

BETSEY STEVENSON
University of Pennsylvania

JUSTIN WOLFERS
University of Pennsylvania

Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox

ABSTRACT The “Easterlin paradox” suggests that there is no link between a society’s economic development and its average level of happiness. We reassess this paradox, analyzing multiple rich datasets spanning many decades. Using recent data on a broader array of countries, we establish a clear positive link between average levels of subjective well-being and GDP per capita across countries, and find no evidence of a satiation point beyond which wealthier countries have no further increases in subjective well-being. We show that the estimated relationship is consistent across many datasets and is similar to that between subjective well-being and income observed within countries. Finally, examining the relationship between changes in subjective well-being and income over time within countries, we find economic growth associated with rising happiness. Together these findings indicate a clear role for absolute income and a more limited role for relative income comparisons in determining happiness.

Economic growth has long been considered an important goal of economic policy, yet in recent years some have begun to argue against further trying to raise the material standard of living, claiming that such increases will do little to raise well-being. These arguments are based on a key finding in the emerging literature on subjective well-being, called the “Easterlin paradox,” which suggests that there is no link between the level of economic development of a society and the overall happiness of its members. In several papers Richard Easterlin has examined the relationship between happiness and GDP both across countries and within

individual countries through time.¹ In both types of analysis he finds little significant evidence of a link between aggregate income and average happiness.

In contrast, there is robust evidence that *within* countries those with more income are happier. These two seemingly discordant findings—that income is an important predictor of individual happiness, yet apparently irrelevant for average happiness—have spurred researchers to seek to reconcile them through models emphasizing reference-dependent preferences and relative income comparisons.² Richard Layard offers an explanation: “people are concerned about their relative income and not simply about its absolute level. They want to keep up with the Joneses or if possible to outdo them.”³ While leaving room for absolute income to matter for some people, Layard and others have argued that absolute income is only important for happiness when income is very low. Layard argues, for example, that “once a country has over \$15,000 per head, its level of happiness appears to be independent of its income per head.”⁴

The conclusion that absolute income has little impact on happiness has far-reaching policy implications. If economic growth does little to improve social welfare, then it should not be a primary goal of government policy. Indeed, Easterlin argues that his analysis of time trends in subjective well-being “undermine[s] the view that a focus on economic growth is in the best interests of society.”⁵ Layard argues for an explicit government policy of maximizing subjective well-being.⁶ Moreover, he notes that relative income comparisons imply that each individual’s labor effort imposes negative

1. Easterlin (1974, 1995, 2005a, 2005b).

2. Easterlin (1973, p. 4) summarizes his findings: “*In all societies, more money for the individual typically means more individual happiness. However, raising the incomes of all does not increase the happiness of all.* The happiness-income relation provides a classic example of the logical fallacy of composition—*what is true for the individual is not true for society as a whole.* The resolution of this paradox lies in the relative nature of welfare judgments. Individuals assess their material well-being, not in terms of the absolute amount of goods they have, but relative to a social norm of what goods they ought to have” (italics in original). Layard (1980, p. 737) is more succinct: “a basic finding of happiness surveys is that, though richer societies are not happier than poorer ones, within any society happiness and riches go together.” For a recent review of the use of reference-dependent preferences to explain these observations, see Clark, Frijters, and Shields (2008).

3. Layard (2005a, p. 45).

4. Layard (2003, p. 17). For other arguments proposing a satiation point in happiness, see Veenhoven (1991), Clark, Frijters, and Shields (2008), and Frey and Stutzer (2002).

5. Easterlin (2005a, p. 441).

6. Layard (2005a). For a concurring view from the positive psychology movement, see Diener and Seligman (2004).

externalities on others (by shifting their reference points) and that these distortions would be best corrected by higher taxes on income or consumption.

Evaluating these strong policy prescriptions demands a robust understanding of the true relationship between income and well-being. Unfortunately, the present literature is based on fragile and incomplete evidence about this relationship. At the time the Easterlin paradox was first identified, few data were available to allow an assessment of subjective well-being across countries and through time. The difficulty of identifying a robust GDP-happiness link from scarce data led some to confound the absence of evidence of such a link with evidence of its absence.

The ensuing years have seen an accumulation of cross-country data recording individual life satisfaction and happiness. These recent data (and a reanalysis of earlier data) suggest that the case for a link between economic development and happiness is quite robust. The key to our findings is a resolute focus on the magnitude of the subjective well-being-income gradient estimated within and across countries at a point in time as well as over time, rather than its statistical significance or insignificance.

Our key result is that the estimated subjective well-being-income gradient is not only significant but also remarkably robust across countries, within countries, and over time. These comparisons between rich and poor members of the same society, between rich and poor countries, and within countries through time as they become richer or poorer all yield similar estimates of the well-being-income gradient. Our findings both put to rest the earlier claim that economic development does not raise subjective well-being and undermine the possible role played by relative income comparisons.

These findings invite a sharp reassessment of the stylized facts that have informed economic analysis of subjective well-being data. Across the world's population, variation in income explains a sizable proportion of the variation in subjective well-being. There appears to be a very strong relationship between subjective well-being and income, which holds for both rich and poor countries, falsifying earlier claims of a satiation point at which higher GDP per capita is not associated with greater well-being.

The rest of this paper is organized as follows. The first section provides some background on the measurement of subjective well-being and economic analysis of these data. Subsequent sections are organized around alternative measurement approaches to assessing the link between income and well-being. Thus, the second section compares average well-being and income across countries. Whereas earlier studies focused on comparisons of small numbers of industrialized countries, newly available data allow comparisons across countries at all levels of development. These comparisons

show a powerful effect of national income in explaining variation in subjective well-being across countries. In the third section we confirm the earlier finding that richer people within a society are typically happier than their poorer brethren. Because these national cross sections typically involve quite large samples, this finding is extremely statistically significant and has not been widely disputed. However, Easterlin and others have argued strongly that the positive relationship between income and subjective well-being *within* countries is much larger than that seen across countries.⁷ This argument is not borne out by the data: the well-being-income gradient measured within countries is similar to that measured between countries. The paper's fourth section extends our analysis to assessing national time-series movements in average well-being and income. Consistent time series measuring subjective well-being data are scarce, and the existing data are noisy. These factors explain why past researchers have not found a link between economic growth and growth in happiness. We reexamine three of the key case studies from previous research and find that a more careful assessment of the experiences of Japan, Europe, and the United States does not undermine the claim of a clear link between economic growth and happiness, a finding supported by repeated international cross-sections. Our point estimates suggest that the link may be similar to that found in cross-country comparisons, although substantial uncertainty remains around these estimates. The fifth section briefly explores alternative measures of well-being.

Some Background on Subjective Well-Being and Income

Our strategy in this paper is to use all of the important large-scale surveys now available to assess the relationship between subjective well-being and happiness. These surveys typically involve questions probing happiness or life satisfaction. The World Values Survey, for example, asks, "Taking all things together, would you say you are: very happy; quite happy; not very happy; not at all happy?" and, "All things considered, how satisfied are you with your life as a whole these days?" Other variants of the question, such as that in the Gallup World Poll, employ a ladder analogy: interviewees are asked to imagine a ladder with each rung representing a successively better life. Respondents then report the "step" on the ladder that best represents their life.

7. Easterlin (1974).

These questions (and many other variants) are typically clustered under the rubric of “subjective well-being.”⁸ Although the validity of these measures remains a somewhat open question, a variety of evidence points to a robust correlation between answers to subjective well-being questions and more objective measures of personal well-being. For example, answers to subjective well-being questions have been shown to be correlated with physical evidence of affect such as smiling, laughing, heart rate measures, sociability, and electrical activity in the brain.⁹ Measures of individual happiness or life satisfaction are also correlated with other subjective assessments of well-being such as independent evaluations by friends, self-reported health, sleep quality, and personality.¹⁰ Subjective well-being is a function of both the individual’s personality and his or her reaction to life events. One would therefore expect an individual’s happiness to be somewhat stable over time, and accurate measurements of subjective well-being to have high test-retest correlations, which indeed they do.¹¹ Self-reports of happiness have also been shown to be correlated in the expected direction with changes in life circumstances. For example, an individual’s subjective well-being typically rises with marriage and income growth and falls while going through a divorce.

Although the results from each of these approaches suggest that cross-sectional comparisons of people within a population have some validity, there is less evidence about the validity of comparisons across populations, which can be confounded by translation problems and cultural differences. Many researchers have argued for the possibility of a biologically based set of emotions that are universal to humans and appear in all cultures.¹² Research has found that people across cultures clearly recognize emotions such as anger, sadness, and joy when displayed in others’ facial expressions.¹³ Studies have also found that when people around the globe are asked about what is required for more happiness or life satisfaction, the

8. Diener (2006, pp. 399–400) suggests that “Subjective well-being refers to all of the various types of evaluations, both positive and negative, that people make of their lives. It includes reflective cognitive evaluations, such as life satisfaction and work satisfaction, interest and engagement, and affective reactions to life events, such as joy and sadness. Thus, subjective well-being is an umbrella term for the different valuations people make regarding their lives, the events happening to them, their bodies and minds, and the circumstances in which they live.”

9. Diener (1984).

10. Diener, Lucas, and Scollon (2006); Kahneman and Krueger (2006).

11. Eid and Diener (2004).

12. Diener and Tov (2007).

13. Ekman and Friesen (1971); Ekman and others (1987).

answers are strikingly uniform: money, health, and family are said to be the necessary components of a good life.¹⁴ Ed Diener and William Tov argue that it is this possibility of biologically based universal emotions that suggests that well-being can be compared across societies.¹⁵

A similar argument applies to making comparisons of subjective well-being within countries over time. One difficulty with time-series assessments is the possibility that small changes in how people perceive or answer questions about their happiness may be correlated with changes in the outcomes—such as income—whose relationship with subjective well-being one wishes to assess. The evidence regarding aggregate changes in happiness over time is inconsistent. Aggregate happiness has been shown to fall when unemployment and inflation rise, and to move in the expected direction with the business cycle.¹⁶ However, on average, women in both the United States and Europe report declining happiness relative to men over recent decades, a finding that is difficult to reconcile with changes in objective conditions.¹⁷ Finally, the present paper is motivated by a desire to better understand the failure of past studies to isolate a link between happiness and economic growth.

A largely underacknowledged problem in making intertemporal comparisons is simply the difficulty in compiling sufficiently comparable data. For instance, Tom Smith shows that small changes in the ordering of questions on the U.S. General Social Survey led to large changes in reported happiness.¹⁸ These same data seem to show important day-of-week and seasonal cycles as well. Another difficulty with intertemporal comparisons is that attempts to cobble together long time series (such as for Japan, the United States, or China) often involve important coding breaks. Many of these issues simply add measurement error, making statistically significant findings more difficult to obtain. However, when scarce data are used to make strong inferences about changes in well-being over decades, even small amounts of measurement error can lead to misleading inferences.

To date, much of the economics literature assessing subjective well-being has tended to use measures of “life satisfaction” and “happiness” interchangeably. The argument for doing so is that these alternative measures of well-being are highly correlated and have similar covariates. However, they capture somewhat different concepts, with happiness more

14. Easterlin (1974).

15. Diener and Tov (2007).

16. Di Tella, MacCulloch, and Oswald (2003); Wolfers (2003).

17. Stevenson and Wolfers (2007).

18. Smith (1986).

related to affect whereas satisfaction is more evaluative. The psychology literature has tended to treat questions probing affect as distinct from more evaluative assessments. We will consider both the income-happiness and income-satisfaction links in parallel. A subtle measurement issue is also involved, in that many of the surveys asking individuals about their happiness provide a shorter scale of answers (such as “very happy,” “pretty happy,” and “not so happy”) than do those asking typical life satisfaction questions (which often use the “ladder” technique described above).

A final measurement issue to consider is the likely functional form of the relationship between subjective well-being and income. Most early studies considered the relationship between the level of absolute income and the level of happiness, and thus often found a curvilinear relationship. In some cases the lack of evidence of a clear linear relationship between GDP per capita and happiness led to theories of a satiation point, beyond which more income would not increase happiness. A more natural starting point might be to represent well-being as a function of the logarithm of income rather than absolute income. And indeed, recent research has shown that *within* countries “the supposed attenuation at higher income levels of the happiness-income relation does not occur when happiness is regressed on log income, rather than absolute income.”¹⁹ However, if happiness is linearly related to log income in the within-country cross section, then cross-country studies should also examine the relationship between average levels of subjective well-being and average levels of log income. If economic development raises individual incomes equiproportionately, then average log income will rise or fall in tandem with the log of average income. Thus, most of our analysis assesses the relationship across countries between well-being and the log of GDP per capita, which is (surprisingly enough) a departure from much of the literature.²⁰ Throughout our analysis we make heavy use of bivariate scatterplots and nonparametric regression techniques in order to allow the reader to assess the appropriate functional forms visually.

Finally, as in the existing literature, our analysis of the relationship between happiness and income involves an assessment of correlations rather than an attempt to establish tight causal links. Thus, our aim is simply to sort out the stylized facts about the link between income and well-being. Several interesting variants of the question could be asked—such as

19. Easterlin (2001, p. 468).

20. Previous authors examining the relationship between well-being and log GDP include Easterlin (1995), Leigh and Wolfers (2006), and Deaton (2008); however, explicit discussion of the appropriate functional form is quite rare.

whether it is GDP, broader measures of economic development, or alternatively, changes in output or in productivity that drive happiness. Unfortunately, we lack the statistical power to resolve these questions.

Cross-Country Comparisons of Income and Well-Being

In his seminal 1974 paper, Easterlin asked whether “richer countries are happier countries.”²¹ Examining two international datasets, he found a relationship across countries between aggregate happiness and income that he described as “ambiguous” and, although perhaps positive, small.²² Subsequent research began to show a more robust positive relationship between a country’s income and the happiness of its people, leading Easterlin to later conclude that “a positive happiness-income relationship typically turns up in international comparisons.”²³ However, this relationship has been argued as prevailing only over low levels of GDP per capita; once wealthy countries have satisfied basic needs, they have been described as on the “‘flat of the curve,’ with additional income buying little if any extra happiness.”²⁴ Although the literature has largely settled on the view that aggregate happiness rises with GDP for low-income countries, there is much less consensus on the magnitude of this relationship, or on whether a satiation point exists beyond which further increases in GDP per capita are associated with no change in aggregate happiness.²⁵

The early cross-country studies of income and happiness tended to be based on only a handful of countries, often with rather similar income per capita, and hence did not lend themselves to definitive findings. In addition, as the relationship between subjective well-being and the log of income is approximately linear, the analysis in terms of absolute levels of GDP per capita likely contributed to the lack of clarity around the relationship between income and happiness among wealthier countries. As we will show, new large-scale datasets covering many countries point to a clear, robust relationship between GDP per capita and average levels of subjective well-being in a country. Furthermore, we find no evidence that coun-

21. Easterlin (1974, p. 104).

22. Easterlin (1974, p. 108).

23. Easterlin (1995, p. 42).

24. Clark, Frijters, and Shields (2008, p. 96).

25. Deaton (2008) finds no evidence of a satiation point. His analysis of the 2006 Gallup World Poll finds a strong relationship between log GDP and happiness that is, if anything, stronger among high-income countries.

tries become satiated—the positive income-happiness relationship holds for both developed and developing nations.

Our macroeconomic analysis focuses on measures of real GDP per capita measured at purchasing power parity. For most countries we use the most recent data from the World Bank’s World Development Indicators database; where we are missing data, we refer to the Penn World Tables (version 6.2) and, failing that, the CIA Factbook. For earlier years we use data from Angus Maddison.²⁶ The average of log income per person may be a more desirable aggregate than the log of average income, and so in some specifications we also account for the difference between these measures (also known as the mean log deviation).

Measuring average levels of subjective well-being is somewhat more difficult, as this typically involves aggregating individual responses to a qualitative question. Moreover, we wish to make comparisons across surveys that contain subjective well-being questions with varying numbers of categories for the responses. To do this, we need to convert the subjective well-being measures to a normalized measure, which we do through the use of ordered probit regressions of happiness on a series of country (or country-year) fixed effects (with no other controls), and then treat these fixed effects as average levels of well-being within a country (or country-year).²⁷ Appendix A compares our ordered probit index with four alternative approaches to cardinalizing both life satisfaction and happiness, demonstrating that these alternatives yield highly correlated well-being aggregates. The distinct advantage of the ordered probit is that coefficients can be interpreted relative to the dispersion of the distribution of latent well-being in the population. As such, our ordered probit index should be interpreted as highlighting differences in average levels of happiness or life satisfaction between countries, relative to the pooled within-country standard deviation.

We present our analysis chronologically, so that the reader may see how the literature has progressed. To allow easy visual comparisons, we use a similar scale when graphing happiness and GDP and try to keep this scale consistent throughout the paper.

26. Maddison (2007). When filling in missing years, we interpolate using the annual percentage changes listed in the Penn World Tables. When filling in missing countries, we apply the ratio of a country’s GDP per capita to U.S. GDP per capita, using data from the Penn World Tables or the CIA Factbook, to the World Bank data.

27. Throughout, we use suggested survey weights to ensure that our estimates are nationally representative for each country in each wave.

The top row of graphs in figure 1 shows the three earliest cross-country comparisons of subjective well-being of which we are aware. Each of these comparisons is based on only four to nine countries, which were similar in terms of economic development. As a consequence, these comparisons yield quite imprecise estimates of the link between happiness and GDP. We have provided two useful visual devices to aid in interpretation: a dashed line showing the ordinary least squares (OLS) regression line (our focus), and a shaded area that shows a central part of the happiness distribution, with a width equal to the cross-sectional standard deviation.

The graphs in the second row of figure 1 show the cross-country comparisons presented by Easterlin.²⁸ Analyzing the 1960 data, Easterlin argues that “the association between wealth and happiness indicated by Cantril’s international data is not so clear-cut. . . . The inference about a positive association relies heavily on the observations for India and the United States.”²⁹ Turning to the 1965 World Survey III data, Easterlin argues that “The results are ambiguous. . . . If there is a positive association between income and happiness, it is certainly not a strong one.”³⁰ Rather than highlighting the positive association suggested by the regression line, he argues that “what is perhaps most striking is that the personal happiness ratings for 10 of the 14 countries lie virtually within a half a point of the midpoint rating of 5 [on the raw 0–10 scale]. . . . The closeness of the happiness ratings implies also that a similar lack of association would be found between happiness and other economic magnitudes.”³¹ The clustering of countries within the shaded area on the chart gives a sense of this argument. However, the ordered probit index is quite useful here in quantifying the differences in average levels of happiness across countries relative to the within-country variation. Unlike the raw data, the ordered probit suggests quite large differences in well-being relative to the cross-sectional standard deviation. Similarly, the use of log income rather than absolute income highlights the linear-log relationship. Finally, Easterlin mentions briefly the 1946 and 1949 data shown in the top row of figure 1,

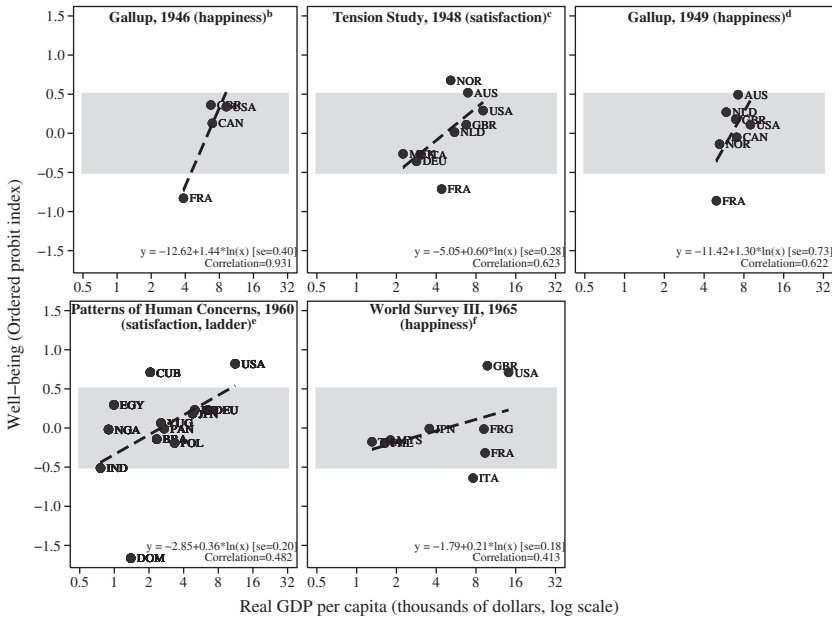
28. Easterlin (1974). We plot the ordered probit index, whereas Easterlin graphs the mean response.

29. Easterlin (1974, p. 105). Following Cantril (1965), Easterlin also notes that “the values for Cuba and the Dominican Republic reflect unusual political circumstances—the immediate aftermath of a successful revolution in Cuba and prolonged political turmoil in the Dominican Republic.”

30. Easterlin (1974, p. 108).

31. Easterlin (1974, p. 106).

Figure 1. Early Cross-Country Surveys of Subjective Well-Being^a



Source: Cantril (1951); Buchanan and Cantril (1953); Strunk (1950); Cantril (1965); Veenhoven (undated); Easterlin (1974, table 7); Maddison (2007).

a. Well-being data are aggregated into an index by running an ordered probit regression of happiness or satisfaction on country fixed effects separately for each survey. Income data were extracted from Maddison (2007) and reflect estimates of real GDP per capita at purchasing power parity in 1990 U.S. dollars. Dashed lines are fitted from OLS regressions of this well-being index on log GDP. Country abbreviations in all figures are standard ISO country codes.

b. Data were extracted from Cantril (1951), who reports on polls by four Gallup affiliates. Countries included are Canada, France, the United Kingdom, and the United States. Respondents were asked, “In general, how happy would you say you are—very happy, fairly happy, or not very happy?”

c. Data were extracted from Buchanan and Cantril (1953), reporting on a UNESCO study of “Tensions Affecting International Understanding.” Countries included are Australia, France, Germany, Italy, the Netherlands, Norway, Mexico, the United Kingdom, and the United States. Respondents were asked, “How satisfied are you with the way you are getting on now?—very, all right, or dissatisfied?”

d. Data were drawn from Strunk (1950). Countries included are Australia, Canada, France, the Netherlands, Norway, the United Kingdom, and the United States. Respondents were asked the same question as in note b.

e. Data were extracted from tabulations by Cantril (1965), as reported in Veenhoven (undated). Countries include Brazil, Cuba, the Dominican Republic, Egypt, Germany, India, Japan, Nigeria, Panama, Poland, the United States, and Yugoslavia; data from the Philippines are missing. Data for the United States were tabulated from the Interuniversity Consortium for Political and Social Research. Surveys were run from 1957 to 1963 using Cantril’s “Self-Anchoring Striving Scale,” which begins by probing about the best and worst possible futures and then shows a picture of a ten-step ladder and asks, “Here is a picture of a ladder. Suppose that we say the top of the ladder [pointing] represents the best possible life for you and the bottom [pointing] represents the worst possible life for you. Where on the ladder [moving finger rapidly up and down ladder] do you feel you personally stand at the present time?”

f. Data were extracted from Easterlin (1974, table 7), who reported cross-tabulations for France, Germany, Italy, Malaysia, the Philippines, Thailand, and the United Kingdom from the World Survey III and added data for the United States from the October 1966 AIPO poll and for Japan from the 1958 survey of Japanese national character. Respondents were asked the same question as in note b. Easterlin reports only the proportion “not very happy” for Japan; hence we infer the well-being index based only on the lower cutpoint of the ordered probit regression run on the eight other countries.

noting that “the results are similar . . . if there is a positive association among countries between income and happiness it is not very clear.”³²

Although the correlation between income and happiness in these early surveys is not especially convincing, this does not imply that income has only a minor influence on happiness, but rather that other factors (possibly including measurement error) also affect the national happiness aggregates. Even so, three of these five datasets suggest a statistically significant relationship between happiness and the natural logarithm of GDP per capita. More important, the point estimates reveal a positive relationship between well-being and income, and a precision-weighted average of these five regression coefficients is 0.45, which is comparable to the sort of well-being-GDP gradient suggested in cross-sectional comparisons of rich and poor people within a society (a theme we explore further below).

We have also located several other surveys from the mid-1960s through the 1970s that show a similar pattern. In particular, the ten-nation “Images of the World in the Year 2000” study, conducted in 1967, and the twelve-nation Gallup-Kettering Survey, from 1975, both yield further evidence consistent with an important and positive well-being-GDP gradient. Subsequent cross-country data collections have become increasingly ambitious, and analysis of these data has made the case for a linear-log relationship between subjective well-being and GDP per capita even stronger, while also largely confirming that the magnitudes suggested by these early studies were quite accurate.

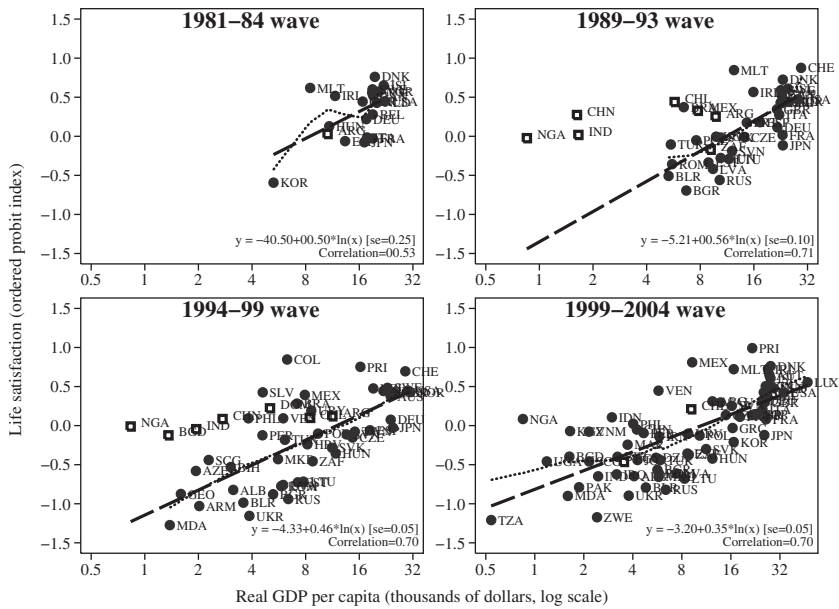
Figure 2 presents data on life satisfaction from each wave of the World Values Survey separately, illustrating the accumulation of new data through time.³³ (We turn to the data on happiness from this survey below, in figure 5.) In the early waves of the survey, the sample consisted mostly of wealthy countries; given the limited variation in income, these samples yielded suggestive, but not definitive, evidence of a link between GDP and life satisfaction. As the sample expanded, the relationship became clearer. In each wave the regression line is upward sloping, and the estimated coefficient is statistically significant and similar across the four waves, with its precision increasing in the later waves. We also plot estimates from locally weighted (or lowess) regressions, to get a sense of whether there are important deviations from the linear-log functional form.³⁴ In the earliest waves the small number of countries and limited heterogeneity in income across

32. Easterlin (1974, p. 108).

33. In order to make these data collections consistent, we analyze only adult respondents.

34. The lowess estimator is a local regression estimator that plots a flexible curve.

Figure 2. Life Satisfaction and Real GDP per Capita: World Values Survey^a



Sources: World Values Survey; authors' regressions. Sources for GDP per capita are described in the text.
 a. Sample includes twenty (1981-84), forty-two (1989-93), fifty-two (1994-99), or sixty-nine (1999-2004) countries; see text for details of the sample. Observations represented by hollow squares are drawn from countries in which the World Values Survey sample is not nationally representative; see appendix B for further details. Respondents are asked, "All things considered, how satisfied are you with your life as a whole these days?" and asks respondents to choose a number from 1 (completely dissatisfied) to 10 (completely satisfied). Data are aggregated into a satisfaction index by running an ordered probit regression of satisfaction on country \times wave fixed effects. Dashed lines are fitted from an OLS regression; dotted lines are fitted from lower regressions. These lines and the reported regressions are fitted only from nationally representative samples. Real GDP per capita is at purchasing power parity in constant 2000 international dollars.

countries made it difficult to make robust inferences about the relationship between life satisfaction and economic development. Nonetheless, pooling data from all four waves and allowing wave fixed effects yields an estimate of the satisfaction-income gradient of 0.40 (with a standard error of 0.04, clustering by country), and an F-test reveals that wave-specific slopes are jointly statistically insignificant relative to a model with a common slope term ($F_{3,78} = 1.98$).

