National Happiness and Genetic Distance: A Cautious Exploration

Revised for the Economic Journal

Eugenio Proto[a,b], Andrew J. Oswald[*a,b]

28 November 2015

[a]Department of Economics, and CAGE Centre, University of Warwick, Coventry,

CV4 7AL, United Kingdom.

[b]IZA Institute, Schaumburg-Lippe Str., 53072 Bonn, Germany

[*]Corresponding author.

Phone: +44 (0) 2476 523523; Fax: +44 (0) 2476 523523; emails andrew.oswald@warwick.ac.uk and e.proto@warwick.ac.uk

Abstract

This paper studies a famous unsolved puzzle in quantitative social science. Why do some nations report such high levels of mental well-being? Denmark, for instance, regularly tops the league table of rich countries' happiness; Britain and the US enter further down; some nations do unexpectedly poorly. The explanation for the long-observed ranking -- one that holds after adjustment for GDP and other socioeconomic variables -- is currently unknown. Using data on 131 countries, the paper cautiously explores a new approach. It documents three forms of evidence consistent with the hypothesis that some nations may have a genetic advantage in well-being.

A large economics and social-science literature exists on international patterns of human happiness and well-being (Diener et al. 1995, Veenhoven and Ehrhardt (1995), Oswald 1997, Di Tella et al. 2001, Easterlin 2013a, Hudson 2006, Graham 2010, Blanchflower and Oswald 2011, Helliwell and Wang 2013, Proto and Rustichini 2013, and Graham and Nikolova 2015). Research into the microeconomic and macroeconomic determinants of mental well-being -- by economists and a range of behavioural and health scientists[1] -- has become *-+65common and is beginning to shape policy-making in the public sphere (Stiglitz et al. 2009)[2]. There also exists a new literature on what might be described as the intra-nation geography of happiness (for example, Aslam and Corrado 2012, Oswald and Wu 2011, Glaeser 2015).

---

[1] Including Deaton (2008), Easterlin (2003, 2013a,b) and Alesina et al. (2004).
[2] Oswald and Wu (2010) uncover a close correspondence between subjective and objective well-being across the US states.

However, a famous lacuna exists in the research literature. A long-observed cross-country pattern remains unexplained. Since the work of Edward Diener in the early 1990s, it has been known, and constantly found in replication studies[3], that nations like Denmark and the Netherlands regularly head the league table of international life-satisfaction. Yet certain other countries, including high-GDP European countries such as France and Italy, come surprisingly far down in an international ranking. The reasons for this are not properly understood (although the innovative work of Senik 2011 provides an analysis of the puzzling French case, and Helliwell and Wang 2013 discuss possible reasons for the international regularities). It is apparently not because of elementary kinds of spurious correlation or measurement error: an equivalent cross-country pattern has been found using data on reported hypertension and on psychiatric health (Blanchflower and Oswald 2008, Ploubidis and Grundy 2009). Moreover, although the existence of a stable international pattern in well-being is to be expected if countries' wealth and institutions matter and are themselves slow-changing, the scientific difficulty is that it has proved impossible in that way to account for all of the empirical cross-national variation. The research paradox therefore continues.

Some influences are known. There is evidence that part of the long-observed ranking can be attributed to GDP levels, the quality of government, and certain welfare-state characteristics (such as in Di Tella et al. 2003, Graham 2010, Helliwell and Wang 2013, and Radcliff 2013). Nevertheless, even after adjusting for a range of such factors, the underlying league-table pattern, with Denmark at or near the top in the world happiness ranking, remains stubbornly in the data. Here we have been particularly influenced by the earlier work -- listed in the references -- of John Hudson, John Helliwell and Shun Wang, and Jan Ott. There is

---

[3] See, for example, the rankings in Helliwell and Wang (2013). Graham (2010) and Blanchflower and Oswald (2011) summarize the literature and give results on modern data.

also important new evidence that, as a statistical matter, Denmark's secret may be related to an avoidance of extreme unhappiness in its citizens (Biswas-Diener et al. 2010). However, the substantive reason for this, if it is to be part of a complete explanation, is itself unknown; so that in turn only pushes the level of explanation one layer further back.

This paper considers an avenue that we originally found implausible. To our own surprise, there is some empirical evidence consistent with a (partial) genetic explanation. It has been known for some time that in individual data on humans there appears to be a genetic element to happiness and well-being: see for example Weiss et al. (2002) and the study of twins by Weiss et al. (2008). Almost no researcher, however, has attempted to explore whether there might be a cross-country equivalent.

The key variable in the first form of evidence used in the paper is a measure of *genetic distance between countries' populations*. The later analysis finds that this variable is correlated with international well-being differences, and that the correlation seems not to be because of potential omitted variables for factors such as prosperity, culture, religion, or geographical position in the world. By using regression equations, this study aims to control for the potential confounding[4] that geneticists sometimes refer to as the 'chopsticks problem' or 'social stratification' (for example, Hamer and Sirota 2000).

---

[4] Loosely, the chopsticks problem is that it would be possible spuriously to identify a gene that appeared to cause the use of chopsticks (whereas the deeper explanation was that cultural factors caused the chopsticks use and those were merely correlated with genetics). For this reason, papers by economists on genetic data, such as Ashraf and Galor (2013), sometimes provoke fierce responses from geneticists. However, both economists and geneticists are well aware of the problem of confounding, and both disciplines attempt to guard against it, if in their different ways and using different jargon.

In a second form of evidence the paper also connects to a modern -- and rather controversial -- literature on depression and happiness levels in individuals that documents statistical evidence for an association between mental well-being and (a mutation in) the length variation in the serotonin-transporter-gene-linked polymorphic region (5-HTTLPR). The protein-encoded serotonin transporter gene influences the reuptake of serotonin, which is believed to be implicated in human mood. To our knowledge, we are among the first to consider this avenue as a possible way to explain the well-being patterns at a country level. We should like to acknowledge, however, that after the first draft of our paper was finished we discovered that related work, though not on exactly the same data, had been presented by Christie Scollon and colleagues in a conference poster session in 2012 (reported in Scollon et al. 2012). Because depression and mental disorder in people has multiplied 'externality' effects on the happiness of others, through families and friendship networks, it is to be expected that the effect of a genotype that influences individual well-being could have larger effects in community-level data than individual data.

This paper also considers a third kind of evidence that potentially links well-being to people's genetic make-up. A later section studies US immigrants' happiness levels as a function of the well-being levels of their families' original home nations. In this segment of the analysis we build upon the simple idea that an American whose family came originally from country Z will carry genes found more commonly in that country.

Genes might matter for mental well-being in two ways: directly or indirectly. Genetic influence could operate in a way orthogonal to other social-science variables or might operate by affecting such factors (or possibly both). In response to a question put by a reviewer of this paper, suppose that, after controlling for many objective factors, we found that the list of

independent variables exhausted the international well-being variance to be explained, so that the coefficient on a variable for genetic influence became approximately zero. What should then be concluded? In such a case, we would not necessarily wish to argue that genetics has no explanatory power for international happiness. We might conclude, instead, that we had potentially discovered an approximate decomposition of the channels from genetics to happiness -- that we had learned, in part, about the likely transmission mechanism from genes to well-being. Nevertheless, as the reader will be able to tell from the paper's later tables and figures, we are not able to exhaust the well-being variance to be explained. One interpretation is that genes may work in a way that is truly orthogonal to (some) social-science influences on human well-being. Another possibility is that one day, perhaps well into the future, researchers may find enough explanatory variables to exhaust the well-being variance. What can be said, currently, is that we have not found a way to drive a genetic variable to statistical insignificance.

This paper builds upon the ideas of earlier scholars. We have been especially influenced by the important research of Spolaore and Wacziarg (2009) and De Neve and colleagues (2011, 2012). We employ data used by the former, elaborated from the original genetic distance data of Cavalli-Sforza, Menozzi, and Piazza (1994), where the focus is the set of 42 world populations for which they report bilateral distances computed from 120 alleles. These populations are aggregated from subpopulations characterized by high genetic similarity. More broadly, our well-being research follows in a tradition exemplified by scholars such as Easterlin (2003), Di Tella et al. (2003), Helliwell (2003), Graham et al. (2004), and Stevenson and Wolfers (2008). Our work also relates to a stream of genetic research (Benjamin et al. 2012, Canli et al. 2005, Caspi et al. 2003, Chen et al. 2013, Chiao and Blizinsky 2010, Clarke et al. 2010, De Neve 2011, De Neve et al. 2012, Fox et al. 2009,

Gigantesco et al. 2011, Kuhnen et al. 2013, Lesch et al. 1996, Risch et al. 2009, Sen et al. 2004, Stubbe et al. 2005, Szily et al. 2008, Weiss et al. 2002, 2008).

For the later analysis, the paper needs two steps. The first is to calculate genetic differences across nations. The second is to calculate whether those differences might have any statistical explanatory power in a regression framework in which cross-country well-being is the dependent variable.

Conceptually, the nature of a genetically homogeneous population Y can be thought of as a vector of allele characteristics, y. We wish to be able to measure the distance between this population and another population, X, with allele characteristics x. Genetic distance, g, has to be captured in our empirical analysis by a scalar. Hence in the case of two populations, X and Y, we need to define some form of mapping:

$$M(g, y, x) = 0 \qquad (1)$$

where g is a scalar for genetic distance, y is a vector of genetic characteristics of country Y, and x is a vector of genetic characteristics of country X. Genetic distance can be thought of as the genetic divergence between different species and their populations. Because there is no unique mathematical way to calculate the distance between two vectors, we follow an approach from within the existing literature. We adopt 'Nei's genetic distance' metric, which can be viewed as the appropriate distance measure when genetic variants come about by genetic drift and mutations. Reassuringly, it is known that the Nei measure is correlated with other distance measures (Nei 1972, Spolaore and Wacziarg 2009). The genetic distance between two countries is a weighted transformation of the distance between homogenous populations; this is done by taking into account the percentage of individuals belonging to the different populations in the two countries, as in Spolaore and Wacziarg (2009).

