4)	GDP per capita	no	no	na	na
Schwarze and Harpfer, 2003	GSOEP 1985–1998	Ordered Probit and OLS	Life satisfaction	Gini (regional)	Gini negative and significant in all equations (ordered logit, pooled, fe, Table 4) Effect explained by 1st, 2nd and 5th quintile (Table 5)	Disposable income (log), disposable income position (quintiles), pregovernment income (log), public transfers, income taxes (percent), payroll taxes (percent).	yes (region)	yes	yes	yes
Clark, 2003	BHPS 1991–2002 (Employed)	Ordered probit, conditional fe logit and re probit	GHQ-12 and Life satisfaction	Gini (reference group)	Gini in GHQ and Lifesat equations positive and significant also for subgroups with ordered probit (Table 1). Gini positive and signif. only in lifesat equation and not for all subroups or estimation model in panel equations (Table 2)	Income and income ref.group	yes (region)	yes	na	yes
Helliwell, 2003	WVS (1980–1982,1990–1991, and 1995–1997 waves)	Oredered probit	Life satisfaction	Gini (World Bank)	Gini non-significant (results not in tables but quoted in text)	Income, relative income	yes	yes (waves)	yes	na
Alesina *et al.*, 2004	US-GSS 1972–1997	Ordered logit	Happiness	Gini (Wu *et al.*, 2002)	Gini negative and significant in 6 of 13 of equations (Tables 1–3 US)	Household Income scale	yes	yes	yes	na
Alesina *et al.*, 2004	EU Eurobarometer 1975–1992	Ordered logit	Happiness	Gini (Deininger and Squire, 1996 for EU)	Gini negative and significant in 7 of 13 equations (Tables 1–3 EU).	Household Income scale	yes	yes	yes	na
Senik, 2004	RLMS 1994–2000	Ordered probit	Life satisfaction	Gini and Stark indices of income overhang (national, regional, and PSU level)	Gini non-significant at all levels. StarkH non-significant at all levels. StarkL positive and significant only at PSU level (Text and Table 11).	Lagged individual income, household income	no	yes	yes	yes
Graham and Felton, 2006	Latinobarometro 2004, 17 LAM countries	Ordered logit	Life satisfaction	Gini	Gini non-significant (description of results in text but not in tables).	Wealth and aver. country wealth	yes	no	na	yes

References

Alesina, A., R. Di Tella, and R. MacCulloch, "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics*, 88, 2009–42, 2004.

Amiel, Y. and F. A. Cowell, "Measurement of Income Inequality: Experimental Test by Questionnaire," *Journal of Public Economics*, 47, 3–26, 1992.

Ball, R. and K. Chernova, "Absolute Income, Relative Income, and Happiness," *Social Indicators Research*, 88, 497–529, 2008.

Blanchflower, D. G. and A. J. Oswald, "A Study of Labour Market and Youth Unemployment in Eastern Europe," *Warwick Economic Research Papers*, 499, 1997.

———, "Well-Being Over Time in Britain and the USA," *Journal of Public Economics*, 88, 1359–86, 2004.

Clark, A. E., "Inequality Aversion and Income Mobility: A Direct Test," *Delta Working Papers*, 11, 2003.

Clark, A. E. and A. J. Oswald, "Unhappiness and Unemployment," *Economic Journal*, 104, 648–59, 1994.

Deininger, K. and L. Squire, "A New Data Set Measuring Income Inequality," *World Bank Economic Review*, 10, 565–91, 1996.

Diener, E., M. Diener, and C. Diener, "Factors Predicting the Subjective Well-Being of Nations," *Journal of Personality and Social Psychology*, 69, 851–64, 1995.

Diener, E., E. Suh, and S. Oishi, "Recent Findings on Subjective Well-Being," *Indian Journal of Clinical Psychology*, 24, 25–41, 1997.

Diener, E., E. M. Suh, R. E. Lucas, and H. L. Smith, "Subjective Well-Being: Three Decades of Progress," *Psychological Bulletin*, 125, 276–303, 1999.

Di Tella, R., R. J. MacCulloch, and A. J. Oswald, "Preferences over Inflation and Unemployment: Evidence from Surveys on Happiness," *American Economic Review*, 91, 335–41, 2001.

Easterlin, R. A., "Does Economic Growth Improve the Human Lot?" in P. A. David and M. W. Reder (eds), *Nations and Households in Economic Growth, Essays in Honour of Moses Abramovitz*, Academic Press, New York, 1974.

———, "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organizations*, 27, 35–47, 1995.

———, "Income and Happiness, Towards a Unified Theory," *Economic Journal*, 111, 465–84, 2001.