In some cases the expansion of the World Values Survey to include poorer countries resulted in explicitly unrepresentative samples.³⁵ For

35. We thank Angus Deaton for alerting us to these limitations in the World Values Survey.

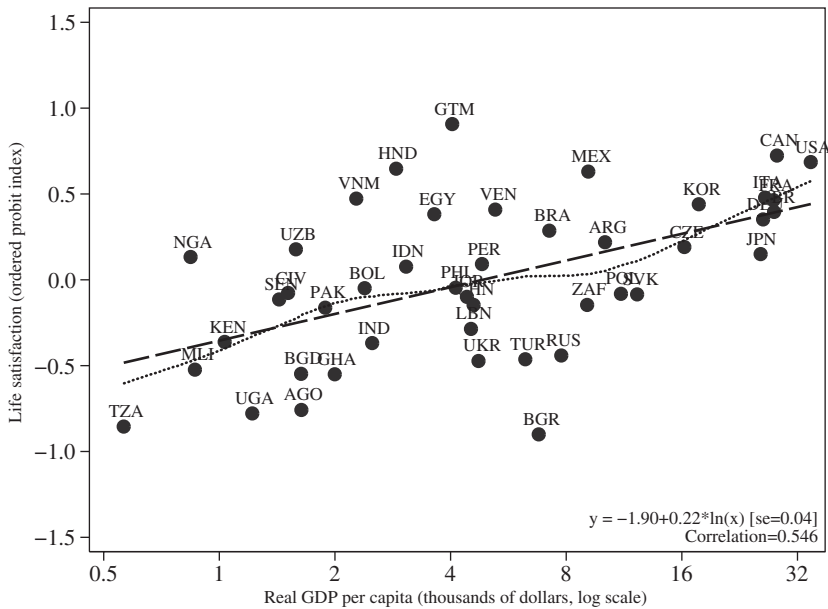
example, Argentina was included in the 1981–84 wave, but the sample was limited to urban areas and was not expanded to become representative of the country overall until the 1999–2004 wave. Chile, China, India, Mexico, and Nigeria were added in the 1989–93 wave, but their samples largely consisted of the more educated members of society and those living in urban areas. These limitations are spelled out clearly in the survey documentation but have been ignored in most subsequent analyses. The nonrepresentative samples typically came from poorer countries and involved sampling richer (and hence likely happier) respondents. Thus, inclusion of these observations imparts a downward bias on estimates of the well-being-income gradient. We therefore exclude from our analysis countries that the survey documentation suggests are clearly not representative of the entire population. Observations for these countries are plotted in figure 2 using hollow squares. As expected, these observations typically sit above the regression line. Appendix B provides a comparison of our results when these countries are included in the analysis, along with greater detail regarding sampling in the World Values Survey.

Subsequently, the 2002 Pew Global Attitudes Survey interviewed 38,000 respondents in forty-four countries across the development spectrum. The subjective well-being question is a form of Cantril's "Self-Anchoring Striving Scale."³⁶ Respondents were shown a picture and told, "Here is a ladder representing the 'ladder of life.' Let's suppose the top of the ladder represents the best possible life for you; and the bottom, the worst possible life for you. On which step of the ladder do you feel you personally stand at the present time?" Respondents were asked to choose a step along a range of 0 to 10. As before, we run an ordered probit of the ladder ranking on country fixed effects to estimate average levels of subjective well-being in each country, and we compare these averages with the log of GDP per capita in figure 3. These data show a linear relationship similar to that seen in figure 2.

The most ambitious cross-country surveys of subjective well-being come from the 2006 Gallup World Poll. This is a new survey designed to measure subjective well-being consistently across 132 countries. Similar questions were asked in all countries, and the survey contains data for each country that are nationally representative of people aged 15 and older. The survey asks a variety of subjective well-being questions, including a ladder question similar to that used in the 2002 Pew survey. As figure 4 shows, these data yield a particularly close relationship between subjective well-

36. Cantril (1965).

Figure 3. Life Satisfaction and Real GDP per Capita: Pew Global Attitudes Survey^a



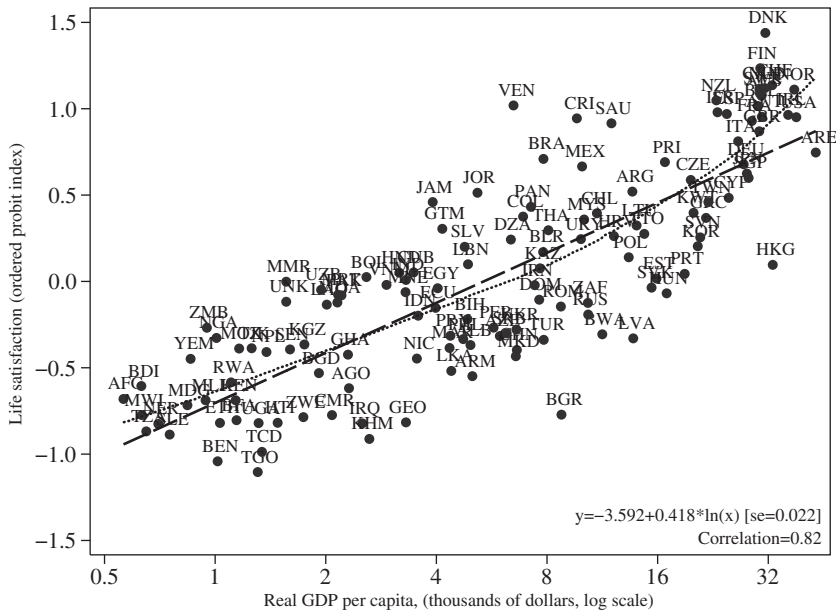
Sources: Pew Global Attitudes Survey, 2002; authors' regressions. Sources for GDP per capita are described in the text.

a. Sample includes forty-four developed and developing countries. Respondents are shown a picture of a ladder with ten steps and asked, "Here is a ladder representing the 'ladder of life.' Let's suppose the top of the ladder represents the best possible life for you; and the bottom, the worst possible life for you. On which step of the ladder do you feel you personally stand at the present time?" Data are aggregated into a satisfaction index by running an ordered probit regression of satisfaction on country fixed effects. Dashed line is fitted from an OLS regression; dotted lines are fitted from lowess regressions. Real GDP per capita is at purchasing power parity in constant 2000 international dollars.

being and the log of GDP per capita. Across the 131 countries for which we have usable GDP data (we omit Palestine), the correlation exceeds 0.8. Moreover, the estimated coefficient on log GDP per capita, 0.42, is similar to those obtained using the World Values Survey, the Pew survey, and the earlier surveys, including those assessed by Easterlin. These findings are also quite similar to those found by Angus Deaton,³⁷ who also shows a linear-log relationship between subjective well-being and GDP per capita using the Gallup World Poll.³⁸ Deaton emphasizes that the clearer

37. Deaton (2008).

38. We estimate a well-being-income gradient that is about half that estimated by Deaton because we have standardized our estimates through the use of ordered probits, whereas Deaton is estimating the relationship between the raw life satisfaction score and log income. Putting both on a similar scale yields similar estimates. Appendix A compares our ordered probit approach with other possible cardinalizations of subjective well-being.

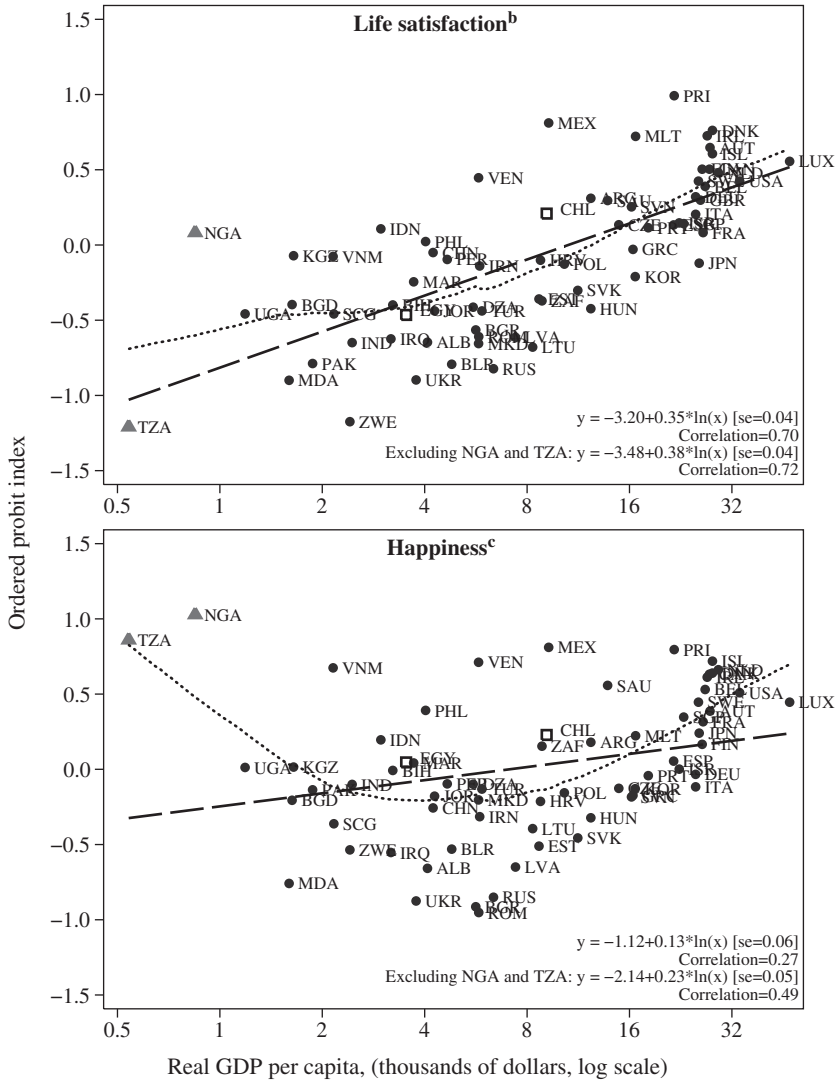
Figure 4. Life Satisfaction and Real GDP per Capita: Gallup World Poll^a

Sources: Gallup World Poll, 2006; authors' regressions. Sources for GDP per capita are described in the text.
 a. Sample includes 131 developed and developing countries. Respondents are asked, "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. Suppose we say that 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?" Dashed line is fitted from the reported ordinary least squares regression; dotted line is fitted from a lower estimation. GDP per capita is at purchasing power parity in constant 2000 international dollars.

relationship in the Gallup data reflects the inclusion of surveys from a greater number of poor countries.

As discussed previously, the economics literature has tended to treat measures of happiness and life satisfaction as largely interchangeable, whereas the psychology literature distinguishes between the two. We now turn to assessing the relationship between measures of happiness and income and comparing these estimates with the estimates of the relationship between measures of life satisfaction and income considered thus far. (We consider additional measures of subjective well-being and their relationship to income in a later section.) Figure 5 investigates both the happiness-GDP link and the life satisfaction-GDP link estimated using the latest wave of the World Values Survey—the life satisfaction data are those discussed above. Happiness is measured using the following question: "Taking all things together, would you say you are: 'very happy,'

Figure 5. Subjective Well-Being and Real GDP per Capita: 1999–2004 World Values Survey^a



Sources: World Values Survey, 1999–2004 wave; authors’ regressions. Sources for GDP per capita are described in the text.

a. Sample includes sixty-nine developed and developing countries. Observations represented by hollow squares are drawn from countries in which the World Values Survey sample is not nationally representative; see appendix B for further details. Dashed lines are fitted from the reported OLS regression; dotted lines are fitted from lowess regressions; both regressions are based only on nationally representative samples. GDP per capita is at purchasing power parity in constant 2000 international dollars.

b. Life satisfaction question asks, “All things considered, how satisfied are you with your life as a whole these days?” and asks respondents to choose a number from 1 (dissatisfied) to 10 (satisfied). Data are aggregated into a satisfaction index by running an ordered probit regression of satisfaction on country × wave fixed effects.

c. Happiness question asks, “Taking all things together, would you say you are: ‘very happy,’ ‘quite happy,’ ‘not very happy,’ [or] ‘not at all happy?’” Data are aggregated into a satisfaction index by running an ordered probit regression of happiness on country × wave fixed effects.

‘quite happy,’ ‘not very happy,’ ‘not at all happy?’” The results suggest that these measures may not be as synonymous as previously thought: happiness appears to be somewhat less strongly correlated with GDP than is life satisfaction.³⁹ Although much of the sample shows a clear relationship between log income and happiness, these data yield several particularly puzzling outliers. For example, the two poorest countries in the sample, Tanzania and Nigeria, have the two highest levels of average happiness, yet both have much lower average life satisfaction—indeed, Tanzania reported the lowest average satisfaction of any country.⁴⁰

This apparent noise in the happiness-GDP link partly explains why earlier analyses of subjective well-being data have yielded mixed results. We reran both the happiness and life satisfaction regressions with Tanzania and Nigeria removed, and it turns out that these outliers explain at least part of the puzzle. In the absence of these two countries, the well-being-GDP gradients, measured using either life satisfaction or happiness, turn out to be very similar. Equally, in these data the correlation between happiness and GDP per capita remains lower than that between satisfaction and GDP per capita.

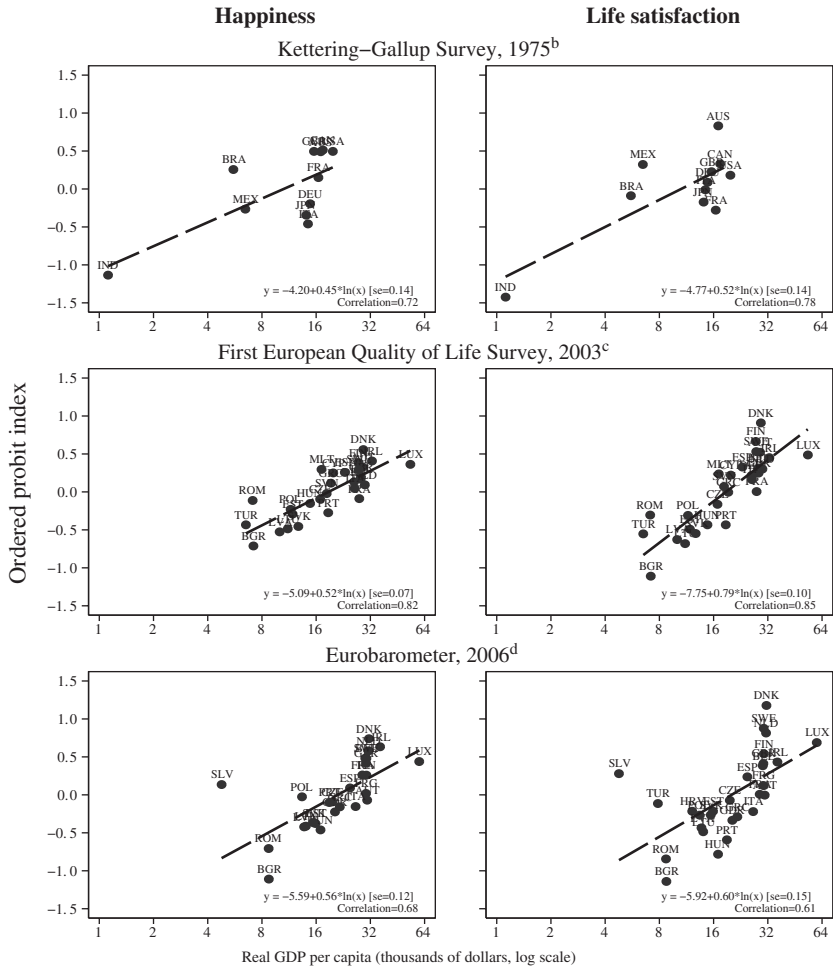
To better understand whether the happiness-GDP gradient systematically differs from the satisfaction-GDP gradient, we searched for other data collections that asked respondents about both happiness and life satisfaction. Figure 6 brings together two such surveys: the 1975 Gallup-Kettering survey and the First European Quality of Life Survey, conducted in 2003. In addition, the bottom panels of figure 6 show data from the 2006 Eurobarometer, which asked about happiness in its survey 66.3 and life satisfaction in survey 66.1. In each case the happiness-GDP link appears to be roughly similar to the life satisfaction-GDP link, although perhaps, as with the World Values Survey, slightly weaker.

Table 1 formalizes all of the analysis discussed thus far with a series of regressions of subjective well-being on log GDP per capita, using data from

39. The contrast in figure 5 probably overstates this divergence, as it plots the data for the 1999–2004 wave of the World Values Survey, whereas table 1 shows that earlier waves yielded a clearer happiness-GDP link.

40. One might suspect that survey problems are to blame, and indeed, the survey notes for Tanzania suggest (somewhat opaquely) that “There were some questions that caused problems when the questionnaire was translated, especially questions related to . . . Happiness because there are different perceptions about it.” We are not aware of any other happiness data for Tanzania, but note that in the 2002 Pew survey Tanzania registered the second-lowest level of average satisfaction among forty-four countries (figure 3). The high levels of happiness recorded in Nigeria seem more persistent: Nigeria also reported the eleventh-highest happiness rating in the 1994–99 wave of the World Values Survey, although it was around the mean in the 1989–93 wave.

Figure 6. Subjective Well-Being and Real GDP per Capita in Selected Surveys^a



Sources: Indicated surveys. Sources for GDP per capita are described in the text.

a. Well-being data are aggregated separately for each indicator in each survey, by running an ordered probit regression of happiness or satisfaction on country fixed effects. Dashed lines are fitted from OLS regressions of this well-being index on log GDP. Real GDP per capita is at purchasing power parity in constant 2000 international dollars.

b. Data were extracted from Veenhoven (undated). Sample includes eleven developed and developing countries. Happiness question asks, “Generally speaking, how happy would you say you are: ‘very happy,’ ‘fairly happy,’ [or] ‘not too happy?’” Life satisfaction question asks, “Now taking everything about your life into account, how satisfied or dissatisfied are you with your life today?” and asks respondents to choose a number from 0 (dissatisfied) to 10 (satisfied).

c. Sample includes twenty-eight European countries. Happiness question asks, “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.” The life satisfaction question asks, “All things considered, how satisfied or dissatisfied are you with your life these days? Please tell me on a scale of 1 to 10, where 1 means very dissatisfied and 10 means very satisfied.”

d. Happiness sample includes thirty European countries drawn from Eurobarometer 66.3. Happiness question asks, “Taking all things together would you say you are: ‘very happy,’ ‘quite happy,’ ‘not very happy,’ [or] ‘not at all happy?’” Life satisfaction sample includes twenty-eight European countries drawn from Eurobarometer 66.1 (missing Croatia and Turkey). The life satisfaction question asks, “On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?”

Table 1. Cross-Country Regressions of Subjective Well-Being on GDP per Capita^a

Survey	Ordered probit regressions, micro data ^b			OLS regressions, national data ^c			Sample size ^e
	Without controls	With controls ^d	All countries	GDP per capita > \$15,000	GDP per capita < \$15,000		
Gallup World Poll, 2006: Ladder question ^f	0.396*** (0.023)	0.422*** (0.023)	0.418*** (0.022)	1.076***† (0.211)	0.348*** (0.037)		139,051 131 countries
World Values Survey: Life satisfaction ^g							
1981–84 wave	0.525** (0.263)	0.291 (0.331)	0.498* (0.252)	1.677** (0.703)	0.722 (0.582)		23,537 (19 countries)
1989–93 wave	0.551*** (0.096)	0.551*** (0.096)	0.558*** (0.096)	0.504 (0.467)	0.391 (0.256)		50,553 (35 countries)
1994–99 wave	0.408*** (0.054)	0.418*** (0.054)	0.462*** (0.051)	0.327 (0.421)	0.394*** (0.084)		65,779 (45 countries)
1999–2004 wave	0.321*** (0.041)	0.329*** (0.041)	0.346*** (0.046)	0.455** (0.223)	0.208** (0.090)		94,224 (67 countries)
Combined, with wave fixed effects	0.373*** (0.038)	0.377*** (0.037)	0.398*** (0.040)	0.477** (0.198)	0.280*** (0.073)		234,093 (79 countries)
World Values Survey: Happiness ^h							
1981–84 wave	0.650*** (0.250)	0.523*** (0.263)	0.569** (0.230)	1.662 (0.987)	0.550 (0.688)		22,294 (18 countries)
1989–93 wave	0.710*** (0.130)	0.725*** (0.128)	0.708*** (0.123)	0.328 (0.475)	0.144 (0.309)		49,281 (35 countries)

1994–99 wave	0.319*** (0.056)	0.335*** (0.056)	0.354*** (0.058)	0.248 (0.235)	0.212** (0.082)	63,785 (46 countries)
1999–2004 wave	0.118* (0.062)	0.138** (0.061)	0.126* (0.073)	0.766***† (0.218)	-0.146 (0.117)	92,799 (66 countries)
Combined, with wave fixed effects	0.229*** (0.055)	0.245*** (0.055)	0.244*** (0.063)	0.612***† (0.170)	-0.015 (0.100)	228,159 (79 countries)
Pew Global Attitudes Survey, 2002: Ladder question ¹	0.223*** (0.041)	0.242*** (0.040)	0.224*** (0.041)	0.466** (0.191)	0.168** (0.082)	37,974 (44 countries)

Source: Authors' regressions.

a. Table reports results of regressions of the indicated measure of well-being on log real GDP per capita. Numbers in parentheses are robust standard errors, clustered by country. Asterisks indicate statistically significant from zero at the *10 percent, **5 percent, and ***1 percent level; † denotes that the coefficient estimate for rich countries is statistically significantly larger than that for poor countries, at the 1 percent level.

b. Ordered probit regressions, using data by respondent, of subjective well-being on log real GDP per capita for the respondent's country, weighting observations to give equal weight to each country × wave.

c. National well-being index is regressed on log real GDP per capita. The well-being index is calculated in a previous ordered probit regression of well-being on country × wave fixed effects.

d. Controls include a quartic in age, interacted with sex, and indicators for missing age or sex.

e. Only nationally representative samples are analyzed, which eliminated seventeen country-wave observations from ten countries in the World Values Survey (see appendix B for further details).

f. Respondents were asked, "Please imagine a ladder with steps numbered from zero at the bottom to ten at the top. Suppose we say that 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?"

g. Respondents were asked, "All things considered, how satisfied are you with your life as a whole these days?" Possible answers range from 1 (dissatisfied) to 10 (satisfied). Respondents were asked, "Taking all things together, would you say you are: (4) very happy; (3) quite happy; (2) not very happy; (1) not at all happy?"

h. Respondents were shown a picture of a ladder with ten steps and asked, "Here is a ladder representing the 'ladder of life.' Let's suppose the top of the ladder represents the best possible life for you, and the bottom, the worst possible life for you. On which step of the ladder do you feel you personally stand at the present time?" Answers are scored from 1 (bottom rung) to 10 (top rung).

the Gallup World Poll, all four waves of the World Values Survey, and the Pew Global Attitudes Survey. The coefficient on log GDP per capita is reported along with its standard error. The first column reports coefficient estimates from ordered probit regressions of individual well-being on the natural log of real GDP per capita, with robust standard errors clustered by country; the second column adds controls for gender and a quartic in age and its interaction with gender. The third column reports the results of a two-stage process: in the first stage we aggregated the data to the country level by running an ordered probit regression of subjective well-being on country fixed effects, which we interpret as a measure of average national happiness. In the second stage we estimated an OLS regression of these country fixed effects on log GDP per capita. The coefficient from this second regression is reported in the third column of table 1. In all the datasets examined, estimates of the relationship obtained from the respondent-level analysis are similar to that obtained through the two-stage process. Moreover, each of these datasets yields remarkably similar estimates of the subjective well-being-GDP gradient, typically centered around 0.4.

The regressions reported in the first three columns of table 1 are performed on the complete sample of countries for each survey; the samples in the remaining two columns consist only of countries with GDP per capita above or below \$15,000 (in 2000 dollars), using the same two-stage process as in the third column, to allow us to assess whether the well-being-GDP gradient differs for rich and poor countries. It has been argued that income is particularly important for happiness when the basic needs of food, clothing, and shelter are not being met, but that beyond this threshold happiness is less strongly related to income. In its stronger form, this view posits a satiation point beyond which more income no longer raises the happiness of a society. For instance, Layard claims that “if we compare countries, there is no evidence that richer countries are happier than poorer ones—so long as we confine ourselves to countries with incomes over \$15,000 per head. . . . At income levels below \$15,000 per head things are different. . . .”⁴¹ Bruno Frey and Alois Stutzer offer a similar assessment of the literature, suggesting that “income provides happiness at low levels of development but once a threshold (around \$10,000) is reached, the average income level in a country has little effect on average subjective well-being.”⁴²

41. Layard (2005b, p. 149).

42. Frey and Stutzer (2002, p. 416).

Employing Layard's cutoff, we find that the relationship between subjective well-being and log GDP per capita is, if anything, stronger rather than weaker in the wealthier countries, although this difference is statistically significant only in a few cases. The point estimates are, on average, about three times as large for those countries with incomes above \$15,000 as for those with incomes below \$15,000.⁴³ We thus find no evidence of a satiation point. Indeed, a consistent theme across the multiple datasets shown in table 1 and figures 1 through 6 appears to be that there is a clear positive relationship between subjective well-being and GDP per capita, even when the comparison is among developed economies only.

The fact that the coefficient on log GDP per capita may be larger for rich countries should be interpreted carefully. Taken at face value, the Gallup results suggest that a 1 percent rise in GDP per capita would have about three times as large an effect on measured well-being in rich as in poor nations.⁴⁴ Of course, a 1 percent rise in U.S. GDP per capita is about ten times as large as a 1 percent rise in Jamaican GDP per capita. Consider instead, then, the effect of a \$100 rise in average incomes in Jamaica and the United States. Such a shock would raise log GDP per capita by ten times more in Jamaica than in the United States, and hence would raise measured well-being by about three times as much in Jamaica as in the United States. For the very poorest countries, this difference is starker. For instance, GDP per capita in Burundi is about one-sixtieth that in the United States; hence a \$100 rise in average income would have a twenty-fold larger impact on measured well-being in Burundi than in the United States.⁴⁵

One explanation for the difference between our findings and earlier findings of a satiation point may be differences in the assumed functional form of the relationship between well-being and GDP. In particular, whereas we have analyzed well-being as a function of log GDP per capita, several previous analyses have focused on the absolute level.⁴⁶ Figure 7

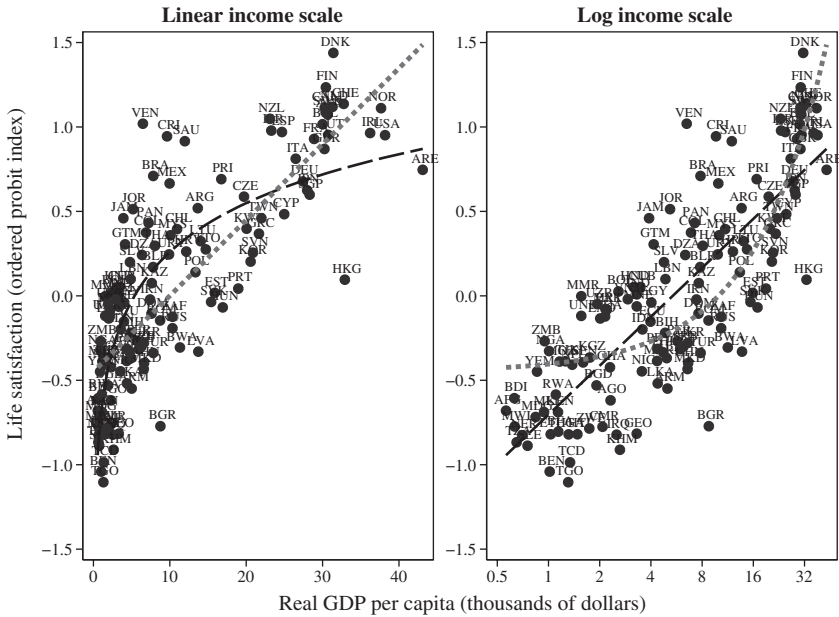
43. This finding is consistent with Deaton (2008).

44. Higher income yields a larger rise in the happiness index, but not necessarily a larger rise in happiness, since we do not know the "reporting function" that translates true hedonic experience into our measured well-being index (Oswald, forthcoming).

45. Using the Gallup World Poll data, we can check whether the log GDP–well-being gradient differs for the very poorest countries. When restricting the sample to countries with GDP per capita below \$3,000, we obtain estimates very similar to those for countries with GDP per capita between \$3,000 and \$15,000. This is also evident in the nonparametric fit shown in figure 4.