As one further check on this study's conclusions, a form of cultural-epidemiological approach (Fernández 2008) is used. This relies on an examination of whether the level of subjective well-being of the descendants of immigrants to a country -- in our case the US -- is correlated with the level of subjective well-being of the original country. A correlation between the two would be consistent with the hypothesis that differences in subjective well-being are heritable. Parents transmit their genetic as well as their cultural (broadly-speaking) traits to their offspring. Since in this paper we hope to isolate the effect of the genes, in these later calculations we control for religion, income, work status, age, and gender.

This paper is organized as follows. In the next three sections we analyze the relationships between well-being and genetic distance, well-being and the 5-HTTLPR polymorphism, and the happiness of US immigrants with their home countries' subjective well-being. Section 5 concludes. The appendix describes the data, and the paper's methods, and shows that even if an alternative dyadic estimation method due to Spolaore and Wacziarg (2009), which treats the data as a set of country pairs, is used instead then the substantive conclusions remain the same.

1. Is There Evidence of a Link Between Genes and Well-being? A First Test

This study draws upon a number of international random-sample surveys and uses those to examine the relationship between well-being differences and genetic distance. The data sources include the World Values Surveys, the Gallup World Poll, the Eurobarometer Surveys, and the European Quality of Life Surveys. Although these data sets together provide information on hundreds of thousands of randomly selected individuals, we are inevitably restricted, in a cross-country analysis, in effective statistical power. The sample size for countries in this paper never exceeds approximately 140 nations.

A flavour of the first evidence is conveyed by Figure I. This is a plot of cross-national data on genetic distance and well-being. The source is data from the Gallup World Poll. On the y-axis of Figure I is a variable for (low) psychological well-being in a country. It is 'Struggling', as defined by Gallup rather than by us, which is a cross-national variable for the percentage of individuals in the country who report that their present life situation is between 5 and 7 on a ten-point scale and who report the perceived quality of their future life as between a 5 and an 8. On the x-axis is plotted Nei's genetic distance measure, which is defined here as distance from Denmark, DK^5. There is a statistically significant positive correlation. In Figure I, the greater is a nation's genetic distance from Denmark, the lower is the reported well-being of that nation (that is, the greater their population's level of struggling).

Notable countries in the data set include the high well-being nations of Netherlands and Sweden; they are depicted as dots in the south-west corner of the graph. These nations, perhaps unsurprisingly given their geographical proximity, have the closest genetic similarity to Denmark. Particularly unhappy countries in Figure I include nations such as Ghana and Madagascar; these have the least genetic similarity to Denmark. Figure II extends the analysis to other well-being measures. 'Thriving' is defined as present life situation of 7 or better and expectations of the next five years as 8 and above; 'Suffering' is defined as present situation and the next five years below 5; high life satisfaction is defined as life satisfaction exceeding 7 on a ten-point scale. In Figure II, in each quadrant, there is an association between greater well-being and having a genetic stock closer in nature to that of Denmark. In both Figures I and II, it is possible to reject the null of zero on each of the five best-fitting

---

[5] This paper's results do not depend on Denmark being the base country. As suggested by referees, we give in an online Additional Appendix some results treating instead Sweden, Norway and the Netherlands as base nation.

lines at the 99.9% confidence level. For transparency, Figure III plots the raw data for each continent.

However, as implied in the introduction, there is an obvious conceptual difficulty with such plots. By their nature, Figures I and II do not control for confounding variables. To economists and economic geographers, the most obvious of these are the prosperity of the countries and the geographical position of the nations. Hence Table I switches to regression equations. In this way, it is possible to probe the robustness of the elementary bivariate correlation between nations' well-being and genes. The five columns of Table I report regression equations in which the sample size is now 131 and the dependent variable is Struggling. Column 1 replicates the pattern of Figure I. Column 2 of Table I then introduces one extra control variable, namely, the PPP-adjusted Gross Domestic Product of each country. The GDP variable is expressed per head of population and, to match the nature of the genetic distance variable, is entered as the absolute difference from Denmark's GDP. As in much previous well-being research, GDP enters strongly positively in column 2 of Table I. The poorer the country (as captured by the distance from Danish Gross Domestic Product), the greater is the degree of psychological struggling. In this case, the coefficient on GDP is 5.63 with a small standard error of 1.05. As would be expected by an economist, when moving from column 1 of Table I to its column 2 the coefficient on Nei Genetic Distance falls. However, the Nei coefficient remains substantial and well-defined statistically. Later columns of Table I add further controls to account for other possible confounders. In column 3, the larger is the geographical distance from Denmark, the greater is the level of Struggling. Its coefficient in the equation is -4.14 with a standard error of 1.72. However, the Nei coefficient on Struggling continues to be positive, at 5.46 with a standard error of 1.58. It is also positive when a set of further dummy variables are included for the different continents.

However, the level of statistical significance falls slightly below the 5% cut-off in column 4 of Table I, once the specification includes all of Nei distance, GDP, geographical distance, and continent dummies. In column 5 of Table I, the Nei genetic distance measure returns to statistical significance, and has a coefficient of 3.61 and a standard error of 1.27.

Geographical forces could, in principle, operate in more subtle ways. However, as a check on whether geographical distance is an inadequate spatial measure, Table A1 in the Appendix shows that the key correlation with genetic distance is unaffected by the inclusion also of measures of longitudinal distance and latitudinal distance.

Table II switches to a different well-being variable, namely, that of Thriving. Consistent with the prior patterns, the correlation between Thriving and genetic distance is negative. It is possible to reject the null of zero at the 99% confidence level for each of the five specifications, including the most demanding one (arguably too demanding given the number of data points) in column 4, within Table II. Furthermore, Table III, for the same group of 131 countries, replicates the equivalent finding when using a Suffering dependent variable. Tables I, II and III thus suggest the same conclusion as the early elementary bivariate graphs.

Tables IV and V switch to traditional life-satisfaction variables. These necessarily have smaller samples. Here the source is not Gallup but instead is information drawn from the World Values Surveys. In the fullest specification, that of column 4 in each table, the same result on countries is found again. Table IV takes as its dependent variable a high level of life satisfaction (numbers over 7 out of 10) whereas Table V's dependent variable is mean life satisfaction. For both tables, the larger the divergence of the genetic stock from that of

Denmark, the lower is the country's life satisfaction. Column 5 in each of Tables IV and V sees a drop in the significance of the Nei Genetic Distance coefficient. That might at a glance be thought a weakness in the argument. However, the data favour the column 4 specifications, which have greater explanatory power. In the fullest specification of Table IV, for instance, the coefficient on Nei distance is -5.20 with a standard error of 2.26.

Are these effect-sizes substantively significant? It is natural to consider within Table IV what a coefficient of more than 5 on the Nei coefficient, in this best-fitting specification, implies. The standard deviation of Log Nei Distance is slightly greater than 1, and the standard deviation of High Life Satisfaction is approximately 12. Hence one standard deviation in genetic distance is associated with more than one third of a standard deviation in national well-being.

Countries differ in many ways more complex than differences in GDP. That leads to other likely sources of confounding. To allow a broader measure of societal prosperity to be included as a control variable, Table VI examines what happens if GDP is replaced by the Human Development Index (HDI) as defined by the United Nations. The level of HDI for a country is an average of its GDP, its educational level, and its average length of life. This could be seen as a fairly severe test for our data to pass. The reason is that HDI could itself be viewed as a measure of human well-being, so some of the variation in the dependent variable itself in a well-being regression equation is being picked up, it might be argued, by having HDI as a control within a subjective well-being equation. Nevertheless, in columns 2 to 5 of Table VI there continues to be evidence of a link between genetic makeup and the happiness of the country. The genetic-distance variable enters in Table VI with a coefficient between 7.11 and, with a longer set of controls, 2.96.

A possible concern is that the high life-satisfaction level observed in Nordic countries is due predominantly to the generosity [6] of their welfare states. The HDI variable implicitly includes education and health levels, so in part provide a control for this as well. Nevertheless, in column 3 of Table VI we do an explicit check. We introduce transfers in terms of social benefits (always in logarithmic distance from Denmark) in order to provide a fuller control for the effects of the welfare state. The social-benefits variable is derived from World Bank data. The correlation with the Nei variable, however, is unaffected by the inclusion of the social-benefits variable. In columns 4 and 5 of Table VI, continent dummies are introduced; the well-being link with the Nei variable again remains. An appendix presents similar tables for other measures.

As explained in the introduction, confounders due to cultural values are also possible. To attempt to check whether genetic distance might be standing in erroneously for such influences, Tables VII to IX reveal, for a set of thriving, struggling and suffering equations, that the Nei variable continues to be statistically significant after controlling for four cultural variables -- religious adherence, colonial origins, language distance, and Hofstede's (2001) cultural-dimensions variable. For example, considering column 4 of Table VII, the coefficient on Nei is -7.25 with a standard error of 2.61.

In these tables, there is some evidence that the coefficient on Nei actually increases. In general, however, the coefficients are largely unchanged by adding Hofstede's cultural variables. This suggests that, even if, as seems likely, our cultural variables are imperfect, the relationship between genetic diversity and subjective wellbeing is not solely explained by

---

[6] This was our presumption before we obtained any genetic data. Di Tella et al (2003) documents evidence that unemployment-benefit generosity affects national well-being. Related arguments about the welfare state were proposed by Richard Easterlin in a 2013 public lecture at Oxford University.

cultural distance. Caution nevertheless remains advisable. Table VII enters a larger number of independent variables than -- for reasons of statistical power -- is ideal with small sample sizes. Such difficulties are inherent in cross-country research, of course, but they are real ones.

A final possibility is that Nordic countries have 'better institutions' in some wide-ranging sense. We probe that possibility. Recently, Helliwell and Wang (2013) calculated the residual life satisfaction (measured with the Gallup Cantrill Ladder) after controlling for the quality of countries' institutions and culture, with variables measuring: perception of corruption, healthy life expectancy, GDP per capita, freedom to make a choice, social support and generosity (in terms of culture for charitable donations). In Table X, therefore, we try this unexplained satisfaction as a dependent variable and show that this well-being residual correlates with the index of genetic distance from Denmark, after controlling for geographic variables. Table X then does similar exercises, in its columns 2 and 3, for unexplained cross-country variation in well-being using also data from Jan Ott and John Hudson. Here the original papers are Hudson (2006) and Ott (2011).

In principle, given the statistical power, the aforementioned checks are taxing ones for the paper's hypothesis. Nevertheless, in each of the three columns of Table X, reasonably persuasive evidence for a correlation with Nei-distance remains visible.