Fordyce, M. A., "A Review of Research on Happiness Measures: A Sixty Second Index of Happiness and Mental Health," *Social Indicators Research*, 20, 355–81, 1988.

Graham, C. and A. Felton, "Does Inequality Matter to Individual Welfare? An Initial Exploration Based on Happiness Surveys from Latin America," *Center on Social and Economic Dynamics Working Papers No. 38*, The Brookings Institution, 2005.

———, "Inequality and Happiness: Insights from Latin America," *Journal of Economic Inequality*, 4, 107–22, 2006.

Greene, W. H., *Econometric Analysis*, Prentice Hall, London, 1997.

Gruen, C. and S. Klasen, "Growth, Inequality, and Welfare, Comparisons Across Space and Time," *Oxford Economic Papers*, 60, 212–36, 2008.

Hagerty, M. R., "Social Comparisons of Income in One's Community: Evidence from National Surveys of Income and Happiness," *Journal of Personality and Social Psychology*, 78, 764–71, 2000.

Helliwell, J. F., "How's Life? Combining Individual and National Variables to Explain Subjective Well-Being," *Economic Modelling*, 20, 331–60, 2003.

Hirschman, A. and M. Rothschild, "The Changing Tolerance for Income Inequality in the Course of Economic Development," *Quarterly Journal of Economics*, 87, 544–66, 1973.

Hopkins, E., "Inequality, Happiness and Relative Concerns, What Actually is Their Relationship?" *Journal of Economic Inequality*, 6, 351–72, 2008.

Inglehart, R. F., *Culture Shift in Advanced Industrial Society*, Princeton University Press, 1990.

Lepper, H. S., "Use of Other Reports to Validate Subjective Well-Being Measures," *Social Indicators Research*, 44, 367–79, 1998.

Mangahas, M., "Self-Rated Poverty in the Philippines, 1981–1992," *International Journal of Public Opinion Research*, 7, 40–55, 1995.

Morawetz, D., E. Atia, G. Bin-Nun, L. Felous, Y. Gariplerden, E. Harris, S. Soustiel, G. Tombros, and Y. Zarfaty, "Income Distribution and Self-Rated Happiness: Some Empirical Evidence," *Economic Journal*, 87, 511–22, 1977.

Ravallion, M. and M. Lokshin, "Identifying Welfare Effects from Subjective Questions," *Economica*, 68, 335–57, 2000.

Runciman, W. G., *Relative Deprivation and Social Justice*, Reports of the Institute of Community Studies, Routledge and Kegan Paul, London, Boston, and Henley, 1966.

Sandvik, E., E. Diener, and L. Seidlitz, "Subjective Well-Being: The Convergence and Stability of Self-Report and Non Self-Report Measures," *Journal of Personality*, 61, 317–42, 1993.
Saris, W. E., A. C. Scherpenzeel, R. Veenhoven, and B. Bunting, *A Comparative Study of Satisfaction with Life in Europe*, Eotvos University Press, Budapest, 1996.
Schwarze, J. and M. Harpfer, "Are People Inequality Averse, and Do They Prefer Redistribution by the State? A Revised Version," *IZA Discussion Paper, No. 974*, 2003.
Senik, C., "When Information Dominates Comparison. Learning from Russian Subjective Panel Data," *Journal of Public Economics*, 88, 2099–133, 2004.
Truglia, R. N. P., "Can a Rise in Income Inequality Improve Welfare?" Mimeo, Universidad de San Andrés, Victoria, Argentina, 2007.
Veenhoven, R., *Happiness in Nations:, Subjective Appreciation of Life in 56 Nations 1946–1992*, Erasmus University Press, Rotterdam, 1993.
———, "Happy Life Expectancy. A Comprehensive Measure of Quality-of-Life in Nations," *Social Indicators Research*, 39, 1–58, 1996.
Wilson, W., "Correlates of Avowed Happiness," *Psychological Bulletin*, 67, 294–306, 1967.
Winkelmann, L. and R. Winkelmann, "Why Are the Unemployed So Unhappy? Evidence from Panel Data," *Economica*, 65(257), 1–15, 1998.
Wu, X., J. Perloff, and A. Golan, "Effects of Government Policies on Income Distribution and Welfare," University of California–Berkeley Working Paper, iirwps-086-02, 2002.
Xu, K., "How Has the Literature on the Gini's Index Evolved in the Past 80 Years?" Working Paper, Dalhousie University, 2004.
Yitzhaki, S., "Relative Deprivation and the Gini Coefficient," *Quarterly Journal of Economics*, 93, 321–4, 1979.