46. In a levels specification, the subjective well-being–income gradient is curvilinear and thus is less steep among wealthier countries. Although the slope is never zero, the flattening out of the curve may be more easily misinterpreted as satiation.

Figure 7. Assessing the Functional Form of the Life Satisfaction–GDP Gradient: Gallup World Poll^a



Source: Gallup World Poll, 2006; authors’ regressions. Sources for GDP per capita are described in the text.
 a. Sample includes 131 developed and developing countries. See figure 4 for wording of the question. In each panel the short- and long-dashed lines are fitted from regressions of satisfaction on GDP per capita and the log of GDP per capita, respectively. Real GDP per capita is at purchasing power parity in constant 2000 international dollars.

shows that the log specification yields a better fit, although the difference is small.⁴⁷ Viewed either way, there remains robust evidence of a strongly positive well-being-income link for rich countries. We have reestimated the well-being-GDP relationship using the level of GDP per capita as the independent variable and found that the well-being-GDP gradient is about twice as steep for poor countries as for rich countries. That is, consistent with our earlier findings, a rise in income of \$100 is associated with a rise in well-being for poor countries that is about twice as large as for rich countries. (A 1 percent rise in GDP per capita is associated with

47. Deaton’s (2008, p. 58) assessment of the functional form for the bivariate well-being-GDP relationship led him to conclude that “the relationship between the log of income and life satisfaction offers a reasonable fit for all countries, high-income and low-income, and if there is any evidence for deviation, it is small and in the direction of the slope being higher among the richer countries.”

much larger income gains, and hence much larger well-being gains for rich countries.)

Thus, our conclusion that there is strong evidence against a satiation point is robust to whether one conceives of well-being as rising with log GDP per capita or with its absolute level. As figure 7 demonstrates, even with observations on 131 countries, we have insufficient data to draw particularly strong inferences about the appropriate functional form, although the evidence is certainly suggestive of a linear-log well-being-income relationship. In the next section we turn to within-country comparisons, and given the much larger samples involved, it will be clear that—at least at the individual level—well-being is best thought of as rising in log income. It is this finding that guides our choice of the appropriate functional form for between-country comparisons.

Income and Happiness: Comparing Within-Country and Between-Country Estimates

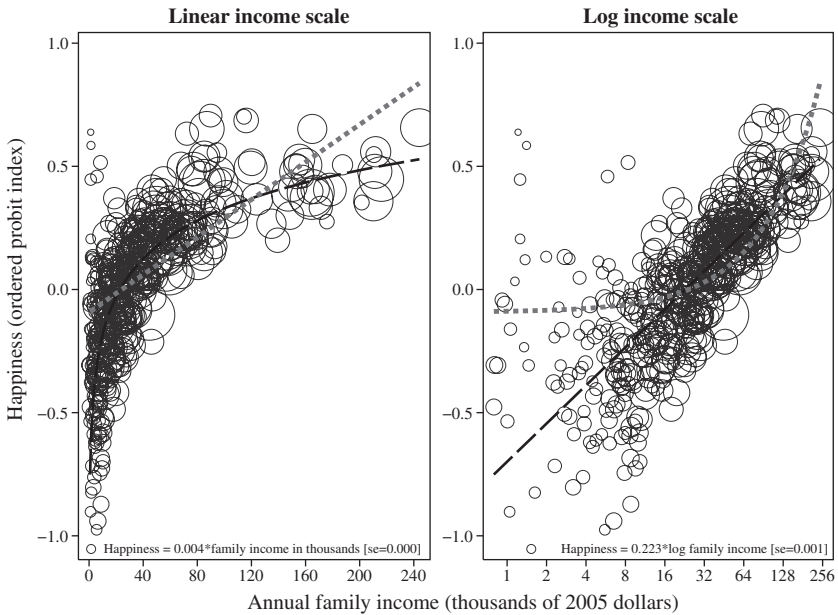
A very simple benchmark for assessing the magnitude of the between-country well-being-GDP gradient measured in the previous section (typically centered around 0.4) is the within-country well-being-income gradient. In particular, Easterlin argued that “the happiness differences between rich and poor countries that one might expect on the basis of the within country differences by economic status are not borne out by the international data.”⁴⁸ Thus, we now turn to comparing the happiness of richer and poorer members of the same society at a single point in time.

On this question there is a clear consensus in the literature, aptly summarized by Easterlin: “As far as I am aware, in every representative national survey ever done a significant bivariate relationship between happiness and income has been found.”⁴⁹ And indeed, we have made similar comparisons in over 100 countries and have yet to find a (statistically significant) exception. Although there has been some debate about the magnitude of this effect, income is clearly an important correlate with happiness. For example, Robert Frank argues for the importance of income for happiness as follows: “When we plot average happiness versus average income for clusters of people in a given country at a given time . . . rich people are in fact a lot happier than poor people. It’s actually an astonishingly large difference. There’s no one single change you can imagine that would

48. Easterlin (1974, pp. 106–07).

49. Easterlin (2001, p. 468).

Figure 8. Assessing the Functional Form of the Happiness-Family Income Gradient: General Social Survey^a



Source: General Social Survey (USA), 1972–2006; authors' regressions.

a. Each circle aggregates income and happiness for one GSS income category in one year, and its diameter is proportional to the population of that income category in that year. The vertical axis in each panel plots the coefficients from an ordered probit regression of happiness on family income category \times year fixed effects; the horizontal axis plots real family income, deflated by the CPI-U-RS. In each panel the short- and long-dashed lines are fitted from regressions of happiness on family income and the log of family income, respectively, weighting by the number of respondents in each income category \times year. Survey question asks, "Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?"

make your life improve on the happiness scale as much as to move from the bottom 5 percent on the income scale to the top 5 percent."⁵⁰

In this spirit we examine the relationship between happiness and income in the United States from 1972 through 2006 using the General Social Survey (GSS), produced by the National Opinion Research Center. Figure 8 plots the coefficients from an ordered probit regression of happiness on income category by year fixed effects against family income. Family income, plotted on the horizontal axis, is converted from income categories by fitting interval regressions to the income data on the assumption that income follows a log-normal distribution.⁵¹ Each circle in the figure

50. Frank (2005, p. 67).

51. We thank Angus Deaton for this suggestion.

represents an income category in a particular year; the diameter of each circle is proportional to the population in the income category in that year. The statistical significance of this relationship is not in doubt, largely because each round of the GSS (as with most happiness surveys) involves over 1,000 respondents. This plot also leaves very little doubt about the functional form: the linear-log relationship between our happiness index and family income is clearly evident throughout the income distribution.⁵²

To further assess the functional form relationship, we investigated the relationship in other datasets for other countries and find similar evidence pointing to a linear-log relationship between subjective well-being and income. In figure 9 we use the Gallup World Poll (as it covers the most countries of any of our datasets) and show estimates from a regression of life satisfaction on separate income category fixed effects for each country, controlling for country fixed effects. We have usable household income data for 113 countries.⁵³ The coefficient estimates on the individual country–household income category fixed effects are plotted against the log of household income, normalized by subtracting off the country average. This figure also points strongly to a linear relationship between subjective well-being and the log of family income, with no evidence of satiation.

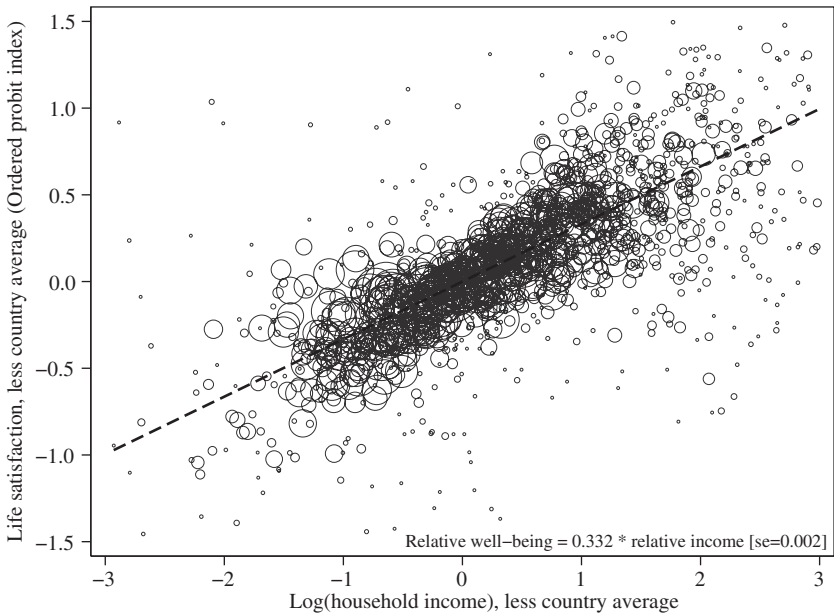
It is the juxtaposition of these statistically significant cross-sectional findings with statistically insignificant cross-country and time-series results that gave rise to the Easterlin paradox. Theories emphasizing relative income comparisons would suggest that the between-country well-being–income gradient would be smaller than the within-country well-being–income gradient (if relative income comparisons are made intranationally). Yet the suggestive comparison of the gradients in figure 4 with figure 9 points to the opposite conclusion: the gradient estimated between countries is larger than that seen within the countries.

Whereas figure 9 plots the gradient seen when examining all of the countries together, it is also worth estimating the within-country well-being–

52. Because the GSS retained the nominal income categories used in 1973, some very low income cells are somewhat off the regression line (the circles to the far left of the graph), reflecting both the fact that small cells yield imprecise happiness estimates and the difficulties in imputing appropriate incomes to the bottom-coded group.

53. We drop Kenya because it lacks labels for income groups, Laos because it contains clearly implausible income groupings, and Uzbekistan because the income categories listed in the data involve overlapping ranges. Respondent-level income data are unavailable for Egypt, Iran, Iraq, Jordan, Kuwait, Latvia, Lebanon, Morocco, Pakistan, Palestine, the Philippines, Saudi Arabia, Sri Lanka, Turkey, the United Arab Emirates, and Yemen. This leaves us with valid household income data for 113 countries.

Figure 9. Within-Country Comparisons of Life Satisfaction and Household Income: Gallup World Poll^a



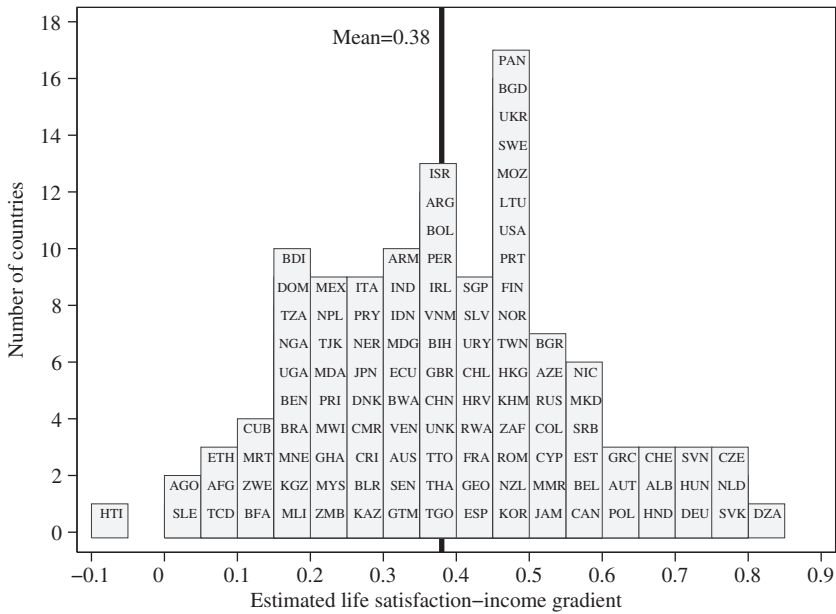
Source: Gallup World Poll, 2006.

a. Each circle aggregates satisfaction in one income category in one country, and its diameter is proportional to the population of that income category in that country. The vertical axis plots the coefficients from an ordered probit regression of life satisfaction on indicator variables for each income category in each country, controlling for country fixed effects; the horizontal axis plots the logarithm of average real household income in each country \times income category, less the country average. The dashed line is from an OLS regression, weighting by the number of respondents in each income category \times country.

income gradient separately for individual countries to see the range of estimated within-country gradients. Thus, for each country we estimate an ordered probit regression of life satisfaction on the natural log of household income, controlling for gender and a quartic in age, entered separately for men and women. The coefficient estimates obtained in each regression (rounded to the nearest 0.05) are displayed in figure 10 as a histogram summarizing the entire sample.⁵⁴ Overall, the average well-being-income gradi-

54. As with the GSS, our various data sources typically report income in categories, rather than as a continuous variable. We follow the same method for each of our datasets, fitting interval regressions to these income data on the assumption that income follows a log-normal distribution. If a dataset contains a bottom income category of zero, we combine it with the succeeding income category. We perform these regressions separately for each country-wave of each dataset.

Figure 10. Distribution of Estimates of the Within-Country Life Satisfaction–Income Gradient^a



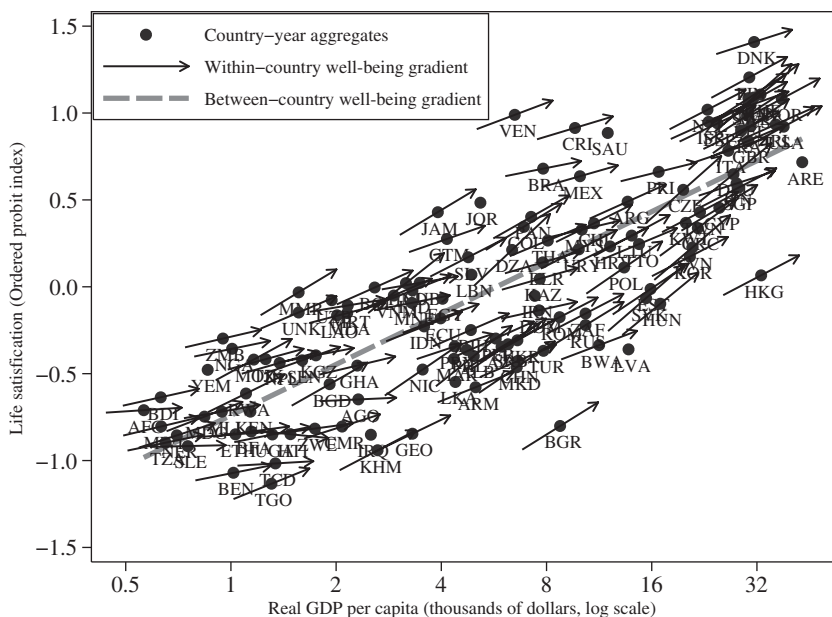
Source: Gallup World Poll, 2006; authors' regressions.

a. Figure plots the distribution of regression coefficients for 113 developed and developing countries from country-specific ordered probit regressions of satisfaction on log household income, controlling for gender, a quartic in age, and their interaction.

ent is 0.38, with the majority of the estimates between 0.25 and 0.45 and 90 percent between 0.07 and 0.72. In turn, much of the heterogeneity likely reflects simple sampling variation: the average country-specific standard error is 0.07, and 90 percent of the country-specific regressions have standard errors between 0.04 and 0.11.

As an alternative representation of these data, figure 11 directly compares within-country and between-country estimates of the well-being–income gradient. Each solid circle plots the GDP per capita and average well-being of a single country (hence the circles suggest the between-country well-being–GDP gradient), and the slopes of the arrows, fitted to each circle, represent the slope of the well-being–income gradient estimated within that country. Not only are the slopes of the arrows remarkably similar across countries; they are also typically quite close to the between-country well-being–GDP slope (the thick dashed line). Figure 12

Figure 11. Within-Country and Between-Country Estimates of the Life Satisfaction–Income Gradient: Gallup World Poll^a



Source: Gallup World Poll, 2006; authors' regressions. Sources for GDP per capita are described in the text.

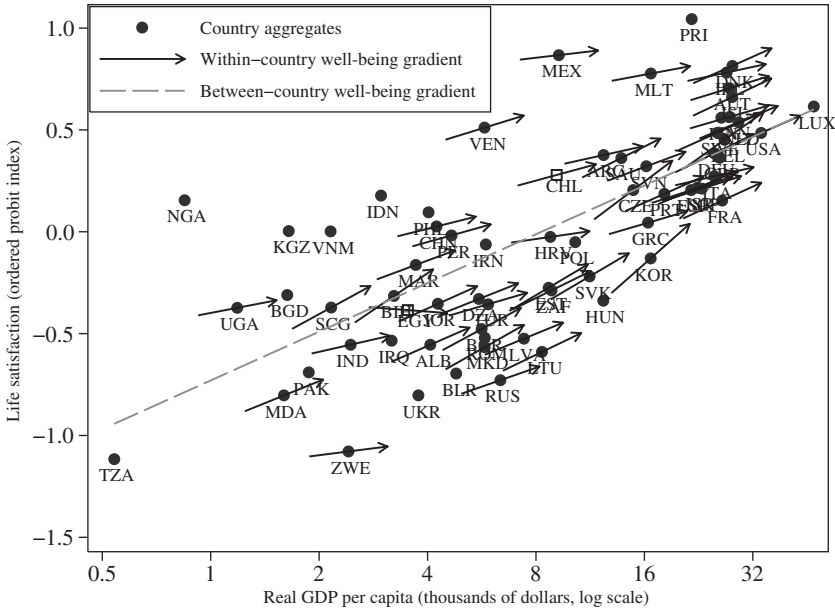
a. Each solid circle plots life satisfaction against GDP per capita for one of 131 developed and developing countries. The slope of the arrow represents the satisfaction-income gradient estimated for that country from a country-specific ordered probit of satisfaction on the log of annual real household income, controlling for gender, a quartic in age, and their interaction. Usable household income data were unavailable for eighteen countries. The dashed line represents the between-country satisfaction-income gradient estimated from an OLS regression of the satisfaction index on the logarithm of real GDP per capita. GDP per capita is at purchasing power parity in constant 2000 international dollars.

repeats this exercise using data from the 1999–2004 wave of the World Values Survey. The household income data in that survey are not as uniform as those in the Gallup World Poll, requiring us to omit several countries.⁵⁵ However, for the countries with sufficient data, the pattern that emerges is similar to that seen in the Gallup data. Repeating the same exercise for the Pew data also yields similar findings (not shown).

Table 2 pools the various national surveys so as to arrive at a summary estimate of the within-country well-being-income gradient. Thus, for each international dataset, we perform an ordered probit of subjective well-

55. In many cases, particularly in earlier waves of the World Values Survey, household income is reported only as an ordinal variable with no information regarding the underlying cardinal measure.

Figure 12. Within-Country and Between-Country Estimates of the Life Satisfaction–Income Gradient: 1999–2004 World Values Survey^a



Source: World Values Survey, 1999–2004 wave; authors’ regressions. Sources for GDP per capita are described in the text.

a. Each solid circle plots life satisfaction against GDP per capita for one of sixty-nine developed and developing countries; hollow squares denote samples that are not nationally representative. The slope of the arrow represents the satisfaction-income gradient estimated for that country from a country-specific ordered probit regression of satisfaction on the log of household income, controlling for gender, a quartic in age, and their interaction, as well as indicator variables for missing age or gender. Usable household income data were unavailable for eighteen countries. The dashed line represents the between-country satisfaction-income gradient estimated from an OLS regression of the satisfaction index on the logarithm of real GDP per capita. GDP per capita is at purchasing power parity in constant 2000 international dollars.

being on log household income, controlling for country (or, for the World Values Survey, country-by-wave) fixed effects, which serve to control for not only the between-country variation in GDP per capita, but also variation in measured income due to differences in exchange rates, purchasing power, or other country-specific factors. The first column shows the results from a simple ordered probit of well-being on log household income, controlling for these fixed effects; the second column adds controls for gender, a quartic in age, and the interaction of these variables. Comparison of these results with the corresponding between-country estimates in table 1 shows them to be roughly similar in magnitude, although as seen in the figures, in most cases the between-country estimates are larger than the within-country estimates, which are centered around 0.3.

Table 2. Within-Country Ordered Probit Regressions of Subjective Well-Being on Income^a

<i>Survey</i>	<i>Without controls</i>	<i>With controls^b</i>	<i>Instrumental variables regression^c</i>	<i>Sample size^d</i>
Gallup World Poll, 2006: Ladder question	0.321*** (0.005)	0.318*** (0.005)	0.592*** (0.014)	102,583 (113 countries)
World Values Survey: Life satisfaction				
1981–84 wave	0.167*** (0.019)	0.199*** (0.022)	n.a.	12,198 (10 countries)
1989–93 wave	0.130*** (0.011)	0.153*** (0.011)	0.001 (0.041)	32,371 (26 countries)
1994–99 wave	0.225*** (0.012)	0.243*** (0.013)	0.233*** (0.021)	11,924 (9 countries)
1999–2004 wave	0.277*** (0.007)	0.286*** (0.007)	0.305*** (0.018)	60,988 (52 countries)
Combined, with country × wave fixed effects	0.232*** (0.007)	0.249*** (0.007)	0.258*** (0.013)	117,481 (62 countries)
World Values Survey: Happiness				
1981–84 wave	0.324*** (0.021)	0.281*** (0.023)	n.a.	12,021 (10 countries)
1989–93 wave	0.198*** (0.012)	0.188*** (0.013)	0.064 (0.047)	31,475 (26 countries)
1994–99 wave	0.208*** (0.013)	0.209*** (0.013)	0.269*** (0.022)	13,176 (10 countries)
1999–2004 wave	0.259*** (0.008)	0.248*** (0.008)	0.292*** (0.020)	60,627 (52 countries)
Combined, with country × wave fixed effects	0.244*** (0.008)	0.234*** (0.008)	0.266*** (0.015)	117,299 (62 countries)
Pew Global Attitudes Survey, 2002: Ladder question	0.320*** (0.008)	0.324*** (0.008)	0.451*** (0.016)	32,463 (43 countries)

Source: Authors' regressions.

a. Table reports results of ordered probit regressions of the indicated measure of well-being on log household income, controlling for country fixed effects or country × wave fixed effects where noted. See the notes to table 1 for wording of survey questions. Observations are weighted to give equal weight to each country × wave. Numbers in parentheses are robust standard errors, clustered by country. Asterisks indicate statistical significance at the *10 percent, **5 percent, and ***1 percent level.

b. Controls include sex, a quartic in age, and their interaction, and indicators for missing age or sex.

c. The first stage instruments for log household income using indicator variables for levels of education, entered separately for each country, controlling for a quartic in age, interacted with gender, and country fixed effects. The second stage is an ordered probit regression of well-being on the predicted values, the residuals, the same controls, and country fixed effects.

d. Samples are restricted to observations with valid household income data, from nationally representative samples (see appendix B). Instrumental variables regressions are further restricted to those countries with valid education data, which further restricts the World Values Survey samples, as valid education data were available for zero, then three, ten, and fifty-two countries in successive waves.

An important issue in considering the within-country cross-sectional relationship between income and subjective well-being is the extent to which measured income differences at a point in time reflect differences in permanent income versus transitory shocks. If people are able to smooth their consumption, then subjective well-being should change little with transitory income changes, and permanent shocks should have a much larger impact. The variation in GDP per capita between countries is likely dominated by variation in permanent income, whereas the variation in annual income within a population likely reflects both permanent and transitory shocks.

A simple back-of-the-envelope calculation can help determine an upper bound on the extent to which these issues are distorting the comparisons in tables 1 and 2. If all cross-country variation in GDP per capita is assumed permanent, and if people are perfect permanent-income consumers, then the coefficients in table 1 can be interpreted as the response of well-being to a shock to consumption. Standard estimates for the United States suggest that around half the variation in annual income in a national cross section is transitory, and a \$1 shock to transitory income typically translates into around a \$0.05 shock to permanent income. Thus, a \$1 change in measured income translates to a roughly \$0.525 change in permanent income. In this case the estimates in table 2 need to be adjusted upward by around 90 percent ($1/0.525$) to be interpreted as the relationship between well-being and permanent income or consumption. If, instead of assuming perfect smoothing, we accept Campbell and Mankiw's estimate that 50 percent of income is earned by "rule-of-thumb" consumers whose propensity to consume from current income is equal to their propensity to consume from permanent income,⁵⁶ the relevant adjustment is closer to 30 percent. This adjustment would make the within- and between-country estimates roughly similar.

We can also address this issue empirically. In an effort to isolate the response of well-being to permanent income, the estimation reported in the third column of table 2 instruments for income using educational attainment, entered separately for each country.⁵⁷ Although we are confident that these instruments isolate variation in permanent rather than tran-

56. Campbell and Mankiw (1990).

57. We follow Rivers and Vuong (1988) in their approach to estimating an instrumental variables ordered probit. Thus, the first stage involves a regression of log household income on indicator variables for each level of education in each country, controlling for country fixed effects as well as gender and a quartic in age, entered separately for each gender, and indicators for missing age or gender. The second stage involves an ordered probit regression of well-being on the predicted values and residuals from the second stage, as well as the same controls, including country fixed effects.

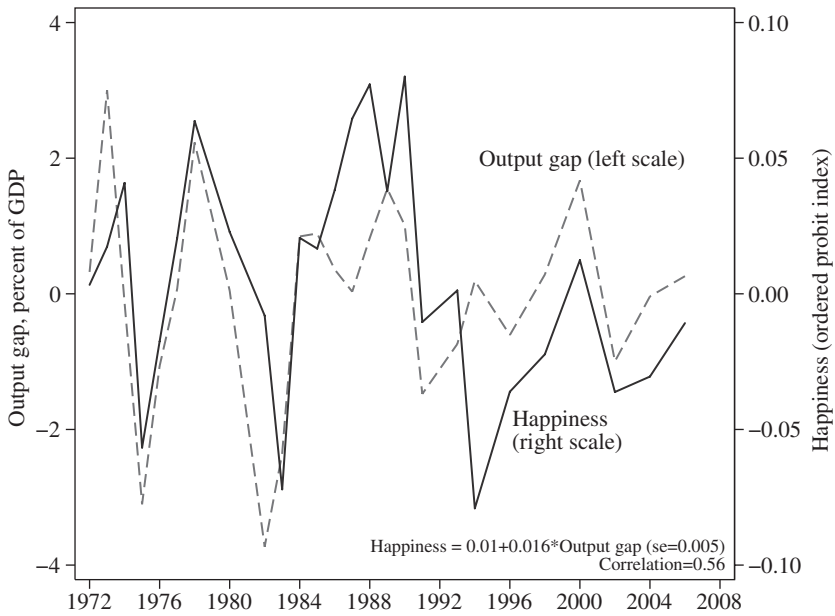
sitory income, we do not hold much faith that the exclusion restriction holds—that education does not have an effect on well-being beyond that mediated by income.⁵⁸ Given that these omitted effects are likely positive, our instrumental variables estimates may overstate the within-country income-well-being gradient. Indeed, in most cases the instrumental variables estimates are larger than the ordered probit estimates of well-being on income. In the largest dataset, the Gallup World Poll, the estimated gradient is 0.6.

The discussion above has been premised on the straightforward view that transitory income shocks yield smaller impacts on well-being than do permanent shocks. Yet the most direct evidence we have on this point—the movement of well-being over the business cycle—in fact suggests the opposite. Figure 13 shows that business-cycle variation in the output gap produces quite large effects on subjective well-being. Indeed, the estimated well-being–transitory income gradient suggested by these shocks is about five times larger than the well-being–GDP gradient estimated in table 1. If this sort of variation is representative of the response of happiness to transitory income, then, paradoxically enough, our findings in table 2 may substantially overstate the within-country well-being–permanent income link.

Although our analysis provides a useful measurement of the bivariate relationship between income and well-being both within and between countries, there are good reasons to doubt that this corresponds to the *causal* effect of income on well-being. It seems plausible (perhaps even likely) that the within-country well-being–income gradient may be biased upward by reverse causation, as happiness may well be a productive trait in some occupations, raising income. A different perspective, offered by Kahneman and coauthors,⁵⁹ suggests that within-country comparisons overstate the true relationship between subjective well-being and income because of a “focusing illusion”: the very nature of asking about life satisfaction leads people to assess their life relative to others, and they thus focus on where they fall relative to others in regard to concrete measures such as income. Although these specific biases may have a more important impact on within-country comparisons, it seems likely that the bivariate well-being–GDP relationship may also reflect the influence of third factors, such as democracy, the quality of national laws or government, health, or even

58. For instance, Lleras-Muney (2005) shows positive impacts of compulsory schooling on health.

59. Kahneman and others (2006).

Figure 13. Happiness and the Output Gap in the United States^a

Sources: General Social Survey, 1972–2006; Bureau of Economic Analysis; authors' calculations.

a. "Output gap" is the difference between real GDP per capita and its trend, estimated using a Hodrick-Prescott filter on annual data on the logarithm of real GDP per capita, with the smoothing parameter set to 6.25. Happiness data are aggregated into a happiness index by running an ordered probit regression of satisfaction on year fixed effects. See figure 8 for wording of the question. See text for details of the sample.

favorable weather conditions, and many of these factors raise both GDP per capita and well-being.⁶⁰ Other factors, such as increased savings, reduced leisure, or even increasingly materialist values, may raise GDP per capita at the expense of subjective well-being. At this stage we cannot address these shortcomings in any detail, although given our reassessment of the stylized facts, we would suggest an urgent need for research identifying these causal parameters.