2. A Second Test: Well-being and the 5-HTT Polymorphism

Because they leave the detailed type of any genetic effect unexplained, the previous results suffer from an important potential weakness. They are in the nature of black-box findings.

In this second section of the paper, we try to respond, if necessarily imperfectly, to such concerns by building on another literature that has previously identified evidence of a specific genetic influence on mental well-being. A large set of writings, triggered in part by a still-controversial paper by Caspi et al. (2003), has studied a particular polymorphism, 5-HTTLPR, at the individual rather than national level. This line of research, by Caspi and many subsequent researchers, suggests that the short and long variants of 5-HTTLPR are correlated with different probabilities of clinical depression. In particular, the short allele has been associated with higher scores on neuroticism and harm avoidance, stronger attentional bias towards negative stimuli, and lower life satisfaction. Consistent with that, the evidence in this section suggests there is a statistical association between lower happiness of nations and the proportion of their population who have the short allele version of the 5-HTTLPR polymorphism, which we will refer to for brevity as (S)5-HTT. Intriguingly, *among the developed nations in our data, it is Denmark and the Netherlands that appear to have the lowest percentage of people with (S)5-HTT*. These findings, we caution, should be treated warily, because when dealing with the countries for which we have 5-HTT data there is a shortage of statistical power.

Across 30 nations on which there is information in Figure IV, the mean of (S) 5-HTT is 49.63, with a standard deviation of 13.09. The short allele is thus found in approximately half the population. For those West European nations on which data are available, the scatter plot in Figure IV depicts the cross-sectional correlation between life-satisfaction and the percentage of citizens in that nation with the (S)5-HTT polymorphism. An inverse relationship exists. Denmark has the highest recorded level of satisfaction with life and it has the lowest % of citizens with (S)5-HTT. Italy has the lowest recorded level of satisfaction with life and the highest % of (S)5-HTT. In Figure IV's scatters, the left-hand set of well-

being data are drawn from the Eurobarometer Surveys. Figure IV also gives an equivalent cross-sectional correlation between mean happiness and the percentage of citizens in that nation with the (S)5-HTT polymorphism. Here the data come from the European Quality of Life Surveys. Figure V uses data on life satisfaction taken from the World Values Survey. This plot expands the previous list of countries to the so-called Western Offshoots and includes New Zealand, the USA, and Australia; for historical reasons these nations are genetically, economically, and politically similar to the Western European countries. The key correlation remains negative and significant. Figure AI in the appendix switches to an alternative well-being measure on the y-axis. It uses a 'ladder of life' well-being question due to Cantril (1965); the exact wording is explained below. Here the statistical result is the same as in the earlier figures. A similar graph can be produced using a so-called Daily Experience index developed by Gallup.

Because the variable (S)5-HTT measures the proportion of individuals with the S allele of 5-HTT, it is perhaps natural to correlate this with a measure of well-being that relates to proportions of individuals rather than to averages (like the mean of life satisfaction, happiness or the Cantril ladder-based index). Figure AII in the Appendix thus refines the Cantril measure to the struggling variable used earlier. This is the proportion of people who are classified by Gallup as having low well-being scores, as assessed by the ladder, both currently and prospectively (for more details see the appendix), in this figure we include all countries a measure of the (S)5-HTT share is available. Consistent with the earlier figures, there is a strong correlation between the percentage of people struggling psychologically in a nation and the percentage of the nation's population who have the short allele of 5-HTT. An equivalent figure can be produced if we consider as an alternative a thriving variable based on the same principles as struggling.

In order to check the consistency of these data with those from the World Values Survey, Figure AIII of the Appendix shows the equivalent relationship between (S)5-HTT and an index 'Very Satisfied' which is a measure of the proportion of individuals reporting life satisfaction larger than 8. A similar pattern emerges if instead the data are on the proportion of individuals reporting life satisfaction larger than 7 (although the p-value on the gradient is then right at the border of the 0.05 cut-off).

Table XI presents simple regression equations for the reduced sample. Here the dependent variable is the percentage of citizens who are defined by Gallup as 'struggling', and the main explanatory variable is the share of the (S)5-HTT polymorphism, which here, to be consistent with the above analysis, is expressed in terms of log of the absolute distance from the values in Denmark. There are only 28 observations, one for each country, so it is necessary to be sparing with the number of independent variables included in these regressions. In column 1, the previous bivariate correlation of the figures is extended by including a variable for Nei genetic distance. It enters with a coefficient of 1.69, with a large standard error of 2.35. In column 1 of Table XI it is the 5-HTT distance variable that is now significant, so this 'horse-race' test, admittedly of a simple kind, seems to suggest it may be a more important explanatory factor than Nei distance per se. Column 2 of Table XII might be viewed as consistent with that. In column 3, the 5-HTT variable survives the inclusion of GDP and geographical distance.

Broadly similar findings are visible in Table XII, where 'thriving' is a dependent variable. It should perhaps be said that, with 4 variables and 28 observations, the regression equations in the last two tables are potentially over-fitted. They should be viewed only as approximate checks.

3. A Third Test: Using Data on US Immigrant Descendants in an Examination of the Possible Genetics of Subjective Well-being

The paper provides a final, and purposely different, form of evidence. One of the unusual advantages of a genetic influence is that in principle it should be visible even if historical measures are used. This is because genetic patterns inherently stem from a previous era.

In Table XIII, we exploit that idea. The table uses data on 29 nations, which is the largest sample available to us for the exercise. The independent variable here is the coefficient obtained from an ordered probit regression, where (current) happiness of individuals who are born in the US is regressed against their family country-of-origin dummies. There are also controls in the regression equation for age (and its square), gender, income, education and religion; these latter corrections are to provide some control for cultural values.[7] The dependent variables -- there are four in Table XIII -- are the same well-being measures used in the paper's earlier tables.

The exact methodology is the following. For Americans who report, say, that they have family origins from Italy, we create *an independent variable derived from the happiness level of current Italian-Americans.* That independent variable is used, in a regression equation, to help explain *the current happiness level of Italy.* In effect, the same procedure is repeated for each country within the data set. Here the ultimate aim is to see whether the current well-being of nations is correlated with the reported well-being of Americans who

---

[7] This procedure is commonly used in the literature to analyze the effect of culture on economic variables like GDP or labour supply (Guiliano 2007).

have ancestors from that nation. The purpose of this statistical exercise is not, of course, to argue that happy Italian-Americans directly cause the happiness of today's Italy. Rather, what the evidence suggests, consistent with the existence of an underlying genetic component in international well-being patterns, is that there is an unexplained positive correlation between the happiness today of Country X and the observed happiness of those Americans *whose ancestors came from Country X*. In the first column of Table XIII, for example, the coefficient is -46.5 with a standard error of 16.8 (the reason the coefficient is negative is that it is for an equation for Struggling rather than well-being). Such evidence is consistent with a genetic influence.

4. Conclusions

This study offers three kinds of evidence in an attempt to make progress on one of the classic puzzles of modern social science. For some decades, a much-replicated international ranking of happiness and well-being has remained predominantly unexplained. The paper has found that the cross-country happiness pattern appears to be correlated with genetic differences. The implied effect-size is apparently not small. The right-hand columns of Table II, for example, reveal that a one standard deviation in genetic distance is associated with more than one third of a standard deviation in country well-being.

The closer a nation is to the genetic makeup of Denmark then the happier is that country. As a raw uncorrected social-science correlation, such a result would not be a persuasive one. However, what is more interesting is that the correlation seems to survive adjustment in the regression equations for many confounding variables (which some kinds of researchers would refer to as adjustment for the 'chopsticks problem'). It is robust, for instance, to the inclusion of controls for

(i)   the GDP of the country,

(ii)   the level of the Human Development Index of the country,

(iii)   the geographical distance of the country from Denmark,

(iv)   a range of cultural and religious variables,

(v)   separate dummy variables for each continent,

(vi)   longitude and latitude variables,

(vii)   indices of nations' institutions and the generosity of their welfare states,

(viii) calculated residuals from the independent country-ranking research of John Hudson, John Helliwell and Shun Wang, and Jan Ott.

Hence the relationship between well-being and genetic distance is not due merely to inherent differences between the world's continents, nor to the obvious fact that, for example, African nations are poor and have different genetic characteristics than rich European countries, nor to some elementary omission of welfare-state variables. It should perhaps also be recorded that if we switch to the estimation method of Spolaore and Wacziarg (2009), which treats the data as a set of country dyadic pairs, the substantive conclusions are unchanged. Tables A2 and A3, in an online appendix, illustrate those results.

The paper's findings are contrary to our own presumptions when we started the inquiry (although studies of twins have concluded that in individual data there is a genetic component to well-being). While we wish to continue to emphasize the need to remain extremely cautious, there are empirical reasons to think that genetic patterns may help researchers to understand international well-being levels. If true -- and other research on the topic is now needed -- this suggests that economists and other social scientists may need to pay greater heed to the role of genetic variation across national populations.

The patterns uncovered in this work should be treated with circumspection. False positives are common in genetic studies. It is valuable to recall especially the strictures of Benjamin et al. (2012), and in particular three concerns: statistical power; the multiple comparisons problem; the comparatively small differences in genetic makeup and (S)5-HTT in industrialized countries.