Economic Growth and Happiness

The last two sections have shown that wealthier societies have greater subjective well-being than poorer societies and that, to a similar degree, wealthier members of a society are happier than their poorer counterparts.

60. Kenny (1999) argues directly for reverse causation running from happiness to income.

This then leads to our final question: do societies get happier through time as they become richer? Easterlin argues that the possibly confounding “cultural influences on international happiness comparisons underscore the importance of national time series evidence . . . for inferring the relationship between subjective well-being and economic development.”⁶¹ Indeed, the core of the Easterlin paradox lies in Easterlin’s failure to isolate statistically significant relationships between average levels of happiness and economic growth through time. Easterlin’s 1974 and 1995 papers contain three important datasets, tracking the time series of happiness within Europe, Japan, and the United States.

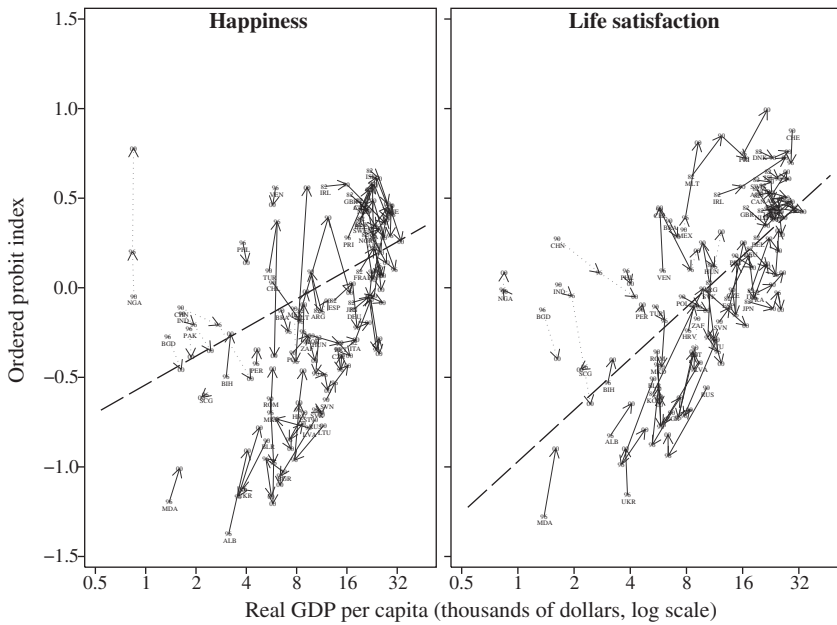
Our analysis is based on three observations about the inferences that existing datasets can support. First, absence of evidence should not be confused with evidence of absence. This is particularly important given both the variability of happiness aggregates between surveys and the limited range of variation in time-series rather than cross-national comparisons of GDP per capita. Second, when we reanalyze these data, we find that happiness has in fact risen in Japan and Europe. The failure of happiness to rise in the United States remains a puzzling outlier, although the extent to which it constitutes a sharp exception should not be overstated. Third, as more data have become available, in the form of both extended national time series and observations from new countries, evidence that happiness rises with GDP per capita has started to accumulate.

Indeed, the World Values Survey has been running since 1981, and across its four waves we now have repeated observations on a large number of countries, spread across several decades. Figure 14 shows the movement of both life satisfaction and real GDP per capita across the waves for all countries for which this survey offers repeated observations. As before, we estimate average well-being in a country-wave as the coefficient from an ordered probit regression of well-being on a saturated set of country-by-wave fixed effects. Arrows link each individual country’s change in well-being-GDP space over time, and so the slope of each arrow corresponds to the well-being-income gradient derived from two consecutive observations in a country’s national time series (dotted arrows connect points where the sampling frame changed, and hence valid time-series comparisons cannot be made).

Several points are evident from figure 14. First, there appears to be a general tendency for economic growth to be accompanied by growth in subjective well-being (arrows tend to point northeast), and economic

61. Easterlin (1995, pp. 43–44).

Figure 14. Subjective Well-Being-Income Gradients across Time: World Values Survey^a



Source: World Values Survey, waves 1-4; authors' regressions. Sources for GDP per capita are described in the text.

a. Arrows show the evolution of measured well-being and real GDP for each country. Dotted arrows join observations based on noncomparable sampling frames (see appendix B). Dashed lines are fitted from an OLS regression of the well-being measure on the natural log of real GDP, estimated from pooling all four waves. See notes to figure 2 for question details and construction of aggregate well-being indices. Real GDP is at purchasing power parity in constant 2000 international dollars.

decline, which is most visible in the former Eastern bloc, has been accompanied by a decline in well-being (arrows pointing southwest). Of the eighty-nine changes shown in the left-hand panel of figure 14, happiness and GDP per capita change in the same direction in sixty-two (fifty-three show growth in both; nine show declines), whereas they move in opposite directions in twenty-seven (of which twenty reflect economic growth unaccompanied by growth in happiness, and seven reflect growing happiness despite economic decline). The life satisfaction data in the right-hand panel yield much weaker results, with satisfaction and GDP per capita moving in the same direction in only forty-six of ninety cases, reflecting generally weaker measured life satisfaction in the two most recent waves of the survey (more on this below).

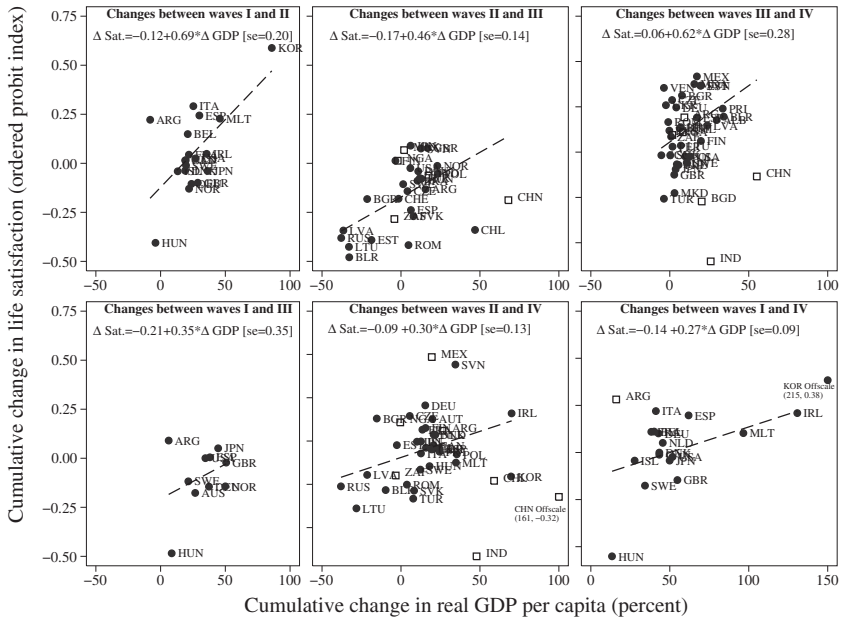
Second, when we average across these country-specific estimates, the well-being-income link within countries through time appears to be roughly

similar to that estimated from the pooled cross-country, cross-time variation (shown as the dashed line in each panel). Third, substantial heterogeneity remains in these estimated responses, although this may reflect the influence of other factors on measured well-being.

Finally, these time-series changes are strongly influenced by the result of common patterns across countries observed in specific waves. We suspect that the trend in life satisfaction has been distorted by changes in question ordering. In particular, in the 1994–99 and 1999–2004 waves, the life satisfaction question was preceded by a question asking, “How satisfied are you with the financial situation of your household?” Respondents typically rate their financial satisfaction substantially lower than their life satisfaction (on the same 1-to-10 scale, responses average about one point lower), and hence this question may have influenced how respondents subsequently reported their life satisfaction. To check this, we assess the (raw) correlation between life satisfaction and financial satisfaction for the eight countries with representative samples in each round of the World Values Survey; this correlation was 0.53 and 0.57 in the two most recent waves, significantly above previous levels (0.45 in the first wave and 0.43 in the second). The happiness question was never proximate to the financial satisfaction question, and the correlation of happiness with financial satisfaction was quite stable across each of the waves (it was recorded as 0.29, 0.30, 0.32, and 0.29 from the earliest to the latest wave). Similarly, in the 1994–99 and 1999–2004 waves, the happiness question was part of a battery of questions probing the importance of friends, family, leisure, politics, and religion, and a similar analysis reveals that the correlation of measured happiness with these variables rose. If these questions prime positive thoughts, this question order change may have inflated measured happiness in the past two waves.

These question-order changes make direct comparisons of countries’ well-being levels across successive waves problematic. However, to the extent that these influences are common across countries, first differencing will yield useful estimates. Thus, in figure 15 we analyze changes in life satisfaction and log GDP between each wave of the panel. The first row shows differences between adjacent waves, and the second row shows longer differences, with the last panel showing differences between the first and last wave. (Because of the uneven participation through time of many countries, these longer-difference panels contain information not shown in the first row). In each comparison of pairs of waves, we find that larger rises in GDP per capita are associated with larger rises in life satisfaction, and the magnitude of these gradients tends to be centered around

Figure 15. Change in Life Satisfaction and Economic Growth: World Values Survey^a



Source: World Values Surveys, waves 1 (1981-84), 2 (1989-93), 3 (1994-99), and 4 (1999-2004); authors' regressions. Sources for GDP per capita are described in the text.

a. Solid circles show changes in life satisfaction and real GDP per capita between various waves of the World Values Survey; hollow squares reflect changes based on noncomparable sampling frames (see appendix B). Dashed lines show the fit from the reported OLS regression of changes in the life satisfaction index on the percent change in GDP, based only on comparable changes in life satisfaction. Graphs in the first row show nineteen, ten, and seventeen comparable short first differences, and those in the second row twenty-five, thirty-two, and thirty-three long first differences. GDP per capita is at purchasing power parity in constant 2000 international dollars.

0.4. A parallel analysis of the happiness data (not shown) yielded roughly similar results (the slope was positive in five of six panels and statistically significant in only one case).

Panel regressions provide an alternative and statistically more efficient way to combine this information, and so in table 3 we turn to analyzing both life satisfaction and happiness measures in the World Values Survey as a country-wave panel dataset. The first column reports the results of respondent-level ordered probit regressions of well-being on log GDP per capita, and the second column aggregates the data to the country-wave level; these are OLS regressions of our well-being index on log GDP per capita. The first row reports results for the simple bivariate well-being-GDP relationship and hence pools both within-country and between-

Table 3. Panel Regressions of Subjective Well-Being on GDP per Capita: World Values Survey^a

<i>Dependent variable and specification</i>	<i>Micro data estimates^b</i>	<i>Macro data estimates^c</i>	<i>Sample</i>
<i>Life satisfaction^d</i>			
Levels	0.386*** (0.039)	0.414*** (0.041)	234,093 (166 country-waves)
Levels with country fixed effects	0.307*** (0.065)	0.301*** (0.091)	234,093 (166 country-waves)
Levels with country and wave fixed effects	0.579*** (0.088)	0.552*** (0.118)	234,093 (166 country-waves)
Short first differences	n.a.	0.596*** (0.082)	87 differences
Long first differences ^e	0.575*** (0.116)	0.314*** (0.072)	133,900 (98 country-years = 49 differences)
<i>Happiness^f</i>			
Levels	0.213*** (0.056)	0.230*** (0.064)	228,159 (165 country-waves)
Levels with country fixed effects	0.388*** (0.093)	0.363*** (0.131)	228,159 (165 country-waves)
Levels with country and wave fixed effects	0.263** (0.111)	0.216 (0.187)	228,159 (165 country-waves)
Short first differences	n.a.	0.215 (0.136)	86 differences
Long first differences ^e	0.305** (0.147)	0.114 (0.103)	132,662 (98 country-waves = 49 differences)

Sources: Authors' regressions using data from World Values Surveys, 1981–2004.

a. Results of regressions of the indicated measure of well-being on log real GDP per capita. Sample pools observations from all nationally representative samples in the four waves of the World Values Survey. Numbers in parentheses are robust standard errors, clustered by country. Asterisks indicate statistical significance at the *10 percent, **5 percent, and ***1 percent level.

b. Ordered probit regression of subjective well-being, using data by respondent, on log real GDP per capita for the respondent's country, weighting observations to give equal weight to each country × wave. Standard errors are clustered by country-wave.

c. National well-being index, using data by country-wave, is regressed on log real GDP per capita. The index is calculated in a previous ordered probit regression of well-being on country × wave fixed effects.

d. See table 1, note g, for wording of survey question.

e. Difference between first and last observation for each country.

f. See table 1, note h, for wording of survey question.

country variation. The estimated coefficients are 0.4 for life satisfaction and 0.2 for happiness. To isolate the within-country time-series variation, the second row adds controls for country fixed effects. Consistent with figure 14, the well-being-GDP gradient estimated from this time-series variation is similar to that estimated from the point-in-time between-country comparisons.

The next row adds further controls for each wave of the World Values Survey, which partial out the changes in well-being that reflect differences in surveys across waves. As might be expected in light of the previous discussion of question order effects, the inclusion of these controls increases the estimate of the time-series life satisfaction–GDP gradient to nearly 0.6 and lowers the estimate of the time-series happiness–GDP gradient to a bit more than 0.2.

In subsequent rows we take short first differences of consecutive country-wave observations, as well as long first differences (subtracting the first from the last observation for each country). Consistent with the analysis in figure 15, there is a clear, and statistically significant, relationship between changes in life satisfaction and log GDP per capita over time in these countries. The estimates for happiness are similar, albeit smaller and less precisely estimated. These repeated international cross sections yield estimates of the time-series well-being–income gradient centered roughly around 0.4, with larger estimates for life satisfaction than for happiness.

Equally, it is worth emphasizing that these estimates are both somewhat imprecisely estimated and fragile. Although the large cross-country datasets allow for useful comparisons between populations in abject poverty and those in industrialized powerhouses, the within-country time-series variation is simply less impressive. Indeed, it is worth noting that the standard deviation of log GDP per capita across countries (in the 1999–2004 wave) is 1.0, whereas the standard deviation of between-wave first differences in log GDP per capita (across all waves) is only 0.2, and hence strong inferences are difficult to draw. Moreover, the inferences one draws from these data are particularly sensitive to the quite unusual economic trajectories of a small number of countries, such as the rapid growth in Korea and Ireland and the decline of the former Eastern bloc countries (figure 15). Even so, most of our approaches to these data yield suggestive evidence falsifying the Easterlin hypothesis that the time-series well-being–GDP gradient is zero. Moreover, even in those cases in which the data fail to falsify the null that the gradient is zero, they also fail to falsify the alternative null hypothesis that this gradient is equal to 0.4, similar to that obtained from our between-country or within-country analyses.

How should our findings be reconciled with earlier reports suggesting no link between changes in GDP per capita over time and life satisfaction? We suspect that the key is simply that our analysis of the satisfaction–income gradient based on both within- and between-country comparisons gives us a specific quantitative yardstick for assessing the importance of (even imprecisely estimated) trends in subjective well-being.

Europe

We turn next to the other major set of repeated international cross-sectional data, the Eurobarometer Survey, a data collection intended to track public opinion across the European Union. We draw our data from the Mannheim Eurobarometer Trendfile, which collects available micro-data from 1970 to 2002, supplemented with data extracted from print editions of the Eurobarometer Reports series from 2002 through 2007. These surveys initially asked respondents in what were then the nine member states of the European Community about their life satisfaction. A life satisfaction question has been asked at least annually (and often semiannually) from 1973 onward (except in 1974 and 1996). The survey has expanded as the European Union itself has expanded, covering fifteen countries by 2002 (with separate surveys for East and West Germany), and it presently includes thirty countries (including three candidate countries), yielding a broad but unbalanced panel. A happiness question was also briefly asked (from 1975 through 1986, except in 1980 and 1981, and in a different format in 2006); given these gaps in the data, we focus on life satisfaction. For the purposes of our analysis, we keep West Germany separate from East Germany, which permits us to analyze a continuous sample of well-being among West Germans over thirty-five years.

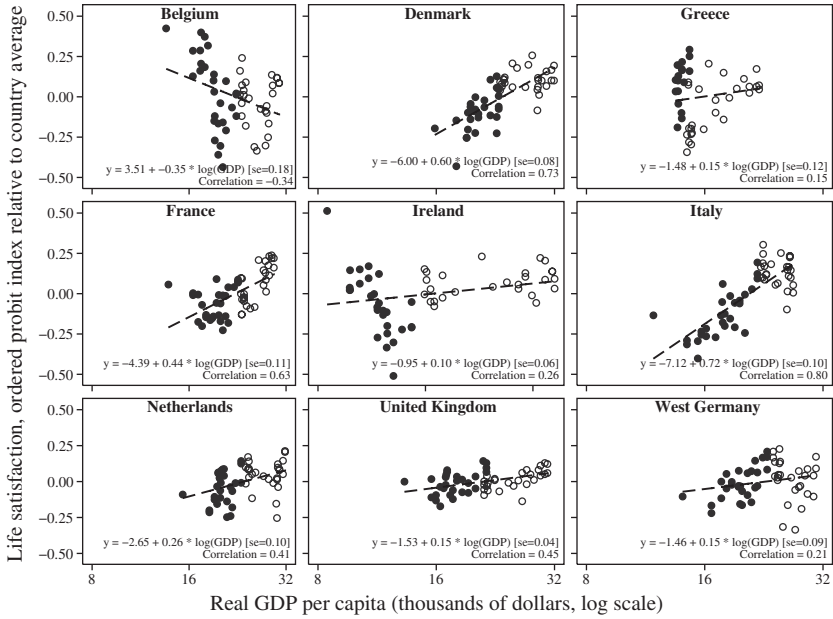
We begin by analyzing the relationship between life satisfaction and GDP for the nine countries that constituted the original 1973 sample. Easterlin analyzed these same nine countries (through to 1989), concluding that “Satisfaction drifts upward in some countries, downward in others. The overall pattern, however, is clearly one of little or no trend in a period when real GDP per capita rises in all of these countries from 25 to 50 percent.”⁶² In a subsequent update, Easterlin maintains that “I think the evidence continues to support my generalization in the 1995 study.”⁶³

Figure 16 updates this analysis, adding a further eighteen years of data (shown with hollow circles). In eight of the nine countries, rising GDP per capita has been associated with rising life satisfaction; the findings are statistically significant in six cases ($p < 0.10$, assessed using Newey-West standard errors accounting for first-order autocorrelation). This figure also suggests a couple of puzzles: a significant declining trend in satisfaction is observed in Belgium, and declining life satisfaction in Ireland during the 1970s and 1980s, although this was quickly followed by rising satisfaction

62. Easterlin (1995, p. 38).

63. Easterlin (2005a, p. 434).

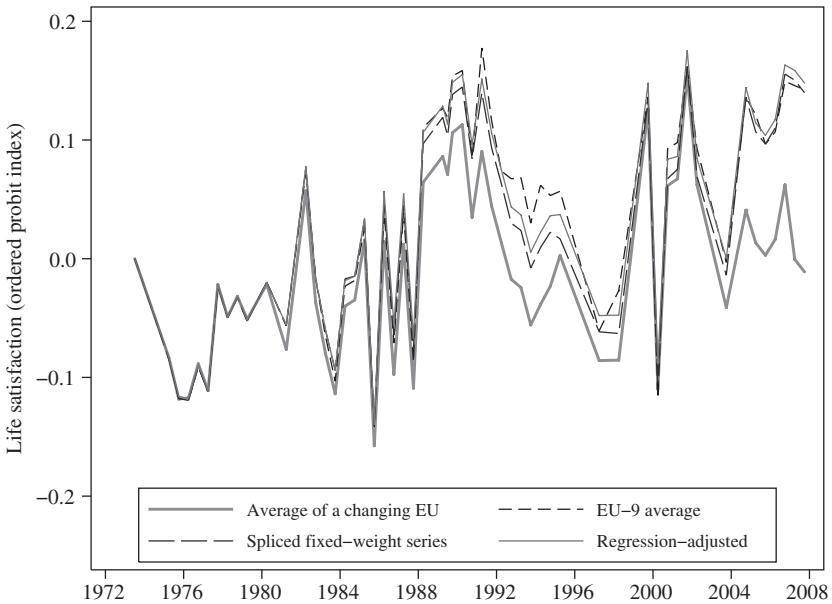
Figure 16. Change in Life Satisfaction and Economic Growth in Europe: Eurobarometer Survey^a



Sources: Eurobarometer Trendfile (for 1973–2002); biannual Eurobarometer reports (for 2002–07); authors’ calculations. Sources for GDP per capita are described in the text.

a. Solid circles represent separate observations from each round of the Eurobarometer survey from 1973 to 1989; these were the data analyzed in Easterlin (1995); open circles extend the sample from 1990 to 2007. Each panel shows data for one of the nine countries analyzed by Easterlin. Data are aggregated into a satisfaction index by running an ordered probit regression of satisfaction on country \times wave fixed effects and subtracting country averages. Dashed lines are fitted from the reported OLS regression; Newey-West standard errors (se) are reported, accounting for first-order autocorrelation. The life satisfaction question asks, “On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?” GDP per capita is at purchasing power parity in constant 2000 international dollars.

during the rapid economic growth associated with the “Irish miracle.” (Satisfaction appeared to be anomalously high in the very first Irish survey; dropping this observation yields, for the entire sample period, a statistically significant coefficient on log GDP per capita of 0.14, with a standard error of 0.05.) Our point is not to count up the number of statistically significant responses one way or the other, but rather to suggest that across the nine large European countries for which we have long time series, life satisfaction has typically risen with GDP per capita. Moreover, estimates of the satisfaction-GDP gradient based on these national time

Figure 17. Trends in Life Satisfaction in the European Union^a

Sources: Eurobarometer, 1973–2007; authors' regressions.

a. Lines depict alternative aggregations of semiannual time series of life satisfaction for each country, derived by running an ordered probit of satisfaction on country \times wave fixed effects. “Average of a changing EU” is calculated by taking a population-weighted average of the satisfaction indices for the set of countries that were members of the European Community or the European Union at the indicated point in time; hence the average is affected by changes in the group’s composition. “EU-9 average” is calculated by taking a fixed-weight average of the satisfaction indices of the nine members of the European Community at the beginning of the sample: Belgium, Denmark, France, West Germany, Ireland, Italy, Luxembourg, the Netherlands, and the United Kingdom; weights reflect the average population share of each country in the group. “Spliced fixed-weight series” simply sums through time first differences in the broadest available fixed-weight average of satisfaction in the member nations; consistent fixed-weight indices were calculated separately for each constellation through the sample: EU-9 (summer 1973–fall 2007), EU-10 (adding Greece, from 1981 onward), EU-12 (adding Portugal and Spain, from 1986 onward), EU-12+ (adding East German surveys, from the fall 1990 survey onward), EU-16 (adding Austria, Finland, and Sweden from 1995 onward), EU-26 (adding Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia from the fall 2004 survey onward), and EU-27 (adding Bulgaria and Romania, from 2007 onward). Population weights for each index reflect the average population share of each country in that aggregate. “Regression-adjusted” index is the series of time fixed effects estimated by running a population-weighted ordered probit regression of satisfaction on survey time fixed effects, controlling for country fixed effects.

series, although quite variable, average around 0.25, with some estimates larger and some smaller.

The upward trend in life satisfaction across the European Union is not widely understood, and in figure 17 we provide some intuition for why this has not been obvious. The simplest approach to building an EU-wide time series of life satisfaction involves taking a population-weighted average of the satisfaction levels of whichever countries happen to be member states

at any point in time. Through this period the European Union has systematically expanded to incorporate poorer countries, which have lower average life satisfaction. This expansion has thus pushed measured average life satisfaction downward, despite the fact that satisfaction rose within most countries. This can most easily be seen by simply examining the nine countries (the EU-9) that have been in the European Union, and hence the Eurobarometer, since 1973. This analysis takes a population-weighted average of the satisfaction indices of the EU-9 and shows rising life satisfaction through time.

In order to use the data from all countries without having the resulting time series driven by compositional changes, we also construct a regression-adjusted series by running an OLS regression of national satisfaction indices on time (survey round) fixed effects, weighting by population and controlling for country fixed effects (thereby adjusting for different average well-being levels among new EU entrants). These time fixed effects are also plotted in figure 17 and clearly suggest a rising trend in life satisfaction similar to that seen in the EU-9 average. Finally, we create a spliced series by summing through time first differences in the broadest available fixed-weight average of satisfaction in EU member nations.⁶⁴ This series is quite similar to the regression-adjusted measure. Clearly the simple average disguises much of the rise in satisfaction occurring within EU members.

Even accounting for these compositional changes, it would be difficult to infer that a positive trend either did or did not exist on the basis of only Easterlin's 1973–89 sample. But over the entire 1973–2007 period, the magnitude of the trend rise in satisfaction turns out to be quite close to what might be expected given underlying GDP trends. Fitting a simple time trend to the composition-adjusted aggregates shown in figure 17 suggests that life satisfaction in Europe rose at a (statistically significant) average rate of about 0.006 per year, compared with a trend rise in log GDP per capita of around 0.020 per year. Considered jointly, these trends point to a long-run satisfaction-GDP gradient of about 0.3 ($= 0.006/0.020$), which both falsifies the null hypothesis of no positive relationship and is roughly consistent with the magnitudes seen in our within- and between-country assessments.

64. The spliced series begins with the 1973 survey and analyzes the EU-9 until 1981, at which point Greece is added. In 1986 Portugal and Spain are added; in 1990 the series is adjusted for German reunification. In 1995 Austria, Finland, and Sweden are added; in 2004 Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia are added; finally, the full EU-27 are included as of 2007 with the addition of Bulgaria and Romania. Population weights for each index reflect the average population share of each country in that aggregate.

To further examine these patterns, table 4 formalizes our findings with a series of panel regressions exploiting all of the Eurobarometer Survey observations across all countries; the top panel analyzes life satisfaction and the bottom panel happiness. (Given the very limited happiness data available, estimates in the bottom panel are extremely imprecise.) As with our panel analysis of the World Values Survey, we begin by including no fixed effects and subsequently add in country fixed effects, and then both country and year fixed effects. The latter two estimates focus on the time-series relationship between satisfaction and GDP and yield coefficients of about 0.2.

The last two rows of each panel attempt to minimize the potential influence of high-frequency variation, by averaging well-being and GDP over five-year periods or over entire decades. These first-difference regressions yield somewhat larger estimates for life satisfaction, although the decadal differences are imprecisely estimated. Because happiness was included only in the early years of the survey, there are many fewer observations, yielding extremely imprecise estimates. Nonetheless, in all cases the estimated coefficients are positive, in some specifications we can falsify the null hypothesis that the well-being-GDP gradient is zero, and in no case can we falsify that it is 0.4.

Japan

Arguably the most persuasive evidence in favor of the Easterlin paradox has come from Japan, which provides a striking case study both because of its dramatic growth in the postwar period (real GDP has risen by a factor of six since World War II), and because it was believed that consistent data on subjective well-being had been continuously collected by the government since 1958 in the “Life in Nation” surveys. Previous researchers have analyzed the simple summary of these questions provided by Ruut Veenhoven,⁶⁵ observing that average levels of well-being had remained flat even in the face of this spectacular growth.⁶⁶

Upon closer inspection, however, these Japanese data are neither as persuasive as many thought, nor is the trend flat. We returned to the original

65. Veenhoven (1993).

66. For instance, Easterlin (1995, pp. 39–40) notes that “Between 1958 and 1987 real per capita income in Japan multiplied a staggering five-fold, propelling Japan to a living level equal to about two-thirds that of the United States. . . . Despite this unprecedented three decade advance in level of living, there was no improvement in mean subjective well-being.” These observations have been cited approvingly by Layard (2005a), Frank (2005), and Kahneman and coauthors (2006), among dozens of others.

Table 4. Panel Regressions of Subjective Well-Being on GDP per Capita: Eurobarometer Survey^a

<i>Dependent variable and specification</i>	<i>Micro data estimates^b</i>	<i>Macro data estimates^c</i>
Life satisfaction, 1973–2007 ^d		
Levels	0.737*** (0.181)	0.769*** (0.177)
Levels and country fixed effects	0.192*** (0.066)	0.194*** (0.059)
Levels and country and wave fixed effects	0.208** (0.099)	0.193** (0.094)
First differences, five-year averages ^f	n.a.	0.579*** (0.181)
First differences, decadal averages ^g	n.a.	0.333 (0.231)
Happiness, 1975–86 ^e		
Levels	0.422 (0.517)	0.448 (0.489)
Levels and country fixed effects	0.554 (0.351)	0.626* (0.346)
Levels and country and wave fixed effects	1.037 (0.993)	1.262 (0.904)
First differences, five-year averages ^f	n.a.	0.107 (0.840)
First differences, decadal averages ^g	n.a.	2.108 (1.678)

Sources: Authors' regressions using 1973–2002 data are drawn from Eurobarometer Trendfile, and those using 2002–07 are from biannual Eurobarometer reports.

a. Results of regressions of the indicated measure of well-being on log real GDP per capita. Sample pools observations from all Eurobarometer samples, using sample weights to typically yield around 1,000 nationally representative respondents in each country and wave (keeping East and West Germany separate). Numbers in parentheses are robust standard errors, clustered by country. Asterisks indicate statistical significance at the *10 percent, **5 percent, and ***1 percent level.

b. Ordered probit regression of subjective well-being, using data by respondent, on log real GDP per capita for the respondent's country, weighting observations to give equal weight to each country × wave. Standard errors are clustered by country-wave.

c. National well-being index, using data by country-wave, is regressed on log real GDP per capita. The index is calculated in a previous ordered probit regression of well-being on country × wave fixed effects.

d. Respondents were asked, "On the whole, are you (4) very satisfied, (3) fairly satisfied, (2) not very satisfied, or (1) not at all satisfied with the life you lead?" Sample yields 850,153 respondents from 776 country × wave observations in 31 countries; 77 five-year first differences; and 37 decadal differences.

e. Respondents were asked, "Taking all things together, how would you say things are these days—would you say that you're (3) very happy, (2) fairly happy, or (1) not too happy these days?" Sample yields 134,590 respondents from 139 country × wave observations in 12 countries; 19 five-year first differences, and 9 decadal first differences (all 1980s compared with 1970s).

f. Data were averaged by country in five-year periods (1973–77, 1978–82, 1983–87, 1988–92, 1993–97, 1998–2002, and 2003–07), and first differences of well-being were regressed against first differences in the log of average real GDP per capita.

g. Data from the 1970s, 1980s, 1990s, and 2000s were averaged separately, and first differences of well-being were regressed against first differences in the log of average real GDP per capita.

codebooks and had the questions translated.⁶⁷ This exercise was quite revealing, suggesting several important series breaks. Accounting for these breaks yields a very different perspective. We provide a full accounting in table 5, which shows both the literal and the idiomatic translations of the survey questions as they have changed.⁶⁸

Three important findings emerge from this table. First, in 1964 the response categories changed dramatically. The top category was changed from the catch-all “Although I am not innumerablely satisfied, I am generally satisfied with life now” to the more demanding “Completely satisfied.” Not surprisingly, the proportion reporting their well-being in this highest category declined from 18.3 percent to 4.4 percent. The second category from the top also became more demanding, changing from “Although I can’t say that I am satisfied, if life continues in this way, it will be okay,” to “Although I can’t say I am completely satisfied, I am satisfied.” In parallel, the bottom category changed from “Life now is very unbearable” to “Completely dissatisfied,” but the proportion choosing this lowest category changed little. Second, the questions asked from 1958 to 1969 focused on feelings about “life at home,” whereas the focus of the relevant question from 1970 onward was on global life satisfaction. Third, the survey question—and the allowable responses—changed again in 1992.

Properly viewed, this leaves us with four periods within which we can make useful assessments of trends in subjective well-being in Japan. A cursory inspection of table 5 suggests an upward trend in well-being in 1958–63, continuing when a new question was asked for the 1964–69 period, followed by a slower rise from 1970 to 1991. This roughly parallels the path of Japanese GDP through these periods. From 1992 until 2007, life satisfaction fell, but this coincides with the end of the Japanese growth miracle and indeed the onset of an economic slump. All told, these findings suggest that subjective well-being in Japan has largely risen with GDP per capita, and that it rose most sharply during the period of rapid growth.

Having established that these data appear qualitatively consistent with a positive satisfaction-GDP gradient, we now turn to a quantitative assessment of the magnitude of this link. One simple approach involves treating these data as four separate datasets and following our earlier style of analysis. Thus, within each continuous subseries we create a time series of average well-being by performing an ordered probit of subjective well-being

67. We thank Michael L. Woodford for his patient assistance with these translations.

68. The original Japanese questions in kanji characters are printed in Stevenson and Wolfers (2008).

Table 5. Subjective Well-Being in Japan: Life in Nation Survey^a

Percent except where stated otherwise

Survey month	Very satisfied	Fairly satisfied	Not very satisfied	Not at all satisfied	Unsure	Don't know or not asked	No. of observations
	<p><i>Literal: "By the way, how do you feel about the way your life is going at home? Which of these is your feeling close to?"</i> <i>Idiomatic: "How do you feel about your circumstances at home? Please choose one of the following."</i></p>						
	<p><i>"Although I am not innumera- bly satisfied, I am generally satisfied with life now"</i> (satisfied)</p>						
	<p><i>"Although I can't say that I am satisfied, if life continues in this way, it will be okay" (not satisfied, not dissatisfied)</i></p>						
	<p><i>"Somewhat dissatisfied with life now"</i> (somewhat dissatisfied)</p>						
	<p><i>"Life now is very unbearable"</i> (extremely dissatisfied)</p>						
	<p><i>"Unclear"</i> (not sure)</p>						
Feb. 1958	16	44	29	9	2		15,941
Jan. 1959	17	49	25	6	3		16,897
Jan. 1960	15	45	28	6	6		17,291
Jan. 1961	14	47	29	5	5		17,103
Jan. 1962	16	45	29	5	5		16,709
Jan. 1963	18.3	45.3	26.1	4.8	5.4		16,007

(continued)

Table 5. Subjective Well-Being in Japan: Life in Nation Survey^a (Continued)

Percent except where stated otherwise

	Very satisfied	Fairly satisfied	Not very satisfied	Not at all satisfied	Unsure	Don't know or not asked	No. of observations
	<p><i>Literal: "How do you feel about your life at home? Please choose the thing that is closest to how you feel."</i></p> <p><i>Idiomatic: "How do you feel about your life at home? Please choose one of the following."</i></p>						
	<p><i>"Although I can't say I am completely satisfied, I am satisfied" (satisfied)</i></p>						
	"Completely satisfied"	"Somewhat satisfied" (satisfied)	"Somewhat dissatisfied"	"Completely dissatisfied"	"Unclear" (not sure)		
Survey month							
Jan. 1964 ^b	4.4	56.6	33.5	3.4	1.9		16,698
Jan. 1965	4.5	55.7	33.8	4.2	1.8		16,145
Jan. 1966	4.5	53.9	34.4	4.9	2.3		16,277
Feb. 1967	5.2	55.4	33.1	4.2	2.2		16,358
Jan. 1968	6.2	57.9	29.8	4.0	2.0		16,619
Jan. 1969	5.7	57.8	31.0	4.0	1.5		16,848
	<p><i>Literal: "How do you feel about your life now? Which of the following?"</i></p> <p><i>Idiomatic: "How do you feel about your life now? Please choose one of the following."</i></p>						
	<p><i>"Although I can't say I am completely satisfied, I am satisfied" (satisfied)</i></p>						
	"Completely satisfied"	"Somewhat satisfied" (satisfied)	"Somewhat dissatisfied"	"Completely dissatisfied"	"Unclear" (not sure)		
Survey month							
Jan. 1970 ^c	6.0	58.9	29.4	3.8	2.0		16,739
Jan. 1971	4.8	52.6	36.0	4.8	1.8		16,399

Jan. 1972	5.4	54.1	34.8	4.5	1.2	16,985
Jan. 1973	10.0	50.5	32.4	5.5	1.6	16,338
Jan. 1974	3.5	50.4	38.0	6.7	1.3	16,552
Nov. 1974	3.8	46.6	39.9	8.0	1.6	8,123
May 1975	5.5	54.8	33.6	4.7	1.4	8,145
Nov. 1975	4.4	53.9	35.1	5.2	1.4	8,188
May 1976	5.8	55.4	33.2	4.6	1.1	8,343
Nov. 1976	4.7	55.6	33.9	4.5	1.4	8,225
May 1977	9.1	55.1	29.7	4.7	1.4	8,219
May 1978	5.4	58.9	30.6	3.8	1.3	8,116
May 1979	7.1	60.4	28.5	3.1	0.9	8,239
May 1980	5.4	57.2	31.7	4.5	1.1	8,373
May 1981	5.4	58.5	30.5	4.5	1.1	8,348
May 1982	5.7	60.1	29.0	4.0	1.2	8,303
May 1983	5.8	59.0	30.2	4.0	0.9	8,106
May 1984	5.8	59.6	29.8	3.9	0.9	8,031
May 1985	7.3	63.3	25.0	3.6	0.9	7,878
May 1986	6.2	62.0	26.9	4.0	0.9	7,857
May 1987	6.0	58.6	30.5	4.1	0.9	7,981
May 1988	6.2	58.4	30.4	4.1	0.9	7,711
May 1989	5.4	57.7	30.8	5.1	1.0	7,735
May 1990	7.1	59.7	27.8	4.3	1.1	7,629
May 1991	6.7	60.4	28.4	3.7	0.8	7,639

(continued)

Table 5. Subjective Well-Being in Japan: Life in Nation Survey^a (Continued)

Percent except where stated otherwise

	Very satisfied	Fairly satisfied	Not very satisfied	Not at all satisfied	Unsure	Don't know or not asked	No. of observations
	<p><i>Literal: "Overall, to what degree are you satisfied with your life now? Which of the following?"</i></p> <p><i>Idiomatic: "Overall, to what degree are you satisfied with your life now? Please choose one of the following."</i></p>						
	<p><i>"You might say I'm satisfied" (somewhat satisfied)</i></p> <p><i>"Somewhat dissatisfied"</i></p> <p><i>"Dissatisfied"</i></p> <p><i>"I can't say any of the above" (none of the above)</i></p> <p><i>"Don't know"</i></p>						
Survey month	"Satisfied"	"You might say I'm satisfied" (somewhat satisfied)	"Somewhat dissatisfied"	"Dissatisfied"	"I can't say any of the above" (none of the above)	"Don't know"	No. of observations
May 1992	9.3	59.9	21.0	6.3	2.9	0.5	7,504
May 1993	10.3	59.5	20.9	6.4	2.6	0.4	7,327
May 1994	8.3	57.0	23.3	7.7	3.3	0.4	7,608
May 1995	10.4	62.4	19.8	4.8	2.4	0.3	7,347
Jul. 1996	10.3	59.6	21.6	6.2	2.0	0.4	7,303
May 1997	9.8	56.7	22.8	7.8	2.5	0.4	7,293
Dec. 1999	9.5	54.2	23.8	10.4	1.8	0.3	7,022
Sep. 2001	8.1	53.4	26.1	10.2	1.8	0.4	7,080
Jun. 2002	7.9	52.9	26.1	10.7	2.1	0.3	7,247
Jun. 2003	7.2	50.9	28.1	11.5	2.0	0.2	7,030
Jun. 2004	7.2	52.6	26.8	10.5	2.4	0.4	7,005
Jun. 2005	7.7	51.8	27.0	10.5	2.6	0.3	6,924
Oct. 2006	9.4	57.1	25.1	7.4	0.9	0.1	5,941
Jul. 2007	8.3	54.4	26.6	9.4	1.0	0.3	6,086

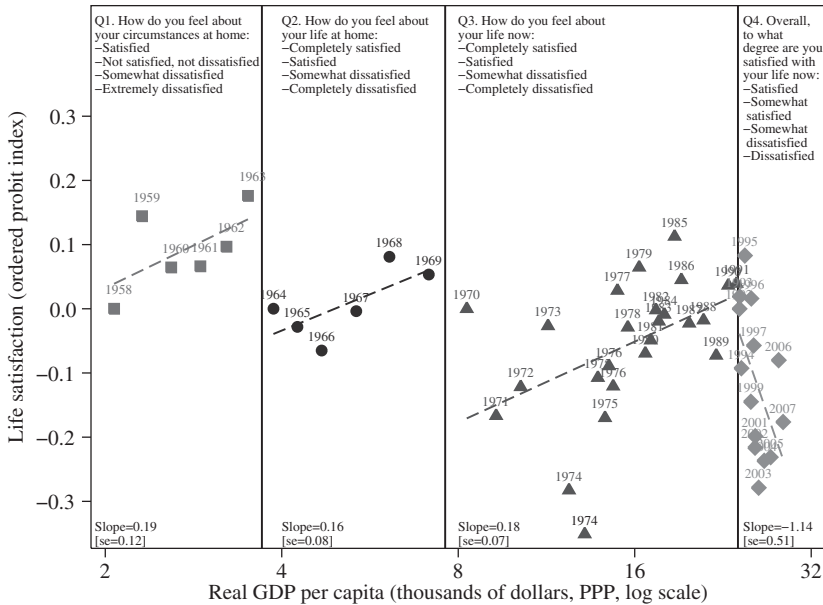
Source: Life in Nation surveys, 1958–2007. www8.cao.go.jp/survey/index-ko.html (in Japanese).

a. Literal translations emphasize the grammatical and lexical form of the source text. Wording in parentheses indicates how the answer was coded; where there is no parenthesis, the answer and the coding are the same. The original Japanese versions of the questions and answers are available in Stevenson and Wolfers (2008).

b. Question is worded literally "How do feel about your life at home? Please choose one answer that is closest to how you feel," yielding an identical idiomatic translation.

c. Question is worded literally "How do feel about your life now? Please choose the answer that is closest to how you feel," yielding an identical idiomatic translation.

Figure 18. Life Satisfaction and GDP per Capita over Time in Japan^a



Source: Life in Nation surveys, 1958–2007.

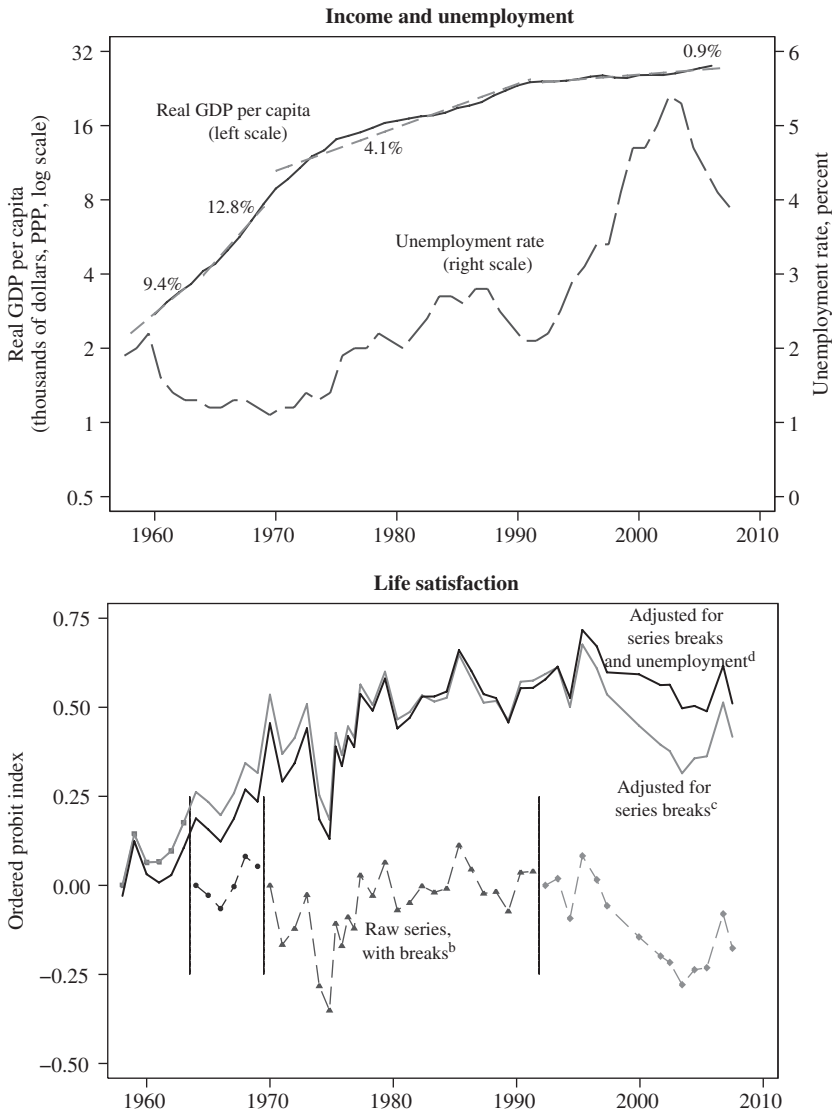
a. The series in each of the four panels reports responses to a different life satisfaction question, and therefore comparisons should be made only within each panel. GDP per capita is at purchasing power parity in constant 2000 international dollars.

on time fixed effects.⁶⁹ By construction, the levels of these four series are not comparable, and hence comparisons within, but not between, series are valid. Figure 18 shows the home/life satisfaction–GDP gradient within each of these periods, and it is clear that throughout the period in which Japan moved from poor to affluent (the first three panels), subjective well-being rose with GDP per capita. The far right panel shows that since 1992 the Japanese economy has shown very little growth, and life satisfaction has fallen sharply.

Figure 19 is a time-series plot of economic progress and subjective well-being in Japan. The top panel shows roughly three episodes in Japanese economic history, corresponding roughly to the changes in the survey questions regarding subjective well-being discussed above: spectacular growth during the period 1958–69, spanning one series break; slower growth from 1970 to 1991; and then anemic growth from 1992 onward, which coincided

69. Table 5 shows the proportions coded as “not sure,” “don’t know,” or “none of the above,” but we simply drop these observations from the rest of the analysis.

Figure 19. Economic Conditions and Life Satisfaction in Japan^a



Source: Life in Nation surveys, 1958–2007; BLS Foreign Labor Statistics; authors' calculations.

a. GDP trend growth rates are estimated by regressing log real GDP per capita on time trends. GDP growth trends are calculated for each period in which comparable satisfaction data were collected (1958–63, 1964–69, 1970–91, 1992–2007).

b. These are four separate raw series, each calculated from an ordered probit run on survey time fixed effects, for the sample periods in which the same survey question was asked (see figure 18). Because the questions changed substantially, these raw indices are comparable only within each period (these levels are normalized so that each series begins at zero). Vertical lines indicate series breaks.

c. The raw life satisfaction indices were pooled, and an OLS regression of satisfaction on series fixed effects, the unemployment rate, and log GDP per capita was run. The adjusted series subtracts the estimate of the series fixed effects from the raw series to yield a regression-adjusted continuous series.

d. The above series is further adjusted for cyclical influences by subtracting the product of the estimated coefficient on unemployment and the contemporaneous unemployment rate.

with the emergence of large-scale unemployment. The symbols in the bottom panel of the figure show the corresponding (and noncomparable across periods) movements in subjective well-being within each of the periods for which consistent data exist.

In an attempt to create a consistent series across the last fifty years, we pool each of these time series and run the following regression to estimate the extent of the relevant series breaks, while controlling for secular and cyclical influences:

$$\begin{aligned} \text{well-being}_t = & -1.67 - 0.26I(1964 \leq \text{year} \leq 1969) \\ & (0.49) \quad (0.07) \\ & - 0.54I(1970 \leq \text{year} \leq 1991) - 0.59I(1992 \leq \text{year}) \\ & (0.11) \quad (0.14) \\ & - 0.063 \text{unemployment rate} + 0.24 \log(\text{GDP per capita})(n = 51). \\ & (0.021) \quad (0.06) \end{aligned}$$

The coefficients on each of the three dummy variables reveal that the changes in the survey question did in fact yield statistically significant (and clearly economically important) changes in estimated well-being. Making adjustments for the series breaks suggested by this regression results in the gray line in the bottom panel of figure 19. This time series suggests that subjective well-being did in fact grow strongly in Japan, at least through the period in which GDP grew most strongly. The regression also finds an important role for unemployment, and this factor explains most of the sharp decline in subjective well-being through the 1990s, as well as the reversal over the past few years as unemployment has started to decline.⁷⁰ The unemployment coefficient is roughly comparable to, although somewhat larger than, estimates for other OECD countries.⁷¹ We can also use this coefficient to back out a “cyclically adjusted” well-being series for Japan, also shown in the bottom panel of figure 19. As should be clear, this series bears a strong relationship with GDP per capita, and indeed, the estimated coefficient, 0.24, is again roughly consistent with our other time-series findings.

Finally, other data also suggest that well-being in Japan has tracked the country’s economic development. For instance, from 1974 through 1991 the same survey also asked, “How do you feel about your life now?” and

70. If instead we estimated this equation without controlling for unemployment, the estimated coefficient on log GDP per capita would be 0.16 (with a standard error of 0.06).

71. Wolfers (2003).

the proportion answering “perfectly complete” or “somewhat complete” trended strongly upward. A somewhat different version of the question was asked from 1992 through 2007, and the proportions feeling perfectly or somewhat complete show a slow decline over this later period. The World Values Survey also provides useful time-series comparisons: in 1981, 16 percent of Japanese respondents reported being very happy, rising to 18 percent in 1990 and then 34 percent in 1995, before falling slightly to 29 percent in 2000. Life satisfaction data from that survey yield a less clear trend, but given the impact of changes in question ordering, it is worth noting that the decline in life satisfaction in Japan was smaller than that experienced in most other countries. Other early assessments of well-being were shown in figure 1: in each of the comparisons in which Japan is included (the 1960 “Patterns of Human Concerns” surveys, the 1965 World Survey, and the 1975 Kettering survey, shown in figure 6), subjective well-being in Japan was consistent with its moderate level of economic development. More recent surveys (such as the World Values Survey or the Gallup World Poll) show that Japan’s well-being is now at a level consistent with its status today as an affluent country.

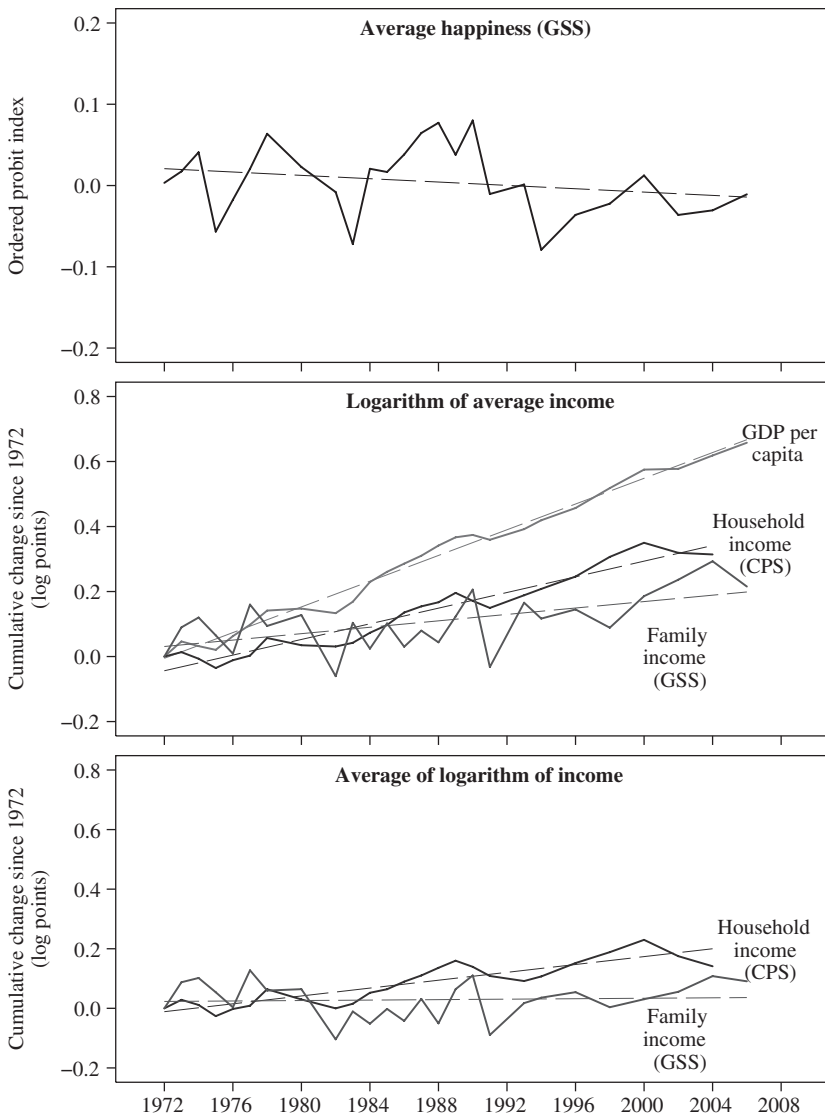
United States

The most widely used dataset for analyzing happiness in the United States is the General Social Survey, conducted on a nationally representative sample of about 1,500 respondents each year from 1972 through 1993 (except 1992), rising to around 3,000 respondents every second year from 1994 through 2004, and 4,500 respondents in 2006. These repeated cross-sectional surveys ask, “Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?” Many researchers have examined the trend in U.S. happiness over this period and all have come to the same conclusion: the United States has not gotten any happier over this time period and has even experienced a mild decline in happiness.⁷² Our analysis turns up similar findings. The top panel of figure 20 plots the coefficients from an ordered probit regression of happiness on year fixed effects.⁷³ These data suggest a very

72. Easterlin (1995) shows a slight decline in happiness between 1972 and 1991 in the United States, and Easterlin (2005a) finds no trend in happiness between 1972 and 2002. Blanchflower and Oswald (2004) report that well-being declined between 1972 and 1998. Stevenson and Wolfers (2007) find that well-being has been flat for men between 1972 and 2006 and has declined among women over this period.

73. We have corrected these data for the biases due to changes in question ordering noted by Smith (1979, 1988). Stevenson and Wolfers (forthcoming) provide a detailed explanation of how these corrections are made and show their impact on individual years.

Figure 20. Happiness and Income over Time in the United States^a



Sources: General Social Survey; Current Population Survey; Bureau of Economic Analysis.
 a. Happiness index is described in figure 13.

mildly declining happiness trend through this period (slope = -0.0010 , with a standard error of 0.0008), which suggests that our happiness index declined by about 0.035 point between 1972 and 2006 (with a 95 percent confidence interval around this decline ranging from -0.09 to $+0.02$).

The middle panel of figure 20 shows that log real GDP per capita rose by 0.66 (or 66 log points) over the same period, and the juxtaposition of this income growth with a roughly flat happiness trend appears to provide useful support for the Easterlin paradox. Indeed, a happiness-income gradient of 0.4 would have led one to expect the happiness index to have risen by 0.26 point. Translating this to the individual happiness categories, we find that U.S. GDP growth from 1972 to 2006 was enough to suggest that by the end of the sample, another 10 percent of the population should have been “very happy,” and the proportions “not too happy” and “fairly happy” should have been about 4 and 6 percentage points lower than actually observed, respectively. Moreover, there is clear evidence of the absence of a time-series happiness-income relationship here—the data clearly reject the view that happiness grew as predicted by the happiness-income gradient estimated within the United States or across countries. Although the U.S. time series is thus a data point supporting the Easterlin paradox, it should be regarded as an interesting exception warranting further scrutiny.

To better understand trends in happiness within the United States and its relationship to recent income growth, we look more closely at the patterns of income growth. In particular, the fruits of economic growth through this period were quite unequally distributed.⁷⁴ From 1972 through 2005, data from the Current Population Survey (CPS) suggest that average real household income grew by only 15 to 20 percent in each of the three bottom quintiles; the fourth quintile experienced growth of nearly 30 percent, and only the top quintile realized income growth of 59 percent.⁷⁵ In turn, the top two quintiles of the household income distribution experienced mild growth in happiness, while happiness actually declined for the bottom three quintiles.

The family income data recorded in the GSS suggest roughly similar real income growth: an average increase of about 32 percent over the full

74. Stevenson and Wolfers (forthcoming) examine how happiness has changed across demographic and socioeconomic groups in the United States. They find that some groups (nonwhites) have gotten happier, while others (those with less than a college degree) have gotten less happy. Overall, however, the variance of happiness has declined both within and between groups.