First, and most seriously, the largest data set at our disposal has 143 cross-national observations, although admittedly it is possible to study country pairs and thereby generate a form of data set with some thousands of data points. An important avenue for future research will be to check that the results can be replicated in other ways -- perhaps across regions within nations. Second, it is known in the field of genetics that the search for patterns can routinely lead to the discovery of illusory Type-I-error associations. For this reason, the first section of the paper examines many possible confounders. To try further to guard against the problem, (i) we followed the lead, in the second part of the paper, of an established literature that previously found at the level of the individual there is evidence to implicate (S)5-HTT polymorphisms in the causes of happiness and depression, and (ii) we documented evidence that that variable goes some way to explaining the statistical significance of the genetic distance variable used in the first part of the paper. Third, the findings in the second section of the paper would imply that noticeable well-being differences across countries could be linked to fairly small differences in the proportion of their populations with the short allele version of 5-HTT. At a glance, that fact -- a spread of only 10 percentage points in the populations -- appears paradoxical. It certainly reduces the plausibility of the paper's second kind of evidence[8]. Perhaps a potential explanation, to be explored in future research, might

---

[8] A critic could argue that it is a statistical fluke that the famously happy country of Denmark has the lowest endowment of the form of 5-HTT polymorphism that, in the eyes of some previous researchers, has been

lie in a form of happiness multiplier within a society. If the happiness of an individual is magnified by social contact with other happy people -- as has been demonstrated by researchers James Fowler and Nicholas Christakis (2008) and seems anecdotally clear from observation of emotional externalities upon members within a family where someone has clinical depression -- then it is straightforward to write down models in which small differences in starting happiness can have larger, multiplied effects throughout a society. A framework of a related kind has been developed for social-science settings (Clark and Oswald 1998). The broad idea of matrix multipliers in social science is an old one and goes back, in a different substantive setting, to the work on input-output theory by the late Wassily Leontief (1936). Here, let h be a vector of happiness levels in the population (where the length of the vector is the number of individuals), A be a matrix of coefficients of happiness interdependence, and e be a vector of genetic endowments of happiness. Then the happiness vector in a society is a fixed-point solution given by equation:

$$h = Ah + e \qquad (2)$$

$$= (I - A)^{-1} e \qquad (3)$$

Happiness vector = multiplier matrix * genetic happiness-endowment vector

where I is the identity matrix. In this framework, a greater genetic endowment of happiness would have magnified effects in society, and these would work through a multiplier matrix given by the inverse of $(I - A)$.

In conclusion, this paper has approached an unexplained phenomenon in a new way. Its three kinds of evidence -- on genetic distance, on 5-HTTLPR, and on immigrants' happiness -- are designed to be seen as complementary; each, singly, would be less persuasive. Caution remains prudent. Much remains to be done, particularly at the intriguing

---

implicated in mental depression (see Figure IV). However, that would leave the paper's other two kinds of evidence, on genetic distance and on immigrants' happiness, still to be explained.

border between social and natural science, before it will be possible to claim a proper understanding of the determinants of nations' well-being.

References

Alesina, A., Di Tella, R. and MacCulloch, R. (2004). 'Inequality and happiness: Are Europeans and Americans different?', *Journal of Public Economics*, vol. 88, pp. 2009-2042.

Ashraf, Q. and Galor, O. (2013). 'The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development', *American Economic Review*, vol.103, pp. 1-46.

Aslam, A. and Corrado, L. (2012). 'The geography of well-being', *Journal of Economic Geography*, vol. 12, pp. 627-649.

Barro, R. (2003). *Religion adherence data-set*. Harvard University, Economics Department.

Benjamin, D.J., Cesarini, D., Chabris C.F., Glaeser, E.L. and Laibson, D.I. (2012). 'The promises and pitfalls of genoeconomics', *Annual Review of Economics*, vol. 4, pp. 627-662.

Biswas-Diener, R., Vitterso, J. and Diener, E. (2010). 'The Danish effect: Beginning to explain high well-being in Denmark', *Social Indicators Research*, vol. 97, pp. 229-246.

Blanchflower, D.G. and Oswald, A.J. (2008). 'Hypertension and happiness across nations', *Journal of Health Economics*, vol. 27, pp. 218-233.

Blanchflower, D.G. and Oswald, A.J. (2011). 'International happiness: A new view on the measure of performance', *Academy of Management Perspectives*, vol. 25, pp. 6-22.

Canli T., Omura K., Haas B.W., Fallgatter, A.J., Constable, R.T. and Lesch, K.P. (2005). 'Beyond affect: A role for genetic variation in the serotonin transporter in neural activation during a cognitive attention task', *Proceedings of the National Academy of Sciences of the USA*, vol. 102, pp. 12224-12229.

Cantril, H. (1965). *The Pattern of Human Concerns*, New Brunswick, NJ, Rutgers University Press.

Caspi, A., Sugden, K., Moffit, T.E., et al. (2003). 'Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene', *Science*, vol. 301, pp. 386-389.

Cavalli-Sforza, L., Menozzi, P. and Piazza, A. (1994). *The History and Geography of Human Genes*, NJ, Princeton University Press.

Chen, H., Pine, D.S., Ernst, M., et al. (2013). 'The MAOA gene predicts happiness in women', *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, vol. 40, pp. 122-125.

Chiao, J.Y. and Blizinsky, K.D. (2010). 'Culture-gene coevolution of individualism-collectivism and the serotonin transporter gene', *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, pp. 529-537.

Clark, A.E. and Oswald, A.J. (1998). 'Comparison-concave utility and following behavior in social and economic settings', *Journal of Public Economics*, vol. 70, pp. 133-155.

Clarke, H., Flint, J., Attwood, A.S. and Munafo, M.R. (2010). 'Association of the 5-HTTLPR genotype and unipolar depression: A meta-analysis', *Psychological Medicine*, vol. 40, pp. 1767-1778.

Deaton, A. (2008). 'Income, health and wellbeing around the world: Evidence from the Gallup World Poll', *Journal of Economic Perspectives*, vol. 22, pp. 53-72.

Diener, E., Suh, E.M., Smith, H. and Shao, L. (1995). 'National differences in reported subjective well-being: Why do they occur?', *Social Indicators Research*, vol. 34, pp. 7-32.

De Neve, J-E. (2011). 'Functional polymorphism (5-HTTLPR) in the serotonin transporter gene is associated with subjective well-being: Evidence from a US nationally representative sample', *Journal of Human Genetics*, vol. 56, pp. 456-459.

De Neve, J-E, Christakis, NA, Fowler, JH and Frey, BS (2012). 'Genes, economics, and happiness', *Journal of Neuroscience, Psychology, and Economics*, vol. 5, pp. 193-211.

Di Tella, R., MacCulloch, R.J., and Oswald, A.J. (2001). 'Preferences over inflation and unemployment: Evidence from surveys of happiness', *American Economic Review*, vol. 91, pp. 335-341.

Di Tella, R., MacCulloch, R.J., and Oswald, A.J. (2003). 'The macroeconomics of happiness', *Review of Economics and Statistics*, vol. 85, pp. 809-827.

Easterlin, R.A. (2003). 'Explaining happiness', *Proceedings of the National Academy of Sciences of the USA*, vol. 100(19), pp. 11176-11183.

Easterlin, R.A. (2013a). 'Happiness, growth, and public policy', *Economic Inquiry*, vol. 51, pp. 1-15.

Easterlin, R.A. (2013b). Public lecture at Oxford University, October 2013.

Fearon, J. (2003). 'Ethnic and cultural diversity by country', *Journal of Economic Growth*, vol. 8, 195–222

Fernández, R. (2008). 'Culture and economics', in the *New Palgrave Dictionary of Economics*, 2nd edition, edited by Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan (Basingstoke and New York).

Fowler, J.H. and Christakis, N.A. (2008). 'Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study', *British Medical Journal*, vol. 337, Article a2338.

Fox, E., Ridgewell, A. and Ashwin, C. (2009). 'Looking on the bright side: Biased attention and the human serotonin transporter gene', *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, pp. 1747–1751.

Gallup (2009). *World Poll Methodology*. Technical Report. Washington, DC.

Gigantesco, A., Stazi, M.A., Alessandri, G., Medda, E., Tarolla, E. and Fagnani, C. (2011). 'Psychological well-being (PWB): A natural life outlook? An Italian twin study on heritability of PWB in young adults', *Psychological Medicine*, vol. 41, pp. 2637-2649.

Glaeser, E.L. (2015). 'Unhappy cities', *Journal of Labor Economics*, forthcoming.

Graham, C., Eggers, A. and Sukhtankar, S. (2004). 'Does happiness pay? An exploration based on panel data from Russia', *Journal of Economic Behavior & Organization*, vol. 55, pp. 319-342.

Graham, C. (2010). *Happiness Around the World: The Paradox of Happy Peasants and Miserable Millionaires*, Oxford University Press.

Graham, C. and Nikolova, M. (2015). 'Bentham or Aristotle in the development process? An empirical investigation of capabilities and subjective well-being', *World Development*, forthcoming.

Guiliano, P. (2007). 'Living arrangements in Western Europe: Does cultural origin matter?', *Journal of the European Economic Association*, vol. 5(5), pp. 927-952.

Hamer, D. and Sirota, L. (2000). 'Beware the chopsticks gene: Population stratification is a potential source of error in psychiatric genetics, *Molecular Psychiatry*, vol. 5, pp. 11-13.

Helliwell, J. (2003). 'How's life? Combining individual and national variables to explain subjective well-being', *Economic Modelling*, vol. 20, pp. 331-360.

Helliwell, J. and Wang, S. (2013). 'World happiness: trends, explanations and distribution', in *World Happiness Report 2013*, United Nations, chapter 2.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, and Organizations Across Nations*. Thousand Oaks. CA 2nd edition. Sage.

Hudson, J. (2006). 'Institutional trust and subjective well-being across the EU', *Kyklos*, vol. 59, pp. 43-62.

Kuhnen, C.M., Samanez-Larkin, G.R. and Knutson, B. (2013). 'Serotonergic genotypes, neuroticism, and financial choices', *PLoS ONE*, vol. 8(1), e54632.

Leontief, W.W. (1936). 'Quantitative input and output relations in the economic system of the United States', *Review of Economics and Statistics*, vol. 18, pp. 105-125.

Lesch, K.P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., et al. (1996). 'Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region', *Science*, vol. 274, pp. 1527–1531.

Nei, M. (1972). 'Genetic distance between populations', *American Naturalist*, vol. 106, pp. 283.

Oswald, A.J. (1997). 'Happiness and economic performance', *Economic Journal*, vol. 107, pp. 1815-1831.

Oswald, A.J. and Wu, S. (2010). 'Objective confirmation of subjective measures of human well-being: Evidence from the USA', *Science*, vol. 327, pp. 576-579.

Oswald, A.J. and Wu, S. (2011). 'Well-being across America', *Review of Economics and Statistics*, vol. 93, pp. 1118-1134.

Ott, J.C. (2011). 'Government and happiness in 130 nations: Good governance fosters higher level and more equality of happiness', *Social Indicators Research*, vol. 102, pp. 3-22.

Ott, J.C. (2012). 'An eye on happiness: Happiness as an additional goal for citizens and governments', Phd dissertation, Rotterdam, November.