75. DeNavas-Walt, Proctor, and Mills (2006).

sample, which was quite unequally distributed, with real declines reported in the bottom quintile. Although the CPS data reported above are surely a more reliable indicator of national trends in the income distribution, the family income data collected in the GSS may speak to the characteristics of the particular sample for whom we have happiness data.

Given these unbalanced gains, it is worth asking how the income-happiness link at the micro level aggregates to yield the macroeconomic income-happiness link. In the simple case in which income gains accrue proportionally across the distribution, individual happiness–log income functions aggregate to a macro-level linear relationship between average log income and happiness aggregates. However, the sharp rise in inequality over recent decades drives a large wedge between the rise in the log of average income (which is what we typically observe in macro data) and the average of log income (which is the relevant aggregate for predicting average happiness).

We computed the rise in income inequality in both the CPS and the GSS samples. From 1972 through 2006, the CPS measure of the log of average real household income rose by 41 log points, while inequality—as measured by the mean log deviation—rose by 19 log points.⁷⁶ Together these numbers imply that the average of log household income rose by only 22 log points over the full sample. For the GSS the rise in the log of average family income is slightly smaller, at 32 log points, and the measured rise in inequality (again measured as the mean log deviation) is 15 log points.

Thus, within the GSS sample, the average of the log of family income has risen by only around 17 log points since 1972 (equivalent to an annual rate of growth of only around 0.5 percent a year).⁷⁷ Based on a happiness-

76. It is worth being a bit more explicit about how reasonably robust economic growth translates into weaker growth in the average log income. From 1972 through 2006, real GDP per capita grew by 93 percent, or 66 log points, and disposable personal income per capita rose by a similar amount. Beyond these aggregate data from the Bureau of Economic Analysis (BEA), the Census Bureau also calculates income per household from the March CPS. These alternative data suggest that income per capita (in 2005 dollars) rose by 66 percent, or by 51 log points. Much of the gap between the BEA and the CPS measures reflects differences in deflators. From 1972 through 2006 the CPI-U-RS (the version of the consumer price index used by the Census Bureau to deflate the CPS data) rose 11 log points more than the GDP deflator. This difference would be even larger (22 log points) if we deflated instead by the official CPI-U series. On a per household basis, the rise in the log of average income was even less impressive, at only 41 log points.

77. We deflate the GSS income data using the CPI-U-RS rather than the CPI-U-X. If instead we used the official deflator, the average log of family income would have registered barely any growth at all.

income gradient of around 0.4, it seems reasonable to expect that happiness in the United States would have been basically flat over the past thirty-five years (or, more precisely, to have risen by only $0.4 \times 0.17 = 0.07$ point). Thus, by refocusing our attention on the appropriate macroeconomic aggregate (in the bottom panel of figure 20), it can be seen that the U.S. experience could be roughly consistent with the accumulated evidence of a robust happiness-income link.

Moreover, many other societal trends beyond trends in GDP per capita may influence trends in happiness. Thus, no single national case study can be dispositive in our effort to understand how national well-being changes with economic development. The Easterlin paradox suggests that on average, countries will not get happier as they get richer. The evidence from the countries for which we have time-series observations, on balance, casts doubt on the Easterlin paradox, with Europe, Japan, and countries in the World Values Survey becoming happier, on average, as income grows.

Alternative Measures of Subjective Well-Being

Our discussion so far has analyzed three basic measures of subjective well-being: reports of happiness, of life satisfaction, and of well-being expressed in terms of a “ladder” with the best and worst possible lives at top and bottom. Yet this still leaves a lot unsaid about the subjectively experienced lives of rich and poor. Fortunately, recent major advances in cross-country data collections have started to paint a broader picture of subjective well-being.

We begin by analyzing the battery of ten questions typically grouped as the Bradburn Affect Balance Scale, which were included in the first two waves of the World Values Survey.⁷⁸ This scale is intended to separately assess both positive and negative affect, by probing direct reports of whether various pleasant and unpleasant feelings have been experienced recently. This battery of questions asks the respondent whether, during the past few weeks, he or she has had any of five positive experiences (“pleased about having accomplished something,” “proud because someone had complimented you on something you had done,” “particularly excited or interested in something,” “things were going your way,” or “on top of the world/feeling that life is wonderful”) and five negative experiences (“bored,” “upset because somebody criticized you,” “so restless you

78. Bradburn (1969).

couldn't sit long in a chair," "very lonely or remote from other people," or "depressed or very unhappy").⁷⁹

We analyze each question separately in table 6, and because our dependent variable is binary (whether or not the respondent reported experiencing each feeling), we use probit regressions. To separately isolate the between-country and the within-country variation, we run one regression in which log GDP per capita is the only independent variable, and then another substituting log household income for GDP per capita, and controlling for country fixed effects. To maintain some consistency in the units, we report actual probit coefficients rather than the elasticity of predicted probabilities.

The first panel of table 6 shows that in both within-country and between-country comparisons, measures of positive affect are all positively, and measures of negative affect negatively, associated with income. Although the within-country comparisons are typically statistically significant, the limited number of country observations gives us less precise and hence less significant estimates (we cluster the standard errors in these regressions by country). The magnitudes of the estimated within- and between-country coefficients are roughly similar, and the within-country estimates typically lie in the 95 percent confidence interval surrounding the between-country estimates. Putting these together into a measure of net affect (the average number of positive experiences less the average number of negative experiences) yields a measure that is strongly related to both log household income and log GDP per capita, and these estimates reflect the impact of income on positive and negative affect in roughly equal measure.⁸⁰

Figure 21 presents the cross-country comparisons graphically. The top row reveals that in richer countries a larger proportion of the population is more likely to report each positive experience (except feeling "particularly excited or interested in something"), and the bottom row shows that a smaller proportion of the population in richer countries typically reports negative experiences. The regressions reported in the figure show how the proportion of the population agreeing with each statement rises or falls with log GDP per capita (and hence these estimates are scaled differently

79. Bradburn (1969, chapter 4) found that among his U.S. sample, within the group of positive or negative questions, responses tended to be highly correlated, but that responses between questions probing "positive affect" and "negative affect" were not closely related. Moreover, individual evaluations of happiness appear to reflect positive and negative affect in roughly equal measure.

80. The summary measure of net affect is computed by adding up the positive and negative measures, with zero indicating an equal number of positive and negative experiences.

Table 6. Probit Regressions of Affect on Income: World Values Survey and Gallup World Poll^a

<i>Affect</i>	<i>Percent reporting affect</i>	<i>Probit coefficients</i>	
		<i>Between-country estimates^b</i>	<i>Within-country estimates^c</i>
<i>World Values Survey: Bradburn Affect Balance Scale</i>			
<i>“During the past few weeks, did you ever feel . . .”</i>			
<i>Measures of positive affect</i>			
Pleased	70.5	0.294** (0.150)	0.140*** (0.012)
Proud	40.8	0.493*** (0.179)	0.084*** (0.011)
Excited or interested	53.2	0.054 (0.124)	0.164*** (0.011)
On top of the world	35.1	0.538*** (0.186)	0.071*** (0.011)
Things going your way	49.6	0.148 (0.155)	0.206*** (0.011)
Total positive ^d	2.50 (s.d. 1.54)	0.571** (0.251)	0.234*** (0.011)
<i>Measures of negative affect</i>			
Bored	23.4	-0.223* (0.120)	-0.129*** (0.012)
Upset/criticized	17.8	-0.157*** (0.059)	-0.050 (0.013)
Restless	30.3	-0.112 (0.095)	-0.092*** (0.011)
Lonely	17.0	-0.160 (0.118)	-0.205*** (0.013)
Depressed	20.6	-0.102 (0.117)	-0.173*** (0.012)
Total negative ^e	1.14 (s.d. 1.30)	-0.182 (0.135)	-0.188*** (0.012)
Net affect balance ^f	1.36 (s.d. 1.97)	0.876*** (0.201)	0.421*** (0.017)
<i>Gallup World Poll, 2006</i>			
<i>“Did you experience [insert feeling here] during a lot of the day yesterday?”</i>			
Enjoyment	72.0	0.154*** (0.021)	0.187*** (0.007)
Physical pain	26.7	-0.125*** (0.015)	-0.139*** (0.007)
Worry	34.6	0.009 (0.020)	-0.123*** (0.007)
Sadness	21.7	-0.040*** (0.015)	-0.181*** (0.007)
Boredom	23.9	-0.036* (0.019)	-0.120*** (0.007)

Table 6. Probit Regressions of Affect on Income: World Values Survey and Gallup World Poll^a (Continued)

<i>Affect</i>	<i>Percent reporting affect</i>	<i>Probit coefficients</i>	
		<i>Between-country estimates^b</i>	<i>Within-country estimates^c</i>
<i>Gallup World Poll, 2006</i>			
<i>“Did you experience [insert feeling here] during a lot of the day yesterday?”</i>			
Depression	14.7	−0.094*** (0.023)	−0.182*** (0.008)
Anger	19.8	−0.021 (0.017)	−0.072*** (0.007)
Love	66.3	0.050 (0.029)	0.128*** (0.007)
<i>“Now, please think about yesterday, from the morning until the end of the day. Think about where you were, what you were doing, who you were with, and how you felt.”</i>			
Would you like to have more days like yesterday?	66.9	0.032** (0.016)	0.120*** (0.007)
Did you feel well rested?	65.7	0.027* (0.014)	0.067*** (0.006)
Were you treated with respect?	84.6	0.146*** (0.028)	0.135*** (0.008)
Were you able to choose how you spent your time all day?	70.0	0.035* (0.018)	0.030 (0.006)
Did you smile or laugh a lot?	70.6	0.103*** (0.017)	0.148*** (0.007)
Were you proud of something you did?	59.3	0.012 (0.023)	0.120*** (0.007)
Did you learn or do something interesting?	52.5	0.029 (0.022)	0.149*** (0.007)
Did you eat good tasting food?	74.1	0.194*** (0.021)	0.222*** (0.007)

Source: Authors' regressions.

a. The (binary) dependent variable in each regression is the answer (yes or no) to the survey question. All regressions control for respondent gender, a quartic in age, their interaction, and indicators for missing age or gender. Numbers in parentheses are robust standard errors. Asterisks indicate statistical significance at the *10 percent, **5 percent, and ***1 percent level. s.d., standard deviation.

b. Probit regression of affect measure on log real GDP per capita, clustering standard errors by country. Sample sizes vary by question, but the Gallup World Poll typically yielded around 134,000 respondents from 130 countries, while the World Values Survey yielded around 66,000 respondents from 31 or 32 countries with nationally representative samples.

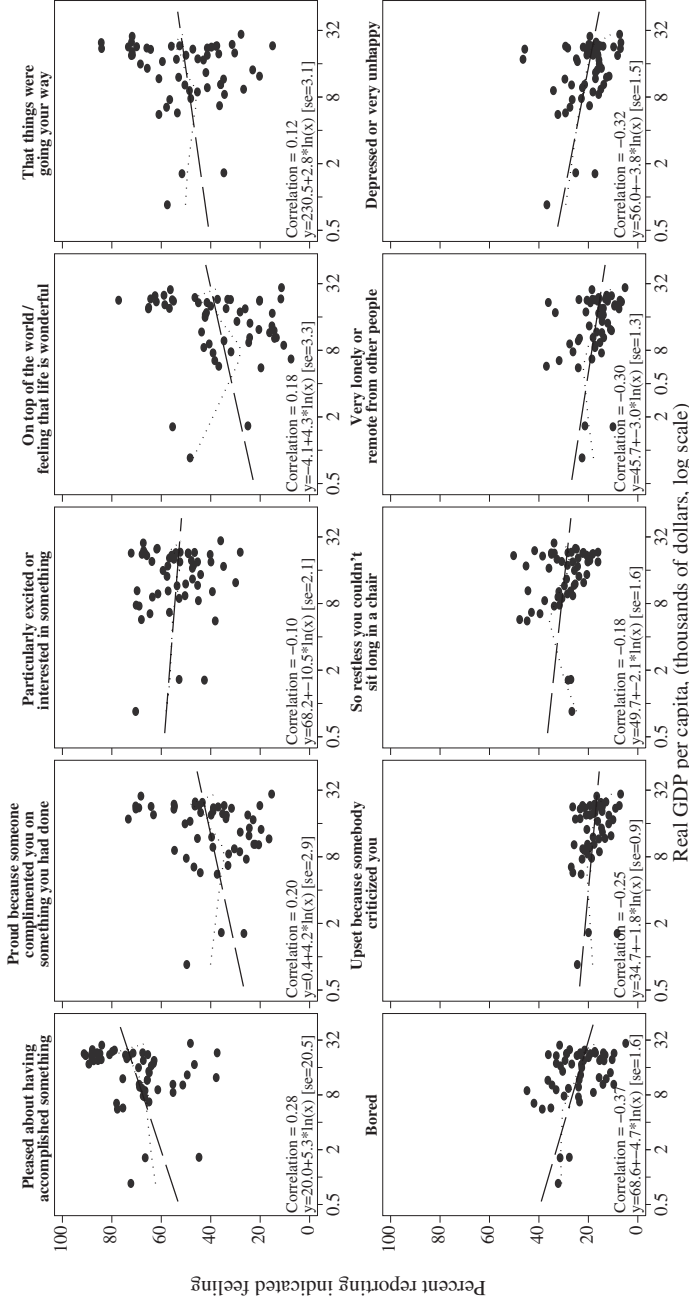
c. Probit regression on log household income, further controlling for country fixed effects (and hence exploiting only within-country income comparisons). Because these regressions also require valid household income data, the sample size was smaller: typically the Gallup World Poll yielded around 100,000 respondents from 113 countries, while the World Values Survey yielded around 42,000 respondents from 24 countries.

d. Number of positive affect questions answered affirmatively (0–5).

e. Number of negative affect questions answered affirmatively (0–5).

f. Total positive affect minus total negative affect.

Figure 21. Cross-Country Measures of Recent Feelings and GDP: World Values Survey^a



Source: World Values Survey, 1981–84 and 1989–93 waves. Sources for GDP per capita are described in the text.

a. Questions are those typically grouped as the Bradburn Affect Balance Scale. Respondents were asked, “We are interested in the way people are feeling these days. During the past few weeks, did you ever feel...?” Top and bottom rows show responses to questions relating to positive and negative affect, respectively. Each observation represents one of fifty-four nationally representative country-wave samples drawn from forty developed and developing countries. Dashed lines are fitted from OLS regressions of the percent agreeing with the statement on log real GDP per capita; dotted lines are fitted from lowest estimations. GDP per capita is at purchasing power parity in constant 2000 international dollars.

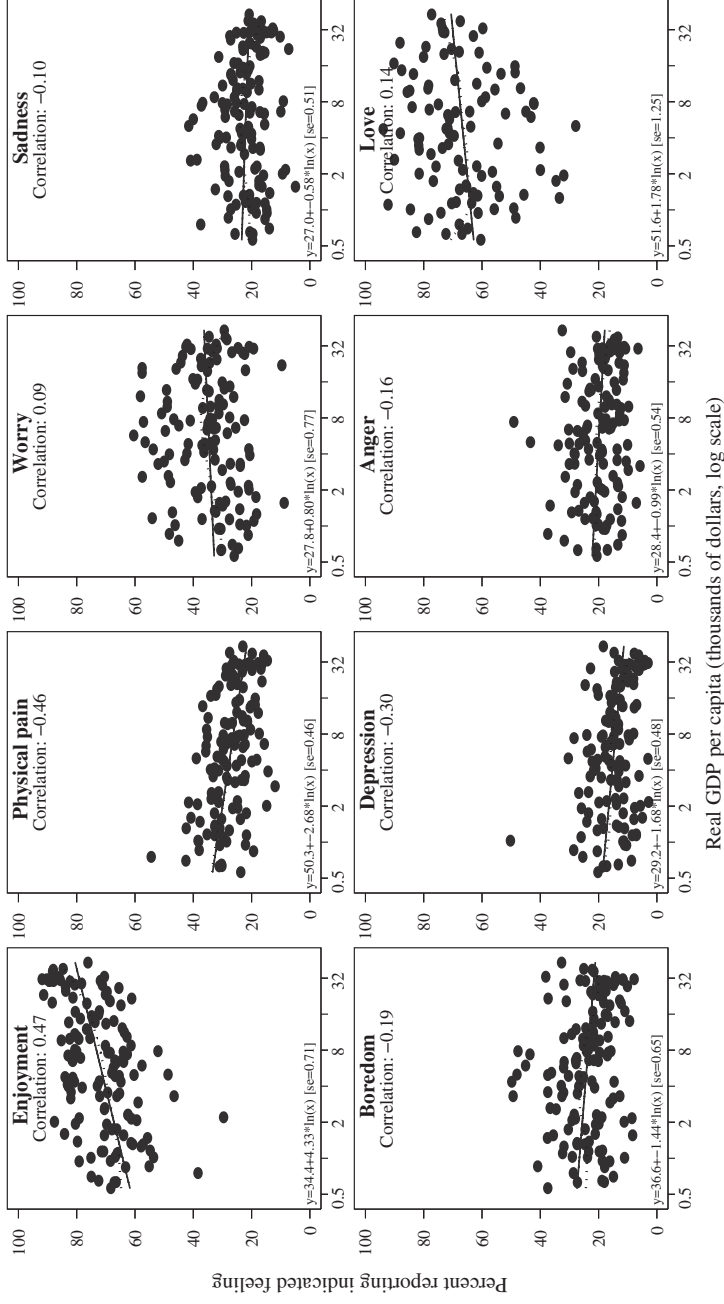
than the probit coefficients reported in table 6). Interestingly, as in the analysis of self-assessed happiness, Nigeria (the poorest country in this figure) is an outlier for all of the measures of positive affect, with Nigerians reporting a much higher likelihood of experiencing positive feelings than residents of other low-income countries. The bottom row shows the relationship between each of the five measures of negative feelings and GDP per capita. In all cases the negative affect–GDP gradient is negatively sloped, with a higher proportion of people in poor countries experiencing negative feelings. (These measures of negative affect suggest that Nigerians have a more typical experience for their income.)

We next turn to a particularly rich series of well-being questions contained in the Gallup World Poll. Respondents are asked to report whether they experienced a given feeling “during a lot of the day yesterday.” The feelings include enjoyment, physical pain, worry, sadness, boredom, depression, anger, and love. The first panel of Gallup results in table 6 shows that among the positive emotions, the enjoyment-income gradient is positive and similar for both the between- and the within-country estimates. More income is clearly associated with more people having enjoyment in their day. Love is less clearly related to income, although within countries, more income is associated with being more likely to experience love. Among the negative emotions, physical pain, boredom, depression, and sadness are all lower at higher levels of income, at both the national and the individual levels. The within-country estimates also reveal that worry and anger fall with income.

Figure 22 allows a fuller examination of the proportion of people in a country experiencing these emotions and GDP per capita. The percent of people in a country who enjoyed the previous day rises from an average of 65 percent in low-income countries to 80 percent in the wealthiest countries. Depression, pain, boredom, and anger all appear to fall linearly with rises in log GDP per capita. The magnitudes of these relationships are large: compared with the poorest countries, those in the wealthiest countries are a third less likely to experience pain or depression and a fifth less likely to report boredom.

The final set of regressions analyze the relationship between income and some more specific experiences in people’s lives, such as feeling respected, smiling, engaging in interesting activities, feeling proud, and learning. Income is positively related to wanting to have more days like yesterday, with feeling well rested, with feeling treated with respect, with being able to choose how to spend one’s time, with smiling or laughing, with feeling proud, with having done something interesting, and with eating

Figure 22. Cross-Country Measures of Recalled Feelings and GDP: Gallup World Poll^a



Source: Gallup World Poll, 2006. Sources for GDP per capita are described in the text.

a. Respondents were asked, "Did you experience [insert feeling here] during a lot of the day yesterday?" Each observation represents one of up to 130 developed and developing countries in the sample (questions were not asked in Iraq). Dashed lines are fitted from ordinary least squares regressions of the percent agreeing with the statement on log real GDP per capita; dotted lines are fitted from lowess estimations. GDP per capita is at purchasing power parity in constant 2000 international dollars.

good-tasting food. For most of these assessments the within-country coefficients are larger than the between-country coefficients, although there are some notable exceptions. Figure 23 plots the proportion of people having these experiences in each country against GDP per capita, and two of these—feeling respected, and having eaten good food—display particularly strong relationships with GDP per capita. The proportion of people smiling or laughing rises with income both within and between countries, although the coefficient estimates point to a somewhat stronger relationship to income within countries. This last measure is particularly interesting, as smiling has been shown to be correlated with reported levels of happiness or life satisfaction. Indeed, in these data, people who report smiling more also tend to report greater life satisfaction. Finally, both table 6 and figure 23 point to an increasing ability to choose how one spends one's time during the day as income rises.

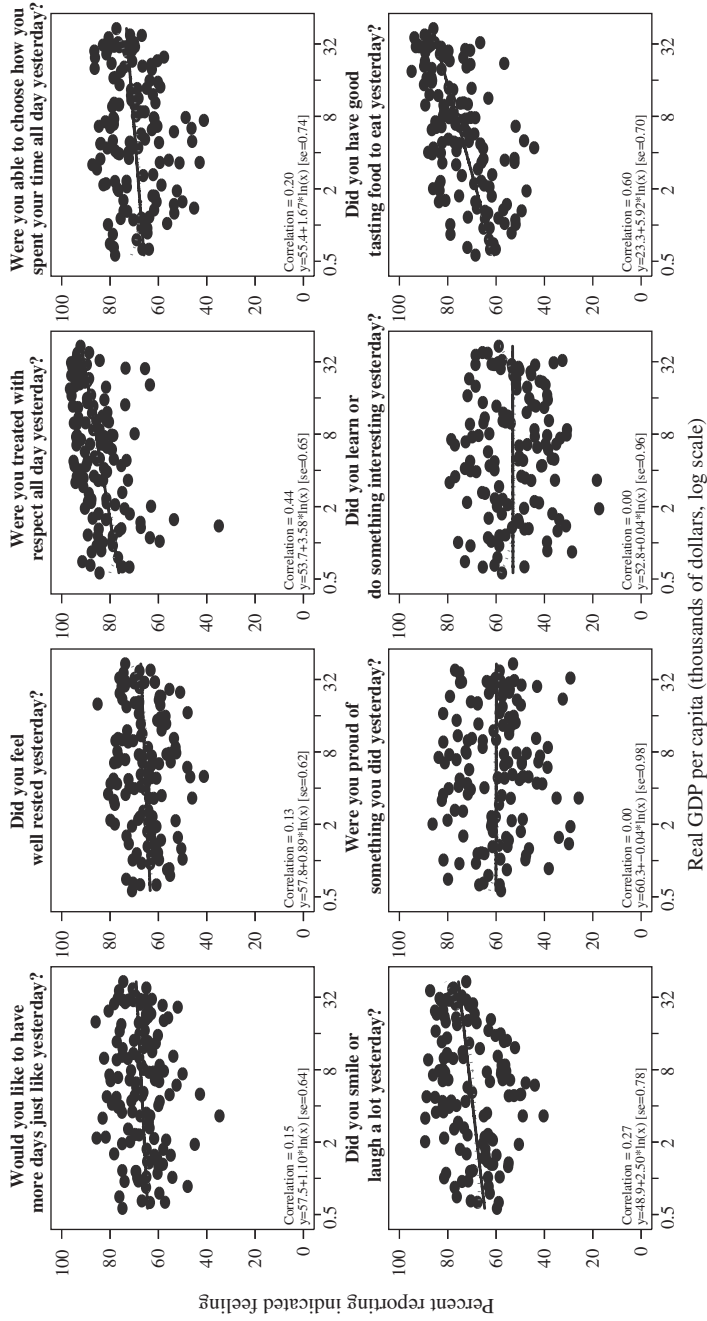
All told, these alternative measures of well-being paint a somewhat more nuanced picture of the different experiences of rich and poor people within countries, and between rich and poor countries. Moreover, these data point to a robust relationship between greater income and greater reported well-being. We suspect that these intriguing new cross-country data collections will launch a productive research program aimed at better understanding the drivers of the robust well-being-income gradient we have identified.

Discussion

This paper has revisited—and where appropriate, revised—the stylized facts regarding the link between subjective well-being and income. Our analysis encompasses virtually all of the extant data linking happiness or life satisfaction to income. Moreover, we have endeavored to place this analysis in a single coherent framework that allows us to make meaningful comparisons across different surveys and different ways of asking about subjective well-being. We were motivated to better understand the Easterlin paradox, and so we have analyzed separately three relationships between income and happiness: that obtained from contrasting rich and poor members of a society, that obtained from contrasting rich and poor countries, and that obtained from observing the paths of average happiness as the average incomes of countries change. Our measurement framework allows us to assess the extent to which these relationships may differ.

Our key contribution is the finding that the relationship between subjective well-being and income within countries (that is, contrasting the

Figure 23. Cross-Country Measures of Daily Experiences and GDP: Gallup World Poll^a



Source: Gallup World Poll, 2006. Sources for GDP per capita are described in the text.

a. Questions were prefaced as follows: "Now, please think about yesterday, from the morning until the end of the day. Think about where you were, what you were doing, who you were with, and how you felt." Each observation represents one of up to 130 developed and developing countries in the sample (questions were not asked in Iraq). Dashed lines are fitted from OLS regressions of the percent agreeing with the statement on log real GDP per capita; dotted lines are fitted from lowest estimations. GDP per capita is at purchasing power parity in constant 2000 international dollars.

happiness of rich and poor members within a country) is similar to that seen between countries, which in turn is similar to the time-series relationship (comparing the happiness of countries at different points in time as they get richer or poorer). In multiple datasets from several decades and covering various populations, we estimate well-being-income gradients that tend to be centered around 0.4. We estimate slightly steeper gradients when comparing well-being between countries, although reading across datasets and taking account of sampling error, we can reject neither the hypothesis that the gradients are the same within and between countries, nor the hypothesis that there are small differences between the two.

The time-series part of our analysis is necessarily only suggestive: repeated (and comparable) surveys of subjective well-being data are both noisy and scarce, and hence they speak less clearly. In many cases we find happiness within a country rising during periods of economic growth and rising most rapidly when economic growth is more rapid. The United States stands out as a notable exception: Americans have experienced no discernable increase in happiness over the past thirty-five years (and indeed, happiness among U.S. women has declined). In contrast, Japan stands out as a remarkable success story, recording rising happiness during its period of rapid economic growth. So, too, life satisfaction has trended upward in Europe, and this trend has been most evident in those countries in which economic growth has been most robust. All told, our time-series comparisons, as well as evidence from repeated international cross sections, appear to point to an important relationship between economic growth and growth in subjective well-being. Quantitatively, the time-series well-being-GDP gradient yields a role for income similar to that seen in our within- and between-country contrasts. Taken as a whole, the time-series evidence is difficult to reconcile with earlier claims that economic growth yields no boost to happiness.

Easterlin and others have argued that comparisons of rich and poor people within a country yield starker happiness differences than comparisons of rich and poor countries, and have cited this as evidence that relative income differences are a key driver of happiness. Carol Graham notes that “a common interpretation of the Easterlin paradox is that humans are on a ‘hedonic treadmill’: Aspirations increase along with income and, after basic needs are met, relative rather than absolute levels matter to well-being.”⁸¹ In its strong form this hypothesis suggests that people (and public policy) are powerless to deliver lasting gains in happiness, because individual happiness

81. Graham (2008, p. 77).

returns inexorably to one's set point of happiness. Our findings clearly falsify this strong form of adaptation: we find that those enjoying materially better circumstances also enjoy greater subjective well-being and that ongoing rises in living standards have delivered higher subjective well-being. However, milder forms of adaptation are potentially consistent with our findings.

Our findings point to an important role for absolute levels of income in shaping happiness and a lesser role for relative income comparisons than was previously thought. Equally, our findings are sufficiently imprecise that they may still admit a role for relative income comparisons in shaping subjective well-being. We find that estimates of the within- and between-country well-being-income gradient tend to cluster around 0.4, and we lack sufficient evidence to say that these gradients are clearly different. Thus, our evidence is consistent with the view that only absolute income matters to happiness (which would imply that the within- and between-country estimates are identical). Indeed, whereas previous analyses of the link between income and happiness suggested a *prima facie* case for relative income playing a dominant role, our updated reanalysis finds no such case.

Equally, our findings do admit the possibility of an interesting role for relative income comparisons. For instance, the within-country coefficient is typically about 0.3 and might be biased downward by the influence of transitory income in the cross section. Thus, perhaps the true within-country coefficient is 0.45, and our estimates are consistent with a view that the between-country coefficient is about 0.36 (with the time-series coefficient a bit weaker still). This is consistent with both absolute and relative income impacting well-being, with the former having a weight about four times as large as the latter.⁸² Thus, our findings should not be interpreted as falsifying the view that relative income plays a role in shaping happiness, although they do bound the extent to which relative income may matter.