Ploubidis, G.B. and Grundy, E. (2009). 'Later-life mental health in Europe: A country-level comparison', *Journals of Gerontology Series B – Psychological Sciences and Social Sciences*, vol. 64, pp. 666-676.

Proto, A. and Rustichini, A. (2013). 'A reassessment of the relationship between GDP and life satisfaction', *PLoS ONE*, vol. 8(11), e79358.

Radcliff, B. (2013). *The Political Economy of Human Happiness*, Cambridge University Press.

Risch, N., Herrell, R., Lehner, T., et al. (2009). 'Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: A meta analysis', *JAMA - Journal of the American Medical Association*, vol. 301, pp. 2462-2471.

Scollon, C.N., Koh, S. and Wirtz, D. (2012). 'The influence of serotonin transporter gene polymorphism (5HTTLPR) and culture on subjective well-being', Poster presented at the Annual Meeting for the Society of Personality and Social Psychology, San Diego, CA.

Sen, S., Burmeister, M. and Ghosh, D. (2004). 'Meta-analysis of the association between a serotonin trans-porter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits', *American Journal of Medical Genetics (Neuropsychiatric Genetics),* vol. 127B, pp. 85–89.

Senik, C. (2011). 'Why are the French so unhappy? The cultural dimension of happiness', PSE working paper, Paris.

Spolaore, E. and Wacziarg, R. (2009). 'The diffusion of development', *Quarterly Journal of Economics*, vol. 124, pp. 469-529.

Stevenson, B. and Wolfers, J. (2008). 'Economic growth and subjective well-being: Reassessing the Easterlin Paradox', *Brookings Papers on Economic Activity*, Spring, vol. 1, pp. 1-102.

Stigliz, J., Sen, A.K. and Fitoussi, J-P. (2009). Commission on the measurent of economic performance and social progress. A report to Nicholas Sarkozy. OECD, Paris. Downloadable at www.stiglitz-sen-fitoussi.fr

Stubbe, J.H., Posthuma, D., Boomsma, D.I. and De Geus, E.J.C. (2005). 'Heritability of life satisfaction in adults: A twin-family study', *Psychological Medicine*, vol. 35, pp. 1581-1588.

Szily, E., Bowen, J., Unoka, Z., Simon, L. and Keri, S. (2008). 'Emotion appraisal is modulated by the genetic polymorphism of the serotonin transporter', *Journal of Neural Transmission*, vol. 115, pp. 819-822.

Veenhoven, R. and Ehrhardt, J. (1995). 'The cross-national pattern of happiness – Test of predictions implied in 3 theories of happiness', *Social Indicators Research*, vol. 34, pp. 33-68.

Weiss, A., King, J.E. and Enns, R.M. (2002). 'Subjective well-being is heritable and genetically correlated with dominance in chimpanzees (Pan troglodytes)', *Journal of Personality and Social Psychology*, vol. 83, pp. 1141-1149.

Weiss, A., Bates, T.C. and Luciano, M. (2008). 'Happiness is a personal(ity) thing – The genetics of personality and well-being in a representative sample', *Psychological Science*, vol. 19, pp. 205-210.

Appendix

A1: Data Appendix and Further Robustness Checks

This appendix is designed as a guide. Figures and Tables of an Additional Appendix (downloadable from either of the authors' websites, including from www.andrewoswald.com) provide further information and robustness tests. Table B6 of that Additional Appendix summarizes the main variables and provides a series of extra checks. Lastly, Tables B7 and B8 of the Additional Appendix give specific robustness tests that were requested by referees.

For the second section of the analysis in the paper, we examined associations between nations' well-being and the prevalence of the short (S) allele of the 5-HTTLPR. Because a cross-national study has relatively few degrees of freedom, we focused on the single hypothesis of a linear relationship between well-being and the proportion of individuals in the population with the short allele 5-HTT. Throughout, significance tests were two-tailed and alpha was set at 0.05.

Our statistics draw upon painstaking data collection by Joan Chiao and Katherine Blizinsky on allelic frequency of 5-HTTLPR among 50135 individuals living in 29 nations + Taiwan (Argentina, Australia, Austria, Brazil, Denmark, Estonia, Finland, France, Germany, Hungary, India, Israel, Italy, Japan, Korea, Mexico, the Netherlands, New Zealand, Poland, People's Republic of China, Russia, South Africa, Slovenia, Singapore, Spain, Sweden, Taiwan, Turkey, UK and USA). Their data set was compiled from 124 peer-reviewed publications.

We combined this genetic information with well-being data taken from various social-science sources. In most cases, we used the original surveys ourselves to calculate the well-being scores. Some of our well-being measures, however, were developed by Gallup, based on the Cantril Self-Anchoring Striving Scale. The Cantril Self-Anchoring Scale consists of the following: *Please imagine a ladder with steps numbered from zero at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time? (ladder-present) On which step do you think you will stand about five years from now? (ladder-future).* Based on statistical studies of the ladder-present and ladder-future scales, Gallup formed an index called Thriving -- well-being that is strong, consistent, and progressing. These respondents have positive views of their present life situation (7+) and have positive views of the next five years (8+). Another index is Struggling -- well-being that is moderate or inconsistent. These respondents have moderate views of their present life situation OR moderate OR negative views of their future; they are either struggling in the present or expect to struggle in the future. The exact cut-offs are that Gallup classifies people in this way if they report current life to be between a 5 and a 7 and their future life between a 5 and an 8. Finally, Suffering includes the individuals who rate both their current and their future satisfaction levels equal to or less than 4.

We complemented Gallup data by using life satisfaction data taken from the World Values Survey (WVS) for both an analysis of 30 countries and for a smaller sample of European nations. In the WVS the variable used to assess personal satisfaction is the answer to the question: "*All things considered, how satisfied are you with your life as a whole these days?*" which is coded on a scale from 1 (dissatisfied) to 10 (satisfied). We consider the data

from the two last waves: 1999-2004 and 2005-2008; we often use the proportion of individuals declaring level of life satisfaction equal to 9, 10 or to 8,9,10.

Finally, we also use data on life satisfaction in 2010 from the Eurobarometer Surveys (The Eurobarometer asks: *'On the whole how satisfied are you: very satisfied(=4); fairly satisfied (=3); not very satisfied (=2) or not at all satisfied (=1) with the life you lead?'*) and data on self-reported happiness from the European Quality of Life Survey, 2007 (*Taking all things together on a scale of 1 to 10, how happy would you say you are? Here 1 means you are very unhappy and 10 means you are very happy*), taken from the coefficients in earlier work on European well-being patterns by Blanchflower and Oswald (2008, Table 4).

Another index of well-being considered in the analysis is the residual of the Gallup Cantril ladder after controlling for healthy life expectancy, perception of corruption, GDP per capita, freedom to make a choice, social support, and generosity -- as developed in Helliwell and Wang (2013).

The country per-capita GDP data are taken from the World Bank World Development Indicators data set and relate to year 2005; they are PPP adjusted and are expressed in constant US Dollars. The social-benefit variable, expressed as a percentage of GDP, relates to year 2008 and is from the World Bank World Development Indicators data set. The United Nation HDI (Human Development Index) relates to 2005.

The cultural variables that we considered include the well-known Hofstede cultural-dimensions variable at the country level.[9] The religion adherence data are from Barro (2003). The index of linguistic distance from Danish follows Fearon[10] (2003), and the data on colonial origins are taken from the CEPII dataset (http://www.cepii.fr/CEPII/en/bdd_modele/bdd.asp) and are expressed as dummy variables indicating, in the case of each country, the long-term colonizer.

For Table XIII, the happiness of individuals born in the United States is available from the General Social Survey database (GSS). This data source covers the period 1972-2012 and provides information on the birthplace and country of origin of the respondent's forebears since 1977. The GSS variable for the country of origin reads as follows: "From what countries or part of the world did your ancestors come?" We use answers to that question as a marker of (a degree of) genetic influence from that country.

NB: All data and codes to produce the following Figures and Tables are available online

FIGURES

Figure 1: *A Multi-Country Scatter Plot of the Relationship Between Psychological 'Struggling' and Genetic Distance from Denmark*

---

[9] They were developed by Hofstede from surveys of IBM employees in approximately 60 countries.
[10] Fearon (2003) used data from Ethnologue to create linguistic trees, thereby classifying languages into common families and displaying graphically the degree of relatedness of world languages. The linguistic tree in this data set contains up to 15 nested classifications. If two languages share many common nodes in the tree, these languages are more likely to trace their roots to a more recent common ancestor language. The number of common nodes in the linguistic tree, then, is a measure of linguistic similarity.
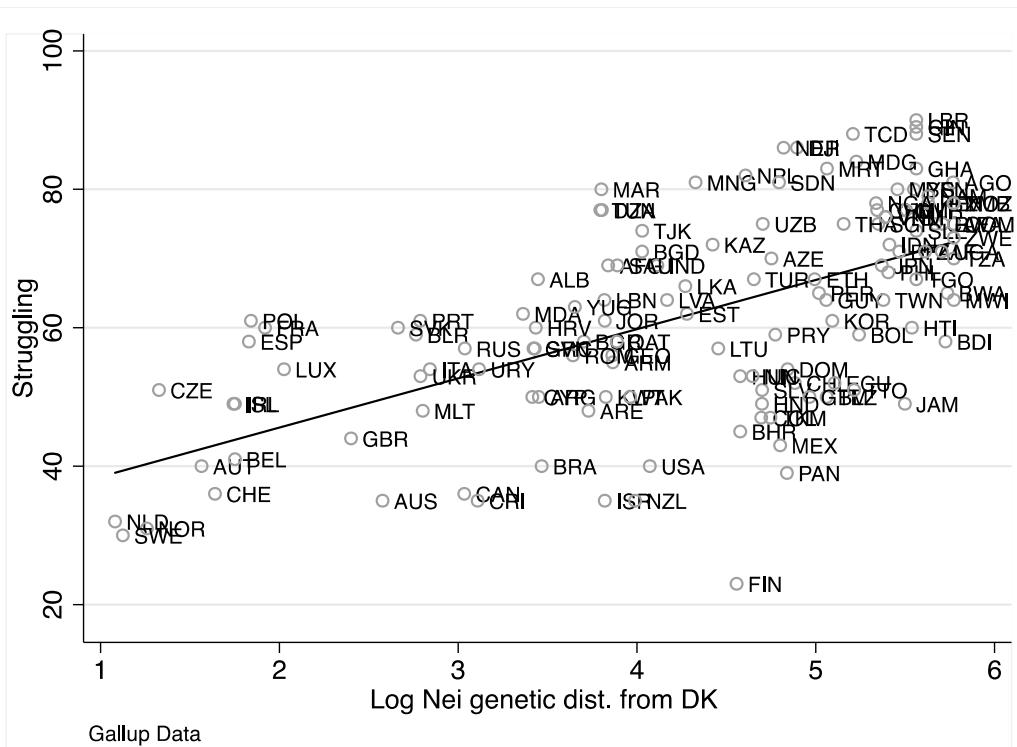
Each dot is a country. Here, and in later figures and tables, the genetic distance variable on the x-axis is calculated with respect to Denmark (denoted DK) as the base. The variable on the y-axis is 'struggling' as defined by Gallup and is a measure of the proportion of people with low mental well-being.

The genetic distance measure here uses the classic definition due to the early work of Masatoshi Nei.

Nei M. Interspecific gene differences and evolutionary time estimated from electrophoretic data on protein identity. *Amer. Naturalist* **105**:385-98, 1971.

> *Nei's distance measure D = -ln I*
> where $I = \Sigma x_i y_i / (\Sigma x_i^2 \, \Sigma y_i^2)^{0.5}$

Figure 2: *Multi-country Scatter Plots of the Relationship Between a Variety of Well-being Variables and the Genetic Distance from Denmark*
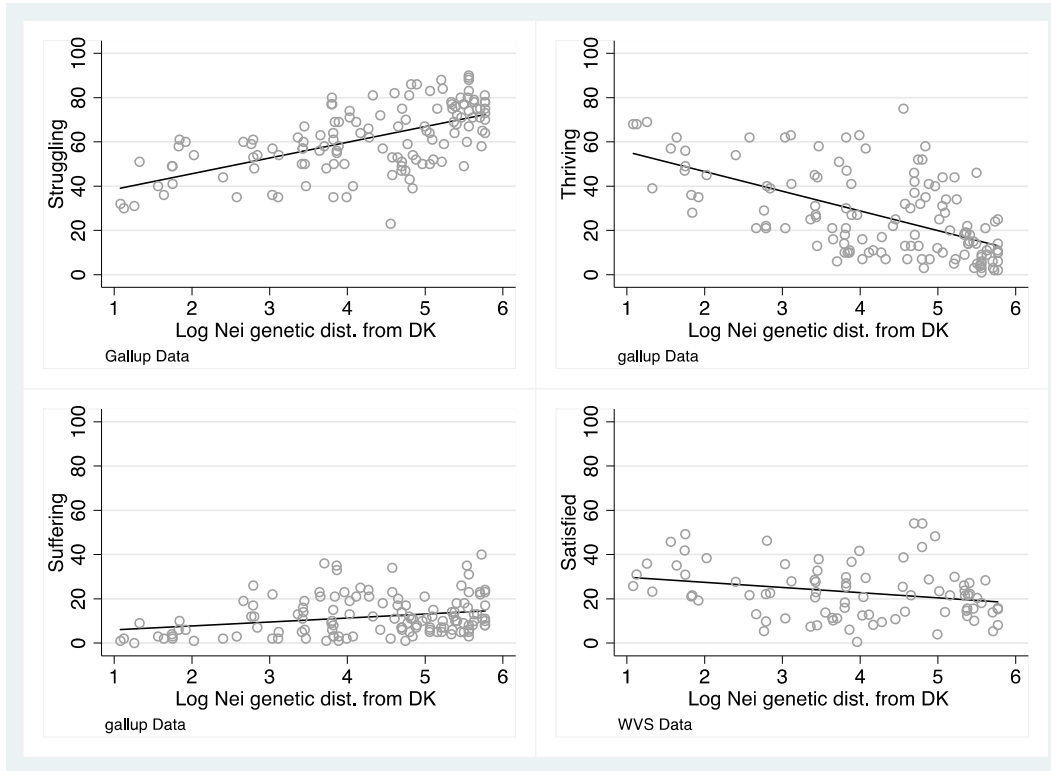
Each dot is a country. These four graphs use four different measures of mental well-being. The variables on the y-axis in each of the two left-hand graphs are measures of the proportion of people with low mental well-being. The variables on the y-axis in each of the two right-hand graphs are measures of the proportion of people with high mental well-being. Data and definitions are from Gallup and the World Values Survey (WVS).

Figure 3: *The Raw Data, in Each Continent, for 'Struggling' and the Genetic Distance from Denmark*
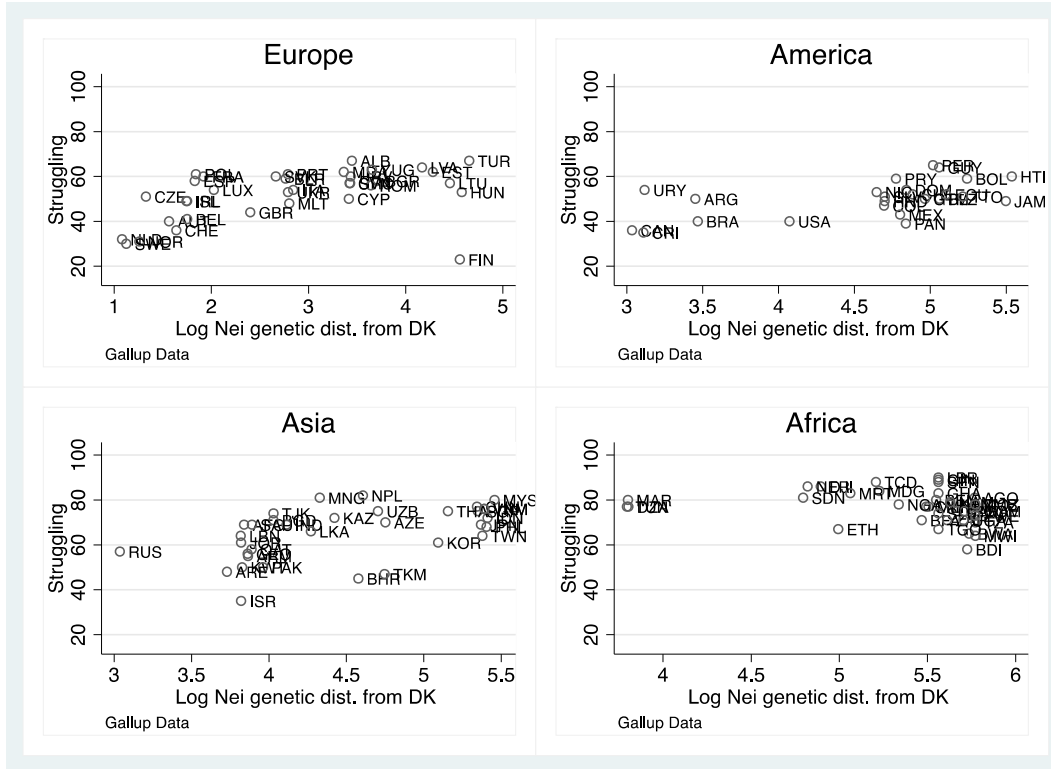
If best-fitting lines are estimated for each of these continents, the lines for Europe, America and Asia have a positive and statistically significant slope (at 95% on a two-tailed test), and the line for Africa has a negative and non-significant slope.

Figure 4: *The Correlation Between Life Satisfaction and Happiness and (S)5-HTT in the West European Countries (from Eurobarometers in the left-hand graph, and European Quality of Life Survey in the right-hand graph)*
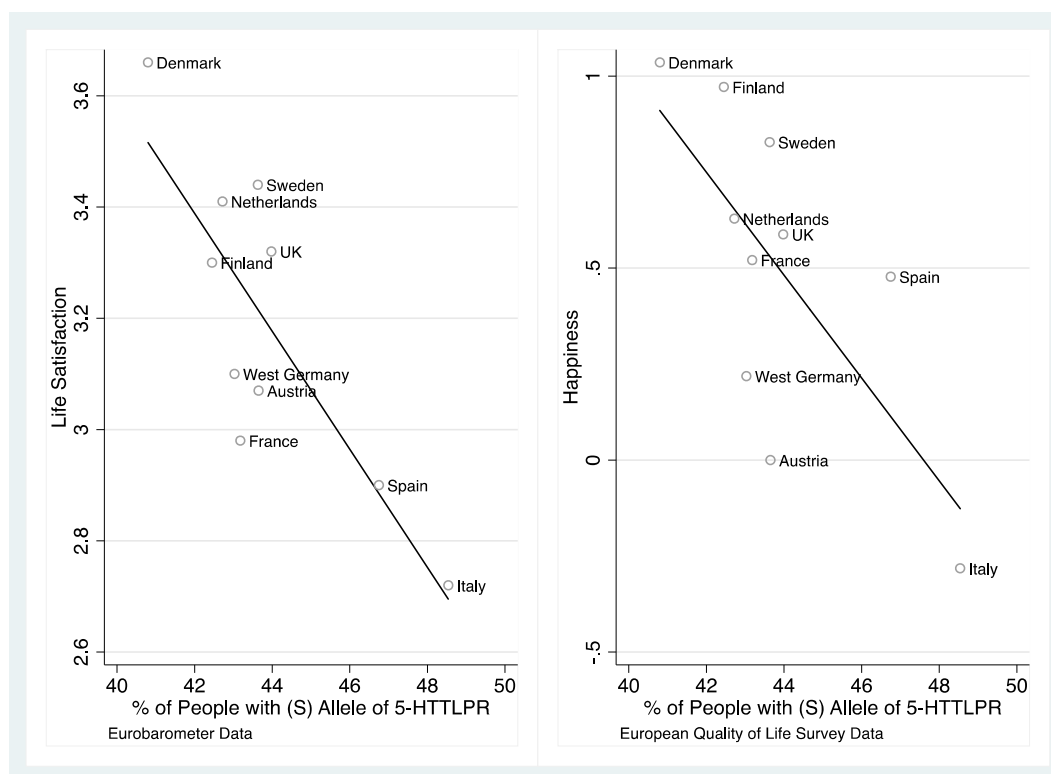
These graphs use genetic data taken from Chiao and Blizinsky (2010).