In light of this range of possible interpretations, we would suggest that more fine-grained evidence on the role of relative income should come from direct evidence of relative income shocks, such as those investigated by Erzo Luttmer.⁸³ In particular, a comparison of the between and within well-being-income gradients casts doubt on a large role for intranational relative income comparisons in determining happiness. Although these comparisons do not speak to the role of international relative income com-

82. For instance, these findings are consistent with a simple happiness function: $\text{happiness} = 0.36 \log(\text{individual income}) + 0.09 \log(\text{individual income} \div \text{average income})$.

83. Luttmer (2005).

parisons, the earliest surveys—conducted in the 1940s—yielded similar estimates of the between-country well-being-income gradient to those seen today. However, these early surveys involved sufficiently few data points that it is impossible to know with precision whether the gradient between subjective well-being and income across countries has changed over time to reflect the growing awareness of the opportunities available to others around the world. In short, the most compelling evidence for the importance of absolute income over relative income in determining happiness may eventually come from the time-series evidence.

Finally, we should note that our analysis has largely focused on establishing the magnitude of the bivariate relationship between subjective well-being and income, rather than tracing the causal effects of income on happiness. We believe that further research aimed at better understanding the causal pathways will be fruitful.

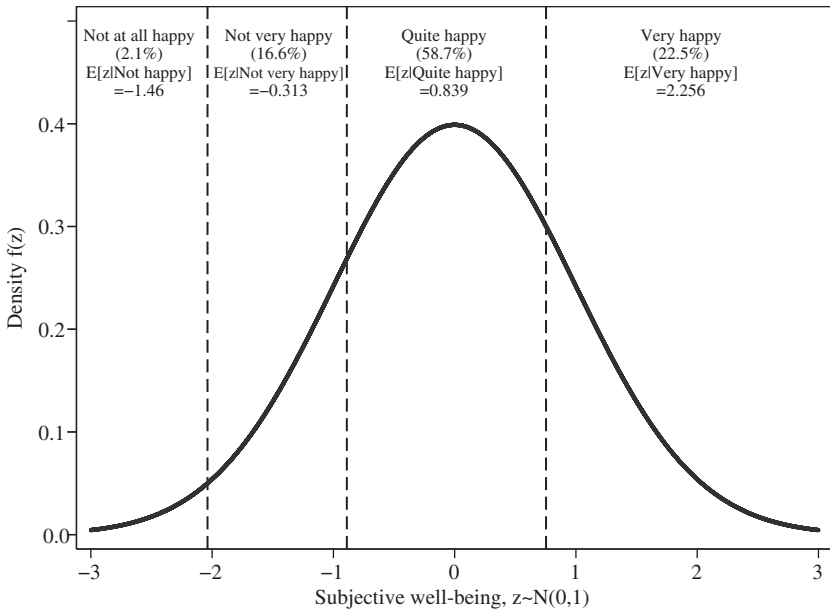
ACKNOWLEDGMENTS The authors would like to thank Gary Becker, Daniel Blanchflower, Angus Deaton, Richard Easterlin, Carol Graham, Daniel Kahneman, Alan Krueger, David Laibson, Andrew Oswald, and Luis Rayo for useful discussions, and Gale Muller and his colleagues at Gallup for access to and help with the Gallup World Poll. Rohak Doshi and Michael L. Woodford provided excellent research assistance. We would like to thank the Zicklin Center for Business Ethics Research and the Zell/Lurie Real Estate Center for generous research support. The data and Stata programs used in this paper are available for download from the authors' homepages.

APPENDIX A

Cardinalizing Happiness and Life Satisfaction

Our approach to constructing an index of average well-being in a country-year (or country-wave) is to report the coefficient from an ordered probit regression of subjective well-being on country by year (or country by wave) fixed effects. This appendix tries to make this approach more transparent, thereby demonstrating how to reconcile our results with alternative approaches.

A simple approach to aggregating data on subjective well-being involves arbitrarily assigning to qualitative categories scores equal to their rank order. Thus, in the World Values Survey, a response of “not at all happy”

Figure A1. Fitting an Ordered Probit^a

Source: World Values Survey, 1981–2004.

a. Happiness question asks, “Taking all things together, would you say you are: ‘very happy,’ ‘quite happy,’ ‘not very happy,’ [or] ‘not at all happy?’” Figure shows the assumed normal distribution of a latent happiness variable when running an ordered probit regression; by a standard normalization, this distribution has a mean of zero and a standard deviation of one. Dashed lines represent the cutpoints estimated from running an ordered probit regression of happiness on country \times wave fixed effects. Thus, the country \times wave fixed effects are equivalent to scoring “not at all happy” as -1.460, “not very happy” as -0.313, “quite happy” as 0.839, and “very happy” as 2.256.

is given a value of 1, “not very happy” a value of 2, “quite happy” a value of 3, and “very happy” a value of 4. Average well-being is then calculated as the simple average of these values. This appears to be the most common approach in the literature.

A key difficulty with this approach is that the scaling of well-being measures from different surveys will vary, depending on whether the question used a three-, four-, five-, seven-, ten-, or eleven-point scale (others also occur). In turn, this approach yields estimates of the well-being-income gradient that are neither comparable across surveys nor have an obvious economic interpretation.

Thus, a somewhat more satisfying index might be constructed by normalizing the dependent variable (subtracting its mean and dividing by its standard deviation), which would yield a common metric. Moreover, this

Table A1. Alternative Scalings of Survey Responses

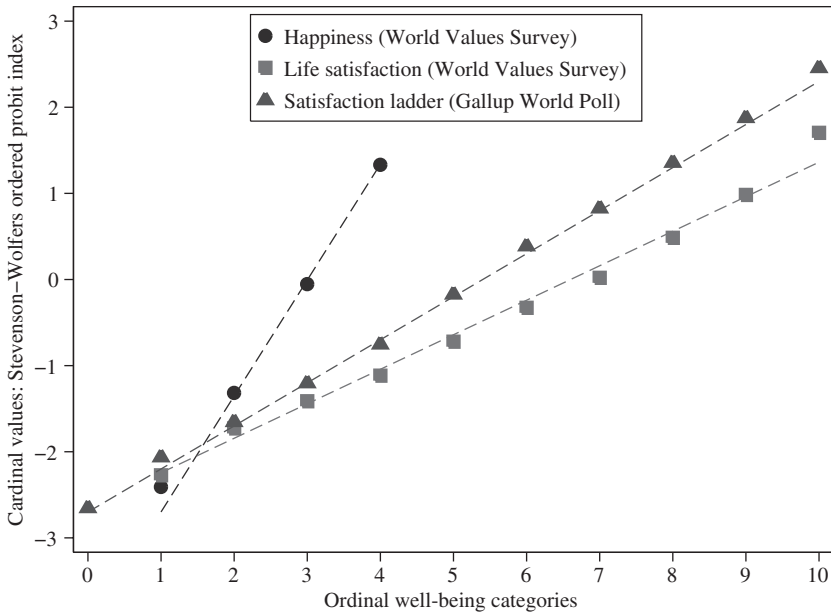
Verbal description	World Values Survey: Happiness		World Values Survey: Life satisfaction			Gallup World Poll: Life satisfaction			
	Simple coding ^a	Standardized ^b	Our method ^a	Simple coding	Standardized	Our method	Simple coding	Standardized	Our method
Not at all happy	1	-2.70	-2.41	1	-2.24	-2.27	0	-2.37	-2.65
Not very happy	2	-1.35	-1.32	2	-1.84	-1.73	1	-1.92	-2.06
Quite happy	3	-0.01	-0.05	3	-1.44	-1.41	2	-1.48	-1.65
Very happy	4	1.34	1.33	4	-1.04	-1.12	3	-1.03	-1.20
				5	-0.64	-0.72	4	-0.59	-0.75
				6	-0.24	-0.33	5	-0.14	-0.17
				7	0.16	0.02	6	0.30	0.39
				8	0.56	0.48	7	0.75	0.83
				9	0.96	0.98	8	1.19	1.36
				10	1.36	1.70	9	1.64	1.88
							10	2.08	2.46

Source: Authors' regressions.

a. Simple coding gives each category a score equal to its ordered rank.

b. Standardized values take the simple coding, subtract its mean, and divide by the standard deviation

c. Our method involves running an ordered probit of well-being on country \times wave fixed effects. Estimates shown are the expected value of a latent happiness index, conditional on being in each category.

Figure A2. Cardinal and Ordinal Measures of Well-Being^a

Sources: Gallup World Poll, 2006; World Values Surveys, 1981–2004; authors' calculations.

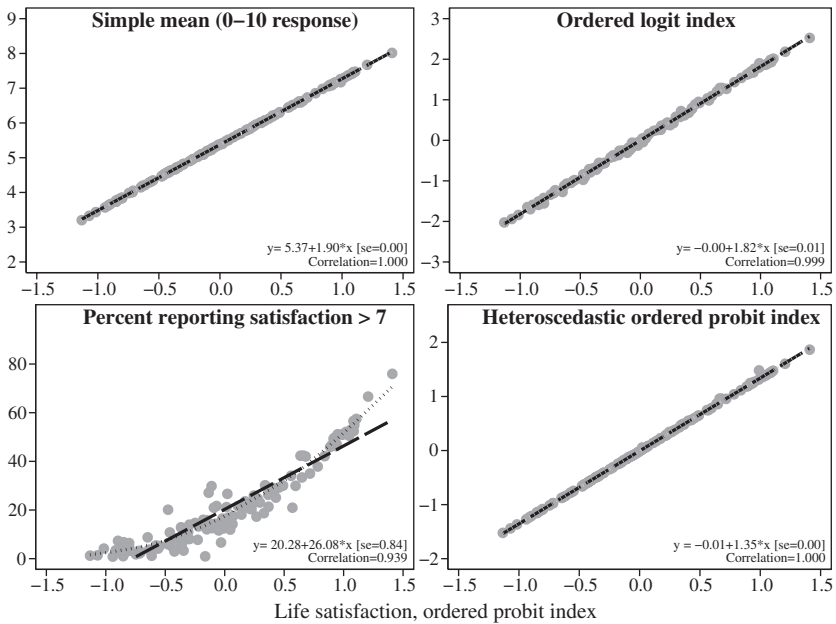
a. The horizontal axis shows each ordered well-being category; the vertical axis shows the cardinal value implicitly given to these values by the ordered probit regression on country \times wave fixed effects (see figure A1). Dashed lines show the cardinalization derived by simply treating the ordinal rankings as cardinal, and standardizing by subtracting the mean and dividing by standard deviation. In the World Values Survey, happiness responses are coded as 1, not at all happy; 2, not very happy; 3, quite happy; 4, very happy; life satisfaction responses are coded on a 1-to-10 scale, with 10 representing most satisfied. In the Gallup World Poll, respondents are shown a picture of a ladder with ten rungs, with the top rung (most satisfied) coded as 10.

metric would have an economic interpretation, scaling differences in well-being relative to its cross-sectional standard deviation. (This approach yields results very close to our approach.)

Even so, the limitation of this approach is that it imposes a linear structure, implying, for example, that the difference between “not very happy” and “not at all happy” is equal to the difference between “quite happy” and “not very happy.” Although psychologists have often been willing to accept that the subjective distances between successive points on category scales are similar, we can use data on the proportions of the population who report themselves as being in each category to relax (or test) this assumption.

To make use of these population proportions, the ordered probit (figure A1) makes a parametric assumption, imposing normality on the distribution of the underlying latent “well-being” measure. Two normalizations

Figure A3. Alternative Estimates of Life Satisfaction: Gallup World Poll, 2006^a



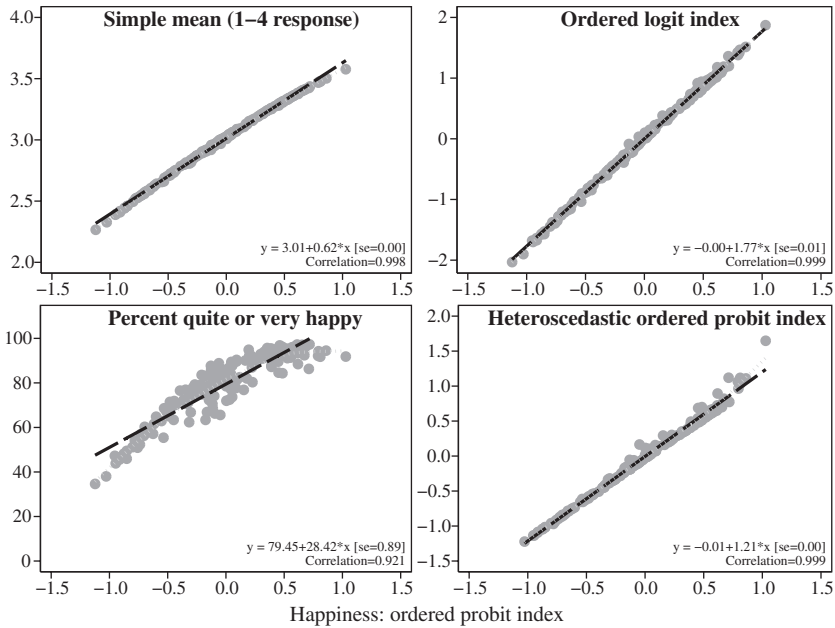
Sources: Gallup World Poll, 2006; authors’ calculations.

a. Each point shows the estimated average level of satisfaction in one of 131 countries, comparing one of four alternative metrics on the vertical axis with the ordered probit index on the horizontal index. Dashed lines are fitted from the reported OLS regression; dotted lines are fitted from lowest estimation.

are also imposed: that the latent variable has a mean of zero and that it has a standard deviation of one. The country or country by wave fixed effects that we estimate (and interpret as well-being) are simply shifts in the mean of this distribution.

There is a very simple mapping between our results and the simple approach described above: whereas the “value” of each categorical answer is simply imposed in the simple approach, in our approach it is equal to the expected value of a standard normal variable, conditional on being between the estimated upper and lower cutoff points. Van Praag and Ferrer-i-Carbonell (2004) describe this as “probit-adapted OLS.” Table A1 reports the mapping between the underlying categorical responses, the standardized categorical responses, and our scaling derived from these ordered probits.

As the table shows, our method yields a cardinalization that is very similar to that obtained simply by standardizing the variables used in the usual

Figure A4. Alternative Estimates of Average Happiness: World Values Survey^a

Sources: World Values Surveys, 1981–2004; authors' calculations.

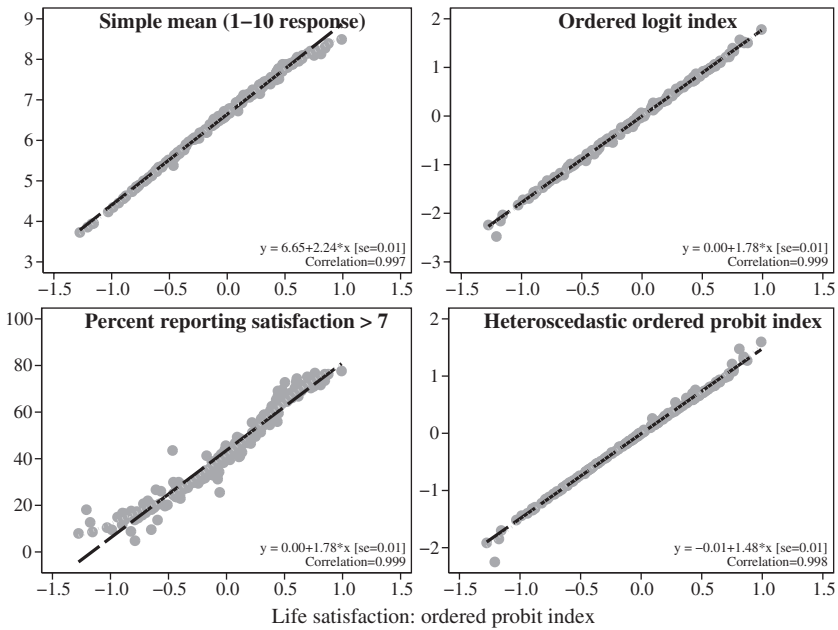
a. Each point shows the estimated average level of happiness in one of 131 countries, comparing one of four alternative metrics on the vertical axis with the ordered probit index on the horizontal index. Dashed lines are fitted from the reported OLS regression; dotted lines are fitted from lowest estimation.

approach. This provides a useful approximation: to map results in other studies onto ours, one can simply divide the estimates of the well-being-income gradient estimated in those studies by the standard deviation of well-being. These results are graphed in figure A2, which shows the cardinalization imposed by our ordered probit procedure in each of three key datasets. As should be clear, our procedure is well approximated by a linear transformation of the simpler approach, which simply analyzes the ordered categories directly.

Next, it is worth assessing this approach relative to four alternative metrics, of which three are typically used in the literature; the fourth is an interesting extension of our approach.

—*Means*: Continuing with the most common approach in the literature, the simplest (and most transparent) approach is to take the ordinal ranking of alternatives as cardinal measures of happiness. This approach

Figure A5. Alternative Estimates of Average Life Satisfaction: World Values Survey^a



Sources: World Values Surveys, 1981–2004; authors’ calculations.

a. Each point shows the estimated average level of satisfaction in one of 131 countries, comparing one of four alternative metrics on the vertical axis with the ordered probit index on the horizontal axis. Dashed lines are fitted from the reported OLS regression; dotted lines are fitted from lowest estimation.

may make more sense when analyzing questions that ask respondents to give a cardinal response (such as the World Values Survey life satisfaction question, which asks for a response on a scale of 1 to 10).

—*Population proportions*: An alternative involves reporting the proportion of the population reporting themselves as, say, “quite happy” or “very happy.” This approach has the advantage that it yields a natural scaling (from 0 to 1) and is directly interpretable. One difficulty is that this approach may lead changes in the dispersion of happiness to be interpreted as changes in the average level of happiness. To minimize this possible confound, one typically chooses a cutoff near the median response. However, the median response in poor countries can turn out to be a far more common response in rich countries.

—*Ordered logits*: The ordered logit is similar to our ordered probit approach but assumes a slightly different (fatter-tailed) distribution of the latent “happiness” in the population. The logistic function also imposes a

standard deviation on the latent variable of $\pi/\sqrt{3}$, which makes the coefficients somewhat differently scaled than with the ordered probit.

—*Heteroscedastic ordered probit*: The ordered probit imposes an equal variance in residual happiness, whereas the heteroscedastic ordered probit allows both the mean and the variance of happiness to vary by country-year. Alternatively phrased, this approach relaxes the assumption of similar cutoff points for each country and year, allowing proportional shifts in these cutoff points, by country-year.⁸⁴

Figures A3 through A5 compare these alternative aggregators with our ordered probit approach, analyzing separately the satisfaction ladder from the Gallup World Poll and the life satisfaction and happiness data, by country and wave, in the World Values Survey. These figures suggest that alternative methods of aggregating subjective well-being all tend to yield highly correlated estimates.

APPENDIX B

Comparing Countries in the World Values Survey

Samples in some low-income countries in certain waves of the World Values Survey were explicitly not designed to be representative of their entire population. These selected samples add measurement error that is, in many cases, correlated with income, education, and other factors related to subjective well-being. In most cases these nonrepresentative samples lead average subjective well-being to be overestimated relative to the population mean. Moreover, nonrepresentative sampling typically occurred in countries with low GDP per capita. For many of these countries, the sampling frame changed in successive waves to become more representative of the entire population (and this change occurred in parallel with rising GDP). Thus, we should expect that for these countries average subjective well-being in the population will decline over time as more rural, low-income, and less educated citizens are included in the sampling frame.

In the results presented throughout the paper, we have excluded a few countries in particular waves because the survey documented that the sampling frame was not representative of the entire country, and no compensatory sampling weights are provided.⁸⁵ In this appendix we detail the

84. Stevenson and Wolfers (forthcoming) provide greater detail on this method.

85. Many samples oversample specific groups, but sampling weights are provided in order to yield nationally representative estimates. Sampling weights cannot adjust for the fact that some groups were not sampled at all.

reasons why these observations were excluded and show how our results are affected when these country-wave observations are included.

We begin by documenting the sampling issues specific to countries that are impacted:

—*Argentina* was surveyed in all four waves; however, in the first three waves sampling was limited to the urbanized central portion of the country and resulted in a wealthier, more educated sample of Argentines than the population average. In the 1999–2004 wave the sample was designed to be representative of the entire country. We include in our analysis only observations from Argentina in the 1999–2004 wave.

—*Bangladesh* was surveyed in two waves. In the 1994–99 wave the survey oversampled men and people in urban areas “to reflect the fact that awareness is relatively more widespread in the urban areas.”⁸⁶ Sampling weights are not provided, and it is therefore not possible to correct for the oversampling. The 1999–2004 wave was designed to be representative. We include in our analysis observations from Bangladesh in the 1999–2004 wave only.

—*Chile* was surveyed in three waves; however, in the 1989–93 and 1994–99 waves the sample was limited to the central portion of the country, which contains slightly fewer than two-thirds of the population and has an average income about 40 percent higher than the national average. In the 1999–2004 survey the sample was drawn from twenty-nine selected cities. As a result of these partial-country samples, we exclude Chile from all of our analysis.

—*China* was surveyed in three waves. For the 1989–93 wave the survey notes state that the researchers “undersampled the illiterate portion of the public and oversampled the urban areas and the more educated strata.” Moreover, the survey notes explicitly state that “the oversampled groups tend to have orientations relatively similar to those found in industrial societies” and that “the data probably underestimate the size of cross-national differences.”⁸⁷ In the 1994–99 wave, a random sample of central China, which contains about two-thirds of the population, was done. Sampling weights are not provided for any of the waves. These two surveys are quite different from each other: in the first wave only 1 percent of the sample were from a town with fewer than 50,000 people, whereas in the second wave 63 percent were. In the first wave 60 percent of respondents were

86. World Values Survey, survey notes for Bangladesh (BD_WVS 1995), available at www.worldvaluessurvey.com.

87. World Values Survey, survey notes for China (CN_WVS 1990).

men; the proportion falls to 53 percent in the second wave. In the 1999–2004 wave the sampling frame was drawn from a previous nationally representative survey and was conducted throughout the entire country, with the exception of six remote provinces with 5.1 percent of the total population. The sample was also limited to persons ages 18 to 65. Despite some limitations, we believe that the last wave is approximately representative and include observations from this last wave in our analysis. Observations from the earlier waves are excluded. If we were to also exclude the final wave, there would be no notable impact on our analysis.

—The *Dominican Republic* was surveyed only in the 1994–99 wave. The sample included only 18- to 49-year-olds and only four communities were chosen to be surveyed. We exclude observations from the Dominican Republic from our analysis.

—*Egypt* was surveyed only in the 1999–2004 wave. The survey notes that a disproportionately large percentage of housewives were included; examining the survey, we find that women are also disproportionately from large urban areas, in particular Cairo. Since no sample weights are provided, we exclude Egypt from our analysis.

—*India* was surveyed in three waves. In the 1989–93 wave the sample was designed such that 90 percent of respondents were literate (compared with a population average of fewer than 50 percent). Interviews were carried out in the eight most widely spoken languages of India, but the rural 10 percent of the sample was confined to the five (out of fourteen) Hindi-speaking states in the sample. In the 1994–99 wave the survey was conducted in Hindi only (the language of fewer than half of the general population), and the sample was stratified to allocate 90 percent of the interviews to urban areas and 10 percent to rural areas. In 1999–2004 the survey was designed to be representative of 97 percent of the population and was conducted in ten languages. Sample weights were not provided for any of the waves. We include only this last wave in our analysis.

—*Nigeria* was surveyed in three waves. The 1989–93 and 1994–99 waves focused on the literate and urban portion of the population: over 40 percent of the respondents in the first wave had attended university. This proportion falls to 23 percent in the second wave, which included a larger rural sample, and to 12 percent in the 1999–2004 wave, which was designed to be representative of the population. We include only this last wave in our analysis.

—*Northern Ireland* was included in the 1999–2004 wave. We exclude Northern Ireland from the analyses because of missing GDP data.

—*Pakistan* was surveyed in two waves; however, in the 1994–99 wave sampling was done only in Punjab, which includes a little over half of

Pakistan's population. In 1999–2004 the sampling frame included the entire country. We include only the 1999–2004 observations in our analysis.

—*South Africa* was surveyed in three waves; however, the first wave, 1989–93, overrepresents minority races, and blacks were sampled only in certain areas. Sampling weights were not provided. The next two waves, 1994–99 and 1999–2004, were designed to be representative of the population. We exclude observations from the first wave only.

Table B1 compares the main coefficient estimates from tables 1 through 3 with those we obtain by including all country-wave observations. The first panel shows the between-country analysis, with the first and third columns reproducing the results in the third column of table 1 for life sat-

Table B1. Influence of Nonrepresentative Samples in the World Values Survey^a

	<i>Life satisfaction</i>		<i>Happiness</i>	
	<i>Representative samples^b</i>	<i>All observations^c</i>	<i>Representative samples^b</i>	<i>All observations^c</i>
<i>Comparisons between countries, coefficient on log real GDP per capita^d</i>				
1981–84 wave	0.498* (0.252)	0.510** (0.230)	0.569** (0.230)	0.596*** (0.193)
1989–93 wave	0.558*** (0.096)	0.210*** (0.073)	0.708*** (0.123)	0.260** (0.098)
1994–99 wave	0.462*** (0.051)	0.323*** (0.069)	0.354*** (0.058)	0.214*** (0.069)
1999–2004 wave	0.346*** (0.046)	0.347*** (0.045)	0.126* (0.073)	0.0125* (0.072)
Combined, with wave fixed effects	0.398*** (0.040)	0.318*** (0.052)	0.244*** (0.063)	0.181*** (0.063)
<i>Comparisons within countries, coefficient on log household income, controlling for country fixed effects^e</i>				
1981–84 wave	0.199*** (0.022)	0.199*** (0.022)	0.281*** (0.023)	0.281*** (0.023)
1989–93 wave	0.153*** (0.011)	0.145*** (0.010)	0.188*** (0.013)	0.190*** (0.011)
1994–99 wave	0.243*** (0.013)	0.250*** (0.012)	0.209*** (0.013)	0.217*** (0.013)
1999–2004 wave	0.286*** (0.007)	0.274*** (0.007)	0.248*** (0.008)	0.245*** (0.008)
Combined, with country × wave fixed effects	0.249*** (0.007)	0.237*** (0.007)	0.234*** (0.008)	0.231*** (0.008)
<i>Comparisons between countries and over time, coefficient on log real GDP per capita^f</i>				
Levels	0.414*** (0.041)	0.339*** (0.053)	0.230*** (0.064)	0.181*** (0.061)
Levels with country fixed effects	0.301*** (0.091)	0.151 (0.136)	0.363*** (0.131)	0.283*** (0.128)

(continued)

Table B1. Influence of Nonrepresentative Samples in the World Values Survey^a (Continued)

	<i>Life satisfaction</i>		<i>Happiness</i>	
	<i>Representative samples^b</i>	<i>All observations^c</i>	<i>Representative samples^b</i>	<i>All observations^c</i>
Levels with country and wave fixed effects	0.552*** (0.118)	0.253* (0.192)	0.216 (0.187)	0.009 (0.170)
Short first differences	0.596*** (0.082)	0.407*** (0.122)	0.215 (0.136)	0.081 (0.122)
Long first differences	0.314*** (0.072)	0.172 (0.118)	0.114 (0.103)	-0.025 (0.117)

Source: Authors' regressions using World Values Surveys data and survey notes, 1981–2004.

a. Table reports results of regressions of well-being on the indicated income variable, both in the nationally representative World Values Survey samples analyzed in the main text and in the entire sample. Numbers in parentheses are robust standard errors. Asterisks indicate statistically significant from zero at the *10 percent, **5 percent, and ***1 percent level.

b. Results in the top, middle, and bottom panels are from tables 1, 2, and 3, respectively. Sample includes only nationally representative country-wave samples, yielding 234,093 life satisfaction respondents (and 228,159 happiness respondents) in 166 (165) country-waves, from 79 countries. Results in the middle panel also require household income data, reducing sample to 117,481 (or 117,299) respondents, 93 (94) country-waves, and 59 countries.

c. Sample includes all country-waves for which data were reported, including non-nationally representative samples described in this appendix. This broader sample yields an extra 25,582 satisfaction respondents (and 26,365 happiness respondents), from 17 (18) more country-waves, adding 1, 7, 8, and 2 countries to the satisfaction (and 1, 7, 8, and 2 to the happiness) samples, respectively, in waves 1–4, yielding 13 (14) more short first differences and 6 (7) more long first differences and resulting in a total sample of 82 countries. Results in the middle panel also require household income data, and so the broader sample yielded only 9,968 extra satisfaction respondents (and 9,946 extra happiness respondents), from 6 country-waves for a total of 61 countries.

d. National well-being index is regressed on log real GDP per capita. The well-being index is calculated in a previous ordered probit regression of well-being on country \times wave fixed effects. Standard errors are clustered by country. See the notes to table 1 for further details.

e. Results of ordered probit regressions of respondent-level well-being on log household income, controlling for country fixed effects, or country \times wave fixed effects where noted, as well as a quartic in age, gender, their interaction, and indicators for missing age or gender. See the notes to table 2 for further details.

f. National well-being index is regressed on log real GDP per capita, pooling together observations from all four waves of the survey. National well-being index is calculated in a previous ordered probit regression of well-being on country \times wave fixed effects. Standard errors are clustered by country. See the notes to table 3 for further details.

isfaction and happiness, respectively. Since the excluded observations typically represent a group with above-average income and education (and hence, likely higher happiness), our expectation is that incorporating these countries will yield lower estimates of the well-being-income gradient. These estimates are shown in the second column of table B1 for life satisfaction, and in the fourth column for happiness. The first and the last wave show little impact on the estimated coefficient, as the samples are largely the same (only Argentina was excluded from the first wave, and only Chile

and Egypt from the last). The 1989–93 and 1994–99 waves yield larger differences, as six countries were excluded from the former and eight from the latter. As expected, including these biased samples attenuates the estimated coefficients substantially, yet in all cases the estimated coefficient remains positive and statistically significant.