Figure 5: *The Correlation Between Life Satisfaction and (S)5-HTT in the West European Countries and Western Offshoots (from the World Values Survey)*
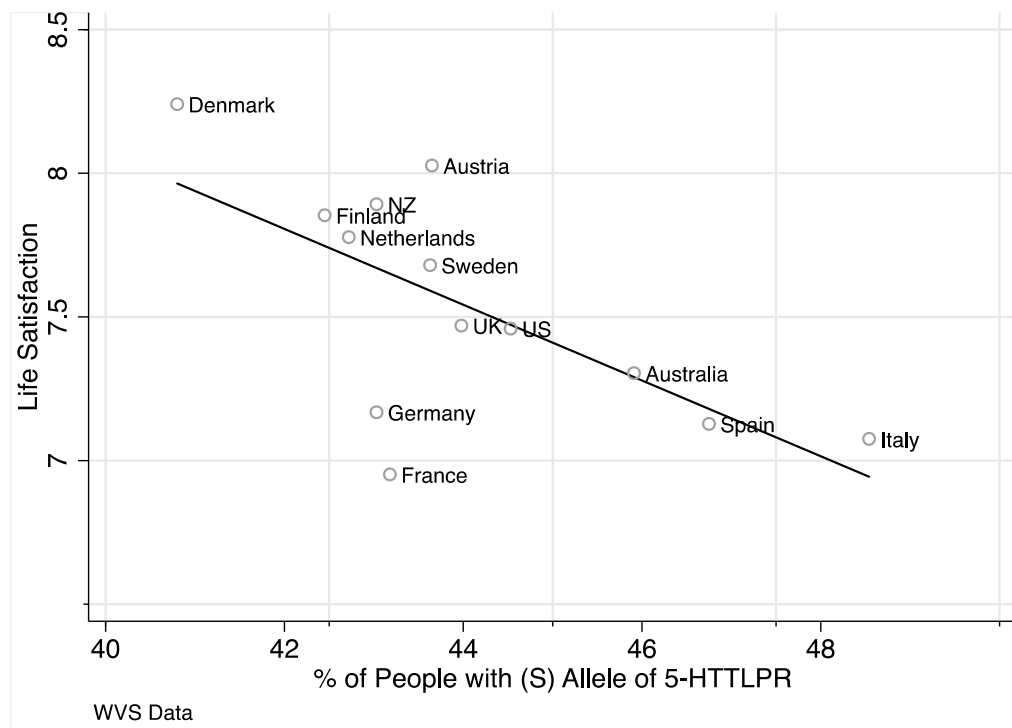
This graph uses genetic data taken from Chiao and Blizinsky (2010).

TABLES

Table 1: *Psychological-Struggling Regression Equations for a Sample of 131 Nations*

(DK here is Denmark. The dependent variable 'Struggling' is defined as present life situation between 5 and a 7 and future life between a 5 and an 8. Gallup data.)

VARIABLES	(1) Struggling	(2) Struggling	(3) Struggling	(4) Struggling	(5) Struggling
Log Nei genetic dist. from DK	7.11***	3.22***	5.46***	2.49*	3.61***
	(0.68)	(1.01)	(1.58)	(1.49)	(1.27)
GDP difference from DK		5.63***	5.46***	3.24***	3.27***
		(1.05)	(1.13)	(0.98)	(1.00)
Log geographic dist. from DK			-4.14**	4.52*	
			(1.72)	(2.51)	
Africa				2.67	7.18*
				(4.60)	(3.85)
America				-19.1***	-11.6***
				(4.33)	(3.04)
Asia				-1.70	3.39
				(3.85)	(3.19)
Oceania				-29.0***	-17.3***
				(6.02)	(2.97)
Constant	31.4***	37.8***	63.0***	11.3	39.9***
	(2.98)	(3.17)	(10.6)	(16.4)	(3.47)
Observations	131	131	131	131	131
R-squared	0.359	0.500	0.529	0.682	0.672

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The variable denoted 'Log Nei genetic dist. from DK' is the logarithm of the genetic distance between each nation and the nation of Denmark (using the method developed by Nei). The paper's results do not depend on Denmark as the choice of the base country.

Table 2: *Thriving Equations for a Sample of 131 Nations*

('Thriving' is defined as present life situation (7+) and the next five years (8+). Gallup data)

VARIABLES	(1) Thriving	(2) Thriving	(3) Thriving	(4) Thriving	(5) Thriving
Log Nei genetic dist. from DK	-8.90***	-3.30**	-8.08***	-6.17***	-6.26***
	(0.87)	(1.27)	(2.04)	(2.18)	(1.71)
GDP diff from DK		-8.12***	-7.74***	-7.08***	-7.08***
		(1.19)	(1.30)	(1.09)	(1.09)
Log geographic dist. from DK			8.81***	-0.34	
			(2.25)	(3.74)	
Africa				8.28	7.94
				(5.96)	(5.36)
America				24.2***	23.7***
				(6.22)	(4.28)
Asia				4.81	4.42
				(5.53)	(4.43)
Oceania				27.2***	26.3***
				(9.57)	(5.13)
Constant	64.4***	55.2***	1.64	60.1**	57.9***
	(4.08)	(4.23)	(14.1)	(24.1)	(4.58)
Observations	131	131	131	131	131
R-squared	0.350	0.532	0.616	0.706	0.706

DK here, and in later tables, is Denmark.

Table 3: *Suffering Equations for a Sample of 131 Nations*

('Suffering' is defined as a present life situation less than 7 and the perceived next five years of less than 8. Gallup data.)

VARIABLES	(1) Suffering	(2) Suffering	(3) Suffering	(4) Suffering	(5) Suffering
Log Nei genetic dist. from DK	1.80***	0.12	2.62***	3.66***	2.65***
	(0.49)	(0.60)	(0.90)	(0.99)	(0.86)
GDP diff from DK		2.43***	2.24***	3.75***	3.73***
		(0.74)	(0.74)	(0.82)	(0.83)
Log geographic dist. from DK			-4.61***	-4.11**	
			(1.05)	(2.06)	
Africa				-10.8***	-14.8***
				(3.42)	(3.13)
America				-5.10	-11.9***
				(4.00)	(2.28)
Asia				-3.06	-7.70***
				(3.69)	(2.81)
Oceania				1.68	-8.97***
				(5.84)	(2.55)
Constant	4.15*	6.93***	35.0***	28.3**	2.18
	(2.12)	(2.09)	(6.86)	(13.4)	(2.14)
Observations	131	131	131	131	131
R-squared	0.067	0.143	0.250	0.333	0.309

Table 4: *High -Life-Satisfaction Equations for a Sample of 86 Nations*

(The dependent variable 'Lfsato8910' is defined here as life satisfaction between 8 and 10. Source: WVS data)

VARIABLES	(1) Lfsato8910	(2) Lfsato8910	(3) Lfsato8910	(4) Lfsato8910	(5) Lfsato8910
Log Nei genetic dist. from DK	-5.40***	-0.90	-5.69***	-5.20**	-3.34*
	(1.14)	(1.45)	(2.07)	(2.26)	(1.80)
GDP diff from DK		-8.35***	-7.98***	-8.56***	-8.38***
		(1.54)	(1.66)	(1.80)	(1.77)
Log geographic dist. from DK			8.56***	7.44*	
			(2.27)	(4.24)	
Africa				4.58	11.3*
				(6.96)	(6.58)
America				9.72	22.5***
				(8.22)	(4.72)
Asia				-4.09	4.61
				(6.37)	(5.05)
Oceania				-6.26	13.0*
				(12.9)	(7.09)
Constant	62.7***	56.6***	5.55	12.6	59.7***
	(4.83)	(4.84)	(14.3)	(26.9)	(5.33)
Observations	86	86	86	86	86
R-squared	0.153	0.355	0.462	0.521	0.499

Table 5: *Life-Satisfaction Equations for a Sample of 86 Nations*

(Life satisfaction, denoted 'Lfsato' here, is the simple mean of life satisfaction. Source: WVS data)

VARIABLES	(1) Lfsato	(2) Lfsato	(3) Lfsato	(4) Lfsato	(5) Lfsato
Log Nei genetic dist. from DK	-0.28***	0.014	-0.30***	-0.31***	-0.16*
	(0.070)	(0.074)	(0.10)	(0.11)	(0.095)
Log GDP diff. from DK		-0.54***	-0.52***	-0.52***	-0.50***
		(0.083)	(0.087)	(0.100)	(0.097)
Log geographic dist. from DK			0.57***	0.57***	
			(0.12)	(0.22)	
Africa				0.017	0.53
				(0.41)	(0.40)
America				0.36	1.35***
				(0.44)	(0.25)
Asia				-0.16	0.51*
				(0.38)	(0.30)
Oceania				-0.75	0.73*
				(0.68)	(0.38)
Constant	7.66***	7.26***	3.85***	3.86***	7.48***
	(0.28)	(0.25)	(0.76)	(1.42)	(0.28)
Observations	86	86	86	86	86
R-squared	0.121	0.375	0.520	0.554	0.515

Table 6: *Struggling Equations for a Sample of 128 Nations (with the HDI Human Development Index and social benefits as control variables)*

(Gallup data)

VARIABLES	(1) Struggling	(2) Struggling	(3) Struggling	(4) Struggling	(5) Struggling
Log Nei genetic dist. from DK	7.11***	3.07**	3.78**	3.61**	2.96**
	(0.68)	(1.45)	(1.54)	(1.53)	(1.40)
Log HDI diff. from DK		5.87***	4.79***	3.71***	4.56***
		(1.63)	(1.55)	(1.34)	(1.38)
Log soc. benefits diff. from DK			0.99	0.92	
			(1.21)	(0.77)	
Africa				8.66**	6.64**
				(3.39)	(3.02)
America				-12.7***	-13.5***
				(2.69)	(2.51)
Asia				3.51	3.13
				(3.10)	(2.79)
Oceania				-16.0***	-15.5***
				(2.50)	(2.32)
Constant	31.4***	60.0***	51.3***	50.0***	58.1***
	(2.98)	(8.91)	(10.6)	(8.28)	(7.40)
Observations	131	128	92	92	128
R-squared	0.359	0.492	0.466	0.712	0.713

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: *Thriving Equations for a Sample of 131 Nations (with culture and religion variables as controls)*

VARIABLES	(1) Thriving	(2) Thriving	(3) Thriving	(4) Thriving
Log Nei genetic dist. from DK	-6.46***	-7.56***	-7.84***	-7.25***
	(1.75)	(1.52)	(1.48)	(2.61)
Log GDP diff. from DK	-5.87***	-5.67***	-5.62***	-5.95*
	(1.05)	(1.01)	(1.03)	(2.93)
Log geographic dist. from DK	1.38	2.03	2.03	0.77
	(3.41)	(3.10)	(3.41)	(7.92)
Differences in the % of				
Catholics	14.0**	12.5	11.4	23.6*
	(5.44)	(7.86)	(8.14)	(13.1)
Protestants	-42.4***	-50.7***	-52.0***	-59.8***
	(12.1)	(11.7)	(12.9)	(19.0)
Other Chr.	32.5***	28.5***	28.7***	14.5
	(8.77)	(10.6)	(10.7)	(20.0)
Orthodoxs	6.57	8.90	8.46	21.2
	(7.31)	(8.85)	(8.99)	(19.8)
Jews	48.9***	45.9***	45.9***	46.9**
	(4.87)	(7.33)	(7.36)	(21.4)
Muslims	15.7***	15.7**	15.6**	22.8
	(5.73)	(7.58)	(7.77)	(14.7)
Buddists	10.1	14.2	14.3	23.2*
	(6.97)	(11.7)	(11.6)	(13.1)
Hinduists	2.15	1.55	1.74	3.80
	(5.68)	(9.23)	(8.97)	(13.4)
Others	24.1***	25.5**	27.2**	37.1
	(8.55)	(10.8)	(11.5)	(67.3)
Language dist.			-1.74	-3.65
			(3.37)	(4.29)
Log Differences in Hofstede index of				
Individualism				-2.22
				(3.05)
Power distance				-2.53
				(4.49)
Uncertainty avoidance				-2.94
				(4.22)
Masculinity				-1.36
				(2.08)
Constant	69.0***	74.4***	77.4***	113**
	(20.6)	(19.3)	(23.5)	(43.6)
Colonial origin	No	Yes	Yes	Yes
Observations	131	131	124	60
R-squared	0.808	0.855	0.850	0.903

Table 8: *Suffering Equations for a Sample of 131 Nations (with culture and religion variables as controls)*

VARIABLES	(1) Suffering	(2) Suffering	(3) Suffering	(4) Suffering
Log Nei genetic dist. from DK	3.82***	4.28***	4.28***	4.54**
	(1.03)	(1.23)	(1.16)	(1.85)
Log GDP diff. from DK	3.17***	3.11***	3.08***	5.70**
	(0.94)	(0.89)	(0.92)	(2.29)
Log geographic dist. from DK	-4.66**	-6.22**	-7.49***	-6.14
	(1.94)	(2.43)	(2.53)	(4.79)
Differences in the % of				
Catholics	5.68	7.38	8.29	-7.47
	(4.51)	(4.94)	(5.15)	(7.16)
Protestants	1.97	6.34	6.07	13.9
	(6.06)	(5.42)	(5.87)	(9.21)
Other Chr.	-7.55	-6.37	-6.31	-1.29
	(6.98)	(7.48)	(7.58)	(10.9)
Orthodoxs	16.6**	16.7**	17.1**	-4.25
	(6.59)	(7.19)	(7.05)	(13.6)
Jews	-7.60*	-6.55	-8.01	-22.5
	(4.03)	(4.90)	(5.02)	(14.8)
Muslims	-3.14	-0.46	-0.77	-13.1
	(4.60)	(4.75)	(4.85)	(10.3)
Buddists	-2.77	-0.94	0.16	-18.9**
	(9.15)	(11.2)	(11.1)	(8.88)
Hinduists	-0.62	-0.67	-1.02	-3.67
	(7.63)	(9.41)	(9.27)	(10.1)
Others	-6.84	1.70	4.11	-46.6
	(9.44)	(9.49)	(9.20)	(35.5)
Language dist.			-0.40	-1.30
			(1.90)	(2.67)
Log Differences in Hofstede index of				
Individualism				0.69
				(2.00)
Power distance				-1.14
				(2.79)
Uncertainty avoidance				4.50
				(3.37)
Masculinity				0.23
				(1.10)
Constant	26.2**	32.5**	41.4**	20.8
	(12.5)	(14.8)	(17.4)	(29.1)
Colonial origin	No	Yes	Yes	Yes
Observations	131	131	124	60
R-squared	0.472	0.554	0.560	0.776

Table 9: *Struggling Equations for a Sample of 131 Nations (with culture and religion variables as controls)*

VARIABLES	(1) Struggling	(2) Struggling	(3) Struggling	(4) Struggling
Log Nei genetic dist. from DK	2.63**	3.26***	3.53***	2.81**
	(1.12)	(0.94)	(0.95)	(1.35)
Log GDP diff. from DK	2.63***	2.48***	2.45***	0.25
	(0.92)	(0.82)	(0.82)	(2.26)
Log geographic dist. from DK	3.32	4.29**	5.55**	5.08
	(2.26)	(2.13)	(2.17)	(4.42)
Differences in the % of				
Catholics	-19.7***	-19.6***	-19.4***	-15.1*
	(5.19)	(5.30)	(5.34)	(8.60)
Protestants	40.5***	43.9***	45.5***	45.5***
	(8.64)	(8.76)	(9.52)	(12.1)
Other Chr.	-25.3***	-21.8**	-22.1**	-12.4
	(8.54)	(8.54)	(8.71)	(14.1)
Orthodoxs	-23.4***	-25.1***	-25.0***	-16.8
	(5.58)	(5.51)	(5.63)	(11.3)
Jews	-41.4***	-39.1***	-37.6***	-23.7*
	(4.50)	(5.56)	(5.58)	(13.3)
Muslims	-12.7**	-15.1***	-14.7**	-9.14
	(5.28)	(5.44)	(5.64)	(8.30)
Buddists	-7.37	-13.0*	-14.3**	-3.85
	(5.97)	(6.70)	(6.94)	(9.85)
Hinduists	-1.48	-0.30	-0.12	0.24
	(6.56)	(6.30)	(6.21)	(10.6)
Others	-17.5*	-27.4***	-31.5***	11.4
	(10.3)	(10.3)	(10.3)	(44.0)
Language dist.			2.12	4.85
			(2.46)	(2.90)
Log Differences in Hofstede index of				
Individualism				1.49
				(1.74)
Power distance				3.80
				(3.08)
Uncertainty avoidance				-1.71
				(2.47)
Masculinity				0.95
				(1.40)
Constant	4.59	-7.30	-19.2	-31.1
	(13.7)	(13.7)	(14.2)	(24.4)
Colonial origin	No	Yes	Yes	Yes
Observations	131	131	124	60
R-squared	0.792	0.845	0.850	0.905

Table 10: *A Check that the Nei Measure Correlates with Adjusted Well-being Rankings in the Existing Published Literature*

VARIABLES	(1) Residual_Helliwell	(2) Residual_Ott	(3) Residual_Hudson
Log Nei genetic dist. from DK	-0.15***	-0.17***	-0.26***
	(0.047)	(0.062)	(0.070)
Log geographic dist. from DK	0.24***	0.36***	0.53***
	(0.069)	(0.091)	(0.13)
Constant	0.63	-2.28***	-3.40***
	(0.43)	(0.55)	(0.82)
Observations	143	91	24
R-squared	0.078	0.177	0.472

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variables in the three columns are unexplained country-residuals from international well-being equations estimated in the work of, respectively, John Helliwell and Shun Wang, Jan Ott, and John Hudson. We are deeply grateful to these scholars for their assistance and for providing their data so generously. See also Ott's dissertation (2012).

Table 11: *Struggling Equations for a Sample of 28 Nations (with HTTLPR5 as a control)*

VARIABLES	(1) Struggling	(2) Struggling	(3) Struggling	(4) Struggling
Log Nei genetic dist. from DK	1.69	5.20***	2.60	4.88*
	(2.35)	(1.63)	(2.55)	(2.76)
Log HTTLPR5 dist.	7.57**		9.07***	
	(3.19)		(2.62)	
Log GDP diff from DK			6.24***	8.10**
			(2.21)	(3.32)
Log geographic dist. from DK			-5.35***	-2.17
			(1.63)	(2.94)
Constant	32.0***	33.3***	65.1***	47.0**
	(5.89)	(6.08)	(12.8)	(18.7)
Colonial Origin	No	Yes	Yes	Yes
Observations	28	28	28	28
R-squared	0.422	0.239	0.575	0.361

Table 12: *Thriving Equations for a Sample of 28 Nations (with HTTLPR5 as a control)*

VARIABLES	(1) Thriving	(2) Thriving	(3) Thriving	(4) Thriving
Log Nei genetic dist. from DK	-4.85	-7.45***	-6.40	-8.46**
	(3.92)	(1.95)	(4.23)	(3.69)
Log HTTLPR5 dist.	-5.61		-8.18*	
	(5.90)		(4.61)	
Log GDP diff from DK			-11.5***	-13.2***
			(3.13)	(4.11)
Log geographic dist. from DK			9.46***	6.59
			(2.68)	(3.95)
Constant	66.5***	65.5***	8.05	24.3
	(7.26)	(7.35)	(18.4)	(24.6)
Observations	28	28	28	28
R-squared	0.304	0.252	0.559	0.470

Table 13: *A Check on Whether the Current Well-being of Nations is Correlated with the Reported Well-being of Americans who have Ancestors from that Nation.*

Source of data on American happiness: General Social Surveys.

VARIABLES	(1) Struggling	(2) Thriving	(3) Suffering	(4) Lfsato8910
Happiness of US-born with different origins	-46.5**	70.4***	-24.4**	62.9**
	(16.8)	(25.1)	(10.1)	(24.6)
Constant	58.1***	27.9***	14.1***	38.8***
	(2.63)	(4.18)	(2.01)	(4.82)
Observations	29	29	29	29
R-squared	0.154	0.165	0.121	0.184

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$