The second panel examines the impact of including the unrepresentative national samples on the within-country cross-sectional estimates. The first and third columns reproduce the coefficients from the second column of table 2. Despite the truncation of the poor in these samples, the fact that subjective well-being is linearly related to log income suggests that excluding a portion of the income distribution does not bias the coefficient estimates. Moreover, as we show in figure 10, most countries have a subjective well-being-income gradient of around 0.4, with little systematic variation. Hence one should expect little difference in the estimates that include observations from all of the country waves. Indeed, the estimated coefficients with the excluded samples, again shown in the second and fourth columns, are little different from those obtained without these countries.

Finally, the last panel reproduces the estimates shown in the second column of table 3, in which we analyze the World Values Survey as a country-wave panel dataset using the country aggregated (macro) data. The table 3 estimates are repeated in the first and third columns, and the comparison estimates including unrepresentative country-wave observations are shown in the second and fourth columns. The first row reports the simple bivariate well-being-GDP relationship and hence pools both within-country and between-country variation. These results are little affected by the inclusion of the unrepresentative country-wave observations. The estimates reported in the second row include country fixed effects and therefore isolate the within-country time-series variation. The inclusion of countries whose sample becomes more representative as GDP grows (second and fourth columns) reduces the estimated coefficient. The third row adds controls for each wave of the World Values Survey in addition to the country controls. Again, the inclusion of the non-representative samples reduces the estimated coefficients. Finally, the last two columns consider both short first differences, that is, those between consecutive country-wave observations, and long first differences, which are those between the first and the last observation for each country. Excluding differences involving countries where the survey frame changed yields robust estimates of a positive relationship between life satisfaction

and income and happiness and income over time. Not surprisingly, including countries whose samples are becoming increasingly representative of the poor over time decreases these estimates substantially. Including the noncomparable intertemporal variation in well-being also yields less precise estimates. Even when these countries are included, the results are still roughly consistent with the null hypothesis that the time-series well-being-income gradient is close to the 0.4 range obtained from our between-country and within-country analyses.

References

- Blanchflower, David, and Andrew Oswald. 2004. "Well-Being over Time in Britain and the USA." *Journal of Public Economics* 88, no. 7–8: 1359–86.
- Bradburn, Norman M. 1969. *The Structure of Psychological Well-Being*. Chicago: Aldine.
- Buchanan, William, and Hadley Cantril. 1953. *How Nations See Each Other: A Study in Public Opinion*. University of Illinois Press.
- Campbell, John Y., and N. Gregory Mankiw. 1990. "Permanent Income, Current Income, and Consumption." *Journal of Business and Economic Statistics* 8, no. 3: 265–79.
- Cantril, Hadley. 1951. *Public Opinion, 1935–1946*. Princeton University Press.
- . 1965. *The Pattern of Human Concerns*. Rutgers University Press.
- Clark, Andrew E., Paul Frijters, and Michael A. Shields. 2008. "Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature* 46, no. 1: 95–144.
- Deaton, Angus. 2008. "Income, Health and Well-Being around the World: Evidence from the Gallup World Poll." *Journal of Economic Perspectives* 22, no. 2: 53–72.
- DeNavas-Walt, Carmen, Bernadette D. Proctor, and Robert J. Mills. 2006. *Income, Poverty and Health Insurance Coverage in the United States: 2005*. Current Population Reports. Washington: U.S. Census Bureau.
- Di Tella, Rafael, Robert J. MacCulloch, and Andrew J. Oswald. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics* 85, no. 4: 809–27.
- Diener, Ed. 1984. "Subjective Well-Being." *Psychological Bulletin* 95, no. 3: 542–75.
- . 2006. "Guidelines for National Indicators of Subjective Well-Being and Ill-Being." *Journal of Happiness Studies* 7, no. 4: 397–404.
- Diener, Ed, and Martin E. P. Seligman. 2004. "Beyond Money: Toward an Economy of Well-Being." *Psychological Science in the Public Interest* 5, no. 1: 1–31.
- Diener, Ed, and William Tov. 2007. "Culture and Subjective Well-Being." In *Handbook of Cultural Psychology*, edited by Shinobu Kitayama and Dov Cohen. New York: Guilford.
- Diener, Ed, Richard E. Lucas, and Christie Napa Scollon. 2006. "Beyond the Hedonic Treadmill: Revising the Adaptation Theory of Well-Being." *American Psychologist* 61, no. 4: 305–14.
- Easterlin, Richard A. 1973. "Does Money Buy Happiness?" *The Public Interest* 30: 3–10.
- . 1974. "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." In *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*, edited by Paul A. David and Melvin W. Reder. Academic Press.
- . 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27, no. 1: 35–48.

- . 2001. "Income and Happiness: Towards a Unified Theory." *Economic Journal* 111, no. 473: 465–84.
- . 2005a. "Feeding the Illusion of Growth and Happiness: A Reply to Hagerty and Veenhoven." *Social Indicators Research* 74, no. 3: 429–43.
- . 2005b. "Diminishing Marginal Utility of Income? Caveat Emptor." *Social Indicators Research* 70, no. 3: 243–55.
- Eid, Michael, and Ed Diener. 2004. "Global Judgments of Subjective Well-Being: Situational Variability and Long-Term Stability." *Social Indicators Research* 65, no. 3: 245–77.
- Ekman, Paul, and Wallace V. Friesen. 1971. "Constants across Cultures in the Face and Emotion." *Journal of Personality and Social Psychology* 17, no. 2: 124–29.
- Ekman, Paul, and others. 1987. "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion." *Journal of Personality and Social Psychology* 53, no. 4: 712–17.
- Frank, Robert H. 2005. "Does Absolute Income Matter?" In *Economics and Happiness: Framing the Analysis*, edited by Pier Luigi Porta and Luigino Bruni. Oxford University Press.
- Frey, Bruno S., and Alois Stutzer. 2002. "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature* 40, no. 2: 402–35.
- Graham, Carol. 2008. "Happiness and Health: Lessons—And Questions—For Public Policy." *Health Affairs* 27, no. 1: 72–87.
- Kahneman, Daniel, and Alan B. Krueger. 2006. "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives* 20, no. 1: 3–24.
- Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur A. Stone. 2006. "Would You Be Happier If You Were Richer? A Focusing Illusion." *Science* 312, no. 5782 (June): 1908–10.
- Kenny, Charles. 1999. "Does Growth Cause Happiness, or Does Happiness Cause Growth?" *Kyklos* 52, no. 1: 3–26.
- Layard, Richard. 1980. "Human Satisfaction and Public Policy." *Economic Journal* 90, no. 363: 737–50.
- . 2003. "Happiness: Has Social Science a Clue." Lionel Robbins Memorial Lectures 2002/3, London School of Economics, March 3–5. cep.lse.ac.uk/events/lectures/layard/RL030303.pdf.
- . 2005a. *Happiness: Lessons from a New Science*. London: Penguin.
- . 2005b. "Rethinking Public Economics: The Implications of Rivalry and Habit." In *Economics and Happiness: Framing the Analysis*, edited by Pier Luigi Porta and Luigino Bruni. Oxford University Press.
- Leigh, Andrew, and Justin Wolfers. 2006. "Happiness and the Human Development Index: Australia is Not a Paradox." *Australian Economic Review* 39, no. 2: 176–84.
- Lleras-Muney, Adriana. 2005. "The Relationship between Education and Adult Mortality in the United States." *Review of Economic Studies* 72, no. 1: 189–221.

- Luttmer, Erzo F. P. 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics* 120, no. 3: 963–1002.
- Maddison, Angus. 2007. "Historical Statistics for the World Economy: 1–2003 AD." www.gdc.net/maddison/Historical_Statistics/horizontal_file_03-2007.xls.
- Oswald, Andrew J. Forthcoming. "On the Curvature of the Reporting Function from Objective Reality to Subjective Feelings." *Economics Letters*.
- Rivers, Douglas, and Quang H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39, no. 3: 347–66.
- Smith, Tom W. 1979. "Happiness: Time Trends, Seasonal Variations, Intersurvey Differences, and Other Mysteries." *Social Psychology Quarterly* 42, no. 1: 18–30.
- . 1986. *Unhappiness on the 1985 GSS: Confounding Change and Context*. GSS Technical Report no. 34. Chicago: NORC.
- . 1988. *Timely Artifacts: A Review of Measurement Variation in the 1972–1988 GSS*. GSS Methodological Report no. 56. Chicago: NORC.
- Stevenson, Betsey, and Justin Wolfers. 2007. "The Paradox of Declining Female Happiness." University of Pennsylvania.
- . 2008. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox." Working Paper 14282. Cambridge, Mass.: National Bureau of Economic Research.
- . Forthcoming. "Happiness Inequality in the United States." *Journal of Legal Studies*.
- Strunk, Mildred. 1950. "The Quarter's Polls." *Public Opinion Quarterly* 14, no. 1: 174–92.
- van Praag, Bernard M. S., and Ada Ferrer-i-Carbonell. 2004. *Happiness Quantified. A Satisfaction Calculus Approach*. Oxford University Press.
- Veenhoven, Ruut. 1991. "Is Happiness Relative?" *Social Indicators Research* 24, no. 1: 1–34.
- . 1993. *Happiness in Nations: Subjective Appreciation of Life in 56 Nations, 1946–1992*. Rotterdam: Erasmus University.
- . Undated. *World Database of Happiness, Trends in Nations*. Rotterdam: Erasmus University. worlddatabaseofhappiness.eur.nl/trendnat/framepage.htm.
- Wolfers, Justin. 2003. "Is Business Cycle Volatility Costly? Evidence from Surveys of Subjective Well-Being." *International Finance* 6, no. 1: 1–26.

Comments and Discussion

COMMENT BY

GARY S. BECKER AND LUIS RAYO In this paper Betsey Stevenson and Justin Wolfers provide convincing evidence that self-reported happiness and measures of life satisfaction rise with income, not only at a moment in time within a country, but also between poorer and richer countries. The income coefficients are in fact quite close for the between- and within-country comparisons, contrary to findings in the literature with more-limited datasets. Depending on the definition of the relevant peer comparison group, these results can mean either that peer comparisons are weaker than previously believed (for example, if the peer comparison group is restricted to an individual's country of residence), or simply that peer comparisons are of similar intensity between and within countries. On the other hand, the impact of changes in income over time is less clear, since the data give a mixed picture and current tests lack statistical power to allow a firm conclusion either way.

What do these results mean for the precise role of peer comparisons and habits in the determination of reported happiness and life satisfaction and, importantly, for the relationship between utility and these measures of well-being? We concentrate our discussion on the latter question, which has received far less attention in the literature. Our conclusion is that although reported happiness and life satisfaction may be related to utility, they are no more measures of utility than are other dimensions of well-being, such as health or consumption of material goods.

We had very helpful discussions with Kevin M. Murphy. We also thank Miguel Diaz for valuable research assistance and gratefully acknowledge The University of Chicago Graduate School of Business for financial support.

To pretty much everyone who has written about happiness, an increase in income raises utility only if it also raises happiness. Similarly, various authors have interpreted the evidence from earlier studies that reported happiness did not rise over time when incomes rose as indicating that the utility of the typical person did not increase. On the basis of this interpretation, some authors have argued that governments should intervene in ways that would effectively discourage income growth.¹

A recent survey of the happiness literature addresses the question of whether happiness refers to utility, subject to reporting error, and mentions various types of possible evidence.² The answers of friends of person A about the happiness of A generally are consistent with A's answers about his own happiness, and his own answers are also consistent over time. Neurological measures, such as asymmetries in prefrontal activity, and physiological measures, such as the stress response, are correlated with the answers to happiness surveys. Some studies have also found that people behave in a manner consistent with their reports, such as trying to avoid situations (for example, unemployment) that reduce their self-reported happiness. Although Stevenson and Wolfers are not explicit about this point, they also seem to view happiness and life satisfaction as noisy measures of utility.

The above evidence indicates that self-reported happiness is indeed meaningful, but it does not imply that it is the same as utility. For example, if someone reports that she is healthy or owns an expensive SUV, her friends are likely to confirm these facts. Moreover, if she continues to remain healthy and to own this car over time, future reports are also likely to confirm these facts. One can also readily find behavior consistent with seeking to improve one's health or to acquire an SUV. However, none of this evidence, no matter how precise, implies that health and SUV ownership can be equated with utility.

These examples suggest an alternative interpretation of the happiness data, namely, that happiness is a commodity in the utility function in the same way that owning a car and being healthy are. Consumption of particular commodities may be correlated with utility, but greater consumption of a commodity is not the same as greater utility. Indeed, following a

1. See, for example, the recommendations in Robert Frank, *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess* (Free Press, 1999), and Richard Layard, *Happiness: Lessons from a New Science* (Penguin Press, 2005).

2. See Andrew E. Clark, Paul Frijters, and Michael A. Shields, "Income and Happiness: Evidence, Explanations, and Economic Implications," PSE Working Paper 2006-24 (Paris: Ecole Normale Supérieure, 2006).

change in consumption opportunities, the consumption even of important commodities may sometimes fall while utility increases. On this interpretation, happiness may be an important determinant of utility, but that does not mean that both coincide.

When one considers the evolutionary origin of human beings and views people's choices in light of their adaptive significance during ancestral times, it should not be surprising that utility has the potential to differ from happiness. A well-accepted proposition of evolutionary psychology is that humans did not evolve to be happy; rather, humans evolved to execute choices that would have favored the multiplication of their genes during ancestral times. Happiness is merely an instrument implicitly used by our genes toward that end.³ In fact, happiness may well be the *primary* instrument guiding our conscious decisions,⁴ but it is hardly the *only* instrument.

The philosopher Blaise Pascal claimed that "All men seek happiness. This is without exception. Whatever different means they employ, they all tend to this end. . . . This is the motive of every action of every man. . . ."⁵ Perhaps Pascal was the victim of a well-known feature of the human brain: the fact that we commonly lack introspective access to the sources of our own behavior.⁶

From a Darwinian perspective, since happiness has limited, if any, direct fitness value, it can easily lose its priority when the right opportunity arises. For instance, if we were to metaphorically "ask" our genes whether we should accept an unpleasant but high-paying job that would increase our social standing, with the only drawback of making us less happy, they would not hesitate for a moment. Thus it is not implausible that we can potentially follow courses of action that would reduce our happiness in exchange for evolutionarily salient goals.

According to the theory of revealed preference, if a person takes the unpleasant but high-status job because it pays well, we would say that he

3. See, for example, Steven Pinker, *How the Mind Works* (Norton, 1997).

4. As modeled in Shane Frederick and George Loewenstein, "Hedonic Adaptation," in *Well-Being: The Foundations of Hedonic Psychology*, edited by Daniel Kahneman, Ed Diener, and Norbert Schwarz (New York: Russell Sage Foundation, 1999); Arthur J. Robson, "The Biological Basis of Economic Behavior," *Journal of Economic Literature*, 39, no. 1 (2001): 11–33; and Luis Rayo and Gary S. Becker, "Evolutionary Efficiency and Happiness," *Journal of Political Economy* 115, no. 2 (2007): 302–37.

5. Cited in Daniel Gilbert, *Stumbling on Happiness* (Knopf, 2006), p. 15.

6. Colin Camerer, George Loewenstein, and Drazen Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature* 43 (March 2005): 9–64; for a discussion of mechanisms other than feelings that guide behavior, see Paul M. Romer, "Thinking and Feeling," *American Economic Review Papers and Proceedings* 90, no. 2 (2000): 439–43.

raises his utility by taking this job, yet it could very well lead him to respond that he is less happy. This example suggests that there are ways to use the theory of consumer utility maximization to test whether happiness data in fact measure utility, or whether instead happiness is a commodity in the utility function. We develop a simple formal analysis to bring out more precisely how to make such tests and show why the distinction between commodity and utility can be important.

Before we do, let us note that some authors recognize that happiness may not always correspond to utility, but they take this to mean that individuals make mistakes in maximizing their utility, and that absent these mistakes, happiness and utility would correspond more or less completely.⁷ Our approach, in contrast, takes the standard position that individual choices do maximize utility, and rather than appeal to mistakes to explain why happiness and utility do not always correspond, we explicitly allow for arguments other than happiness in the utility function. It is also worth noting that we do not take a normative position on what individuals should or should not be maximizing.

To develop a more formal analysis, consider a utility function

$$(1) \quad U = U(Z, H),$$

where H is happiness and Z is another commodity. For simplicity, we consider a static environment. We assume that

$$(2) \quad \partial U / \partial H > 0, \text{ and } \partial U / \partial Z > 0.$$

The usual identification of reported happiness with utility assumes

$$(3) \quad U = U(H), \text{ so that } dU/dH > 0.$$

That equation 2 uses the partial derivative and equation 3 the total derivative of U with respect to H is crucial: in equation 2, Z is held constant when happiness changes, whereas in equation 3 nothing is held fixed because nothing else affects utility. Equation 2 implies, for example, that an increase in happiness over time could be associated with lower utility if Z decreases enough, and that an individual who reports less happiness could have higher utility than another individual if the individual with lower H had sufficiently higher Z .

7. For example, Bruno Frey and Alois Stutzer, "Economic Consequences of Mispredicting Utility," IEW Working Paper 218, University of Zurich and University of Basel, 2004.

Obviously, one cannot directly buy happiness in the marketplace. So we assume that both H and Z are not directly purchased but have to be produced by each individual according to household production functions, using market goods, time, and other inputs. These production functions are

$$(4) \quad H = F(x, h_h, E), \text{ and } Z = G(y, h_z, E),$$

where x and y refer to inputs of various goods, the h 's are household time inputs, and E refers to environmental variables. These environmental inputs include the education of the individual, shocks to his or her health, the H or Z of other individuals (to allow for social interactions), and command over technology that affects production of H and Z . For example, an award for achievement, or a better job, or declines in the consumption or happiness of others, might raise H .

Budget constraints are the third building block of the analysis. The goods constraint is

$$(5) \quad p_x x + p_y y = wl + R = I,$$

where the p 's are market prices, w is the wage rate, l is hours worked ($l = 1 - h_h - h_z$), and R is nonwage income. This equation can be manipulated to give the "full-income" budget constraint

$$(6) \quad \pi_z Z + \pi_h H = w + R = S,$$

where the π 's are average shadow prices of producing H and Z , and S is full income that is independent of the allocation of time between the market and household sectors. These shadow prices depend on the prices of the goods inputs (the p 's), the wage rate (w), and the productivity of household production, which depends on the various individual-specific variables (E). This analysis of household production indicates that the production of happiness has important personal components as well as objective market components, such as income and success.

We assume that individuals maximize their utility, subject to their budget constraints and household production functions. If the utility function is that given by equation 1, the resulting Hicks demand function for H is

$$(7) \quad H = H(U, \pi_z, \pi_h) = H(U, p_x, p_y, w, E).$$

An increase in H would necessarily correspond to an increase in U only if the π 's, or the p 's, w , and E , are held constant. Also note that a rise in the individual's nonwage income R (with no change in prices) increases U , and this rise increases H as well if happiness is a normal good.

In contrast, if the utility function is given by equation 3, the Hicks demand function for H is simply the inverse of the utility function in that equation:

$$(8) \quad H = U^{-1}(U).$$

In this case, which is the usual one in the happiness literature, there is a one-to-one correspondence between happiness and utility.

Fortunately, empirical tests can help distinguish whether the Hicks demand function in equation 7 or that in equation 8 is more relevant for interpreting the happiness data. These tests are fundamentally different from the typical tests in the happiness literature that try to determine whether reported happiness can be taken as an accurate measure of “hedonic well-being,” but make no reference to choices.

To consider a test that uses data on income, suppose a person’s income increased because her hourly wage rate increased as a result of exogenous factors. If her nonwage income and other prices did not change, equation 8 predicts that her happiness would rise. However, if at the same time as her wage increased, R were reduced so that she could just continue to buy the initial level of leisure and goods, then consumer theory predicts that her utility either would be unaffected (for small wage changes) or would increase. This means that if equation 8 describes utility, her degree of happiness should also be unaffected (for small wage increases) or increase (for larger changes) but should not decrease.

Similarly, voluntary migration from a poorer to a richer country would raise utility even if the immigrant’s relative income in the country that he moved to would be lower than in the country he moved from. Therefore equation 8 would predict an increase in the reported happiness of immigrants.

The prediction of equation 7 about the effect of these changes on happiness is more open-ended, since it depends on the household production function for happiness. For example, if the production of happiness is highly intensive in the individual’s own time, which seems reasonable to us, then $\partial H/\partial w < 0$ in equation 7 (in which utility is held constant). That is, under these conditions a compensated increase in the wage rate would reduce happiness, since the production of happiness is assumed to be time intensive relative to the production of other commodities. In fact, the demand function in equation 7 allows for the possibility that an increase in the wage rate reduces happiness even if the individual also experiences higher utility.

With regard to the migration prediction, if happiness is sensitive to the sizable adjustment costs of a major move, then reported happiness could

very well go down for a while as utility went up. Eventually, happiness might also increase after an adjustment is made to the new situation.

Bruno Frey and Alois Stutzer argue that individuals sometimes fail to maximize their own happiness, such as when seeking higher income at the expense of longer commuting times.⁸ The interpretation offered by these authors is that individuals in fact seek to maximize happiness but systematically underestimate the negative effects of commuting, while overestimating the value of enjoying higher wages. In contrast, under the framework we have proposed, we would interpret their results as confirming that happiness is not all that individuals appear to care about.

The literature on self-reported health illustrates the type of approach to happiness that we are advocating. The papers that try to explain answers to questions about a person's quality of health do not assume that these answers report utility, although they recognize that health is important to utility. Rather, they use these health reports to construct household production functions for health, and demand functions for health that depend on various prices and individual-specific characteristics.⁹ If answers about health are treated only as an input into utility, albeit an important input, why should answers about happiness be treated any differently?

There are several important examples in the literature where consumption of a good varies with respect to income at a moment in time very differently than it varies with income as income changes over time. For example, it was at first considered rather paradoxical that the share of income saved rises strongly with income when comparing individuals at a moment in time, whereas the average share of income saved hardly changed as average income increased over time. It was later found that this paradox could be resolved either by the permanent income hypothesis,¹⁰ or by the assumption that saving rates depend on one's income relative to peers.¹¹

Similarly, in early studies on the labor force participation of married women, the propensity of wives to participate in the labor force tended, in cross-sectional comparisons, to be lower when their husbands' income was higher. Yet over time the participation of wives tended to rise as the aver-

8. Frey and Stutzer, "Economic Consequences of Mispredicting Utility."

9. See, for example, Donna B. Gilleskie and Amy L. Harrison, "The Effect of Endogenous Health Inputs on the Relationship between Health and Education," *Economics of Education Review* 17, no. 3 (1998): 279–95.

10. Milton Friedman, *A Theory of the Consumption Function: A Study by the National Bureau of Economic Research, New York* (Princeton University Press, 1957).

11. James S. Duesenberry, *Income, Saving and the Theory of Consumer Behavior* (Harvard University Press, 1949).

age income of their husbands increased. Here the reconciliation proposed was that the hourly earnings of women rose over time, too, and that higher wages induced more married women to enter the labor force.¹²

Apropos of the question that motivates the reexamination of the happiness-income relation in the present study, if happiness refers to utility, it is somewhat devastating to the usual welfare implications of economic growth if average happiness does not increase as average incomes rise over time, or if happiness is not greater on average in richer than in poorer countries. However, if happiness is just one (important) commodity in the utility function, then the absence of a positive connection between happiness and income over time or across nations is consistent with an increase in utility when income is higher.

Happiness as a commodity may not increase when utility does, perhaps because happiness is very habitual, even addictive, or because a person's happiness may be affected by the achievements and happiness of others. Or, going back to the discussion of the Hicks demand function for happiness in equation 7, perhaps happiness is time intensive, so that a rise in wage rates over time raises both income and the relative cost of producing happiness.

That the authors do find a positive connection between income and happiness within and between countries, and possibly over time within a country, is important and reasonable, but it does not speak to the question of whether happiness is identical with utility. The next step is to integrate their findings into a more comprehensive theory of utility maximization that can discern the precise role played by happiness in people's decisions.

COMMENT BY

ALAN B. KRUEGER As of this morning, there were 1,790 references to the Easterlin paradox according to Google Scholar. It is therefore a real challenge to add something new in this field. Yet Betsey Stevenson and Justin Wolfers have succeeded in raising doubts about the validity of the Easterlin paradox. This is an achievement. The real contribution of this paper, it seems to me, is in precisely defining and estimating the relationship between subjective well-being and income, using cross-sectional data within countries, cross-sectional data across countries, and (to a lesser extent) data on within-country changes over time. My colleague

12. See, for example, Jacob Mincer, "Labor Force Participation of Married Women: A Study of Labor Supply," in *Aspects of Labor Economics: A Conference of the Universities—National Bureau Committee for Economic Research*, edited by H. G. Lewis (Princeton University Press, 1962).

Angus Deaton has recently shown that life satisfaction and the “ladder of life” measures of well-being move with the logarithm of GDP per capita across countries.¹ By making the Easterlin hypothesis a precise statement about gradients that can be compared statistically across samples, estimation techniques, and sources of variation, Stevenson and Wolfers have provided a valuable service.

Together with Deaton’s work, the findings in this paper pose a strong challenge to the Easterlin paradox that should not be ignored. Indeed, perhaps Easterlin’s claim of no connection between increases in income and increases in happiness should now be called the “Easterlin hypothesis” or the “Easterlin conjecture.” I’m not ready to call it a nonparadox just yet, however. After reading the paper and reanalyzing the data, I think the available evidence supports a positive relationship between various measures of subjective well-being and log income, although the jury is still out in one important respect: As I explain below, the time-series evidence strikes me as providing an indecisive test one way or the other.

Before I turn to the evidence in the paper, it is worth considering some biases (or “tendencies”) in psychological measures of subjective well-being. People do not think about their life satisfaction or level of happiness in the same way they think about their mailing address or years of schooling. When asked, they construct an answer on the spot. They often use rules of thumb for providing their answers. They are also affected by their current mood and thoughts. In an ingenious experiment to demonstrate the importance of transitory mood on reported life satisfaction, Norbert Schwarz invited subjects to fill out a satisfaction questionnaire.² Before answering the questionnaire, however, he asked them to make a photocopy of the questionnaire. For half of the subjects, a dime was planted on the copy machine. Reported life satisfaction was a point higher for those who encountered a dime! Clearly, their mood improved by finding the small amount of money, leading them to report higher satisfaction with their lives over all.

The topic on a person’s mind at the time of answering a life satisfaction or happiness question also affects his or her response. Sometimes people’s predicament naturally suggests a topic to focus on. For example, people in Minnesota report that they believe people in California are happier than they themselves are, because they naturally focus on the weather when

1. Angus Deaton, “Income, Aging, Health and Wellbeing around the World: Evidence from the Gallup World Poll,” Working Paper 13317 (Cambridge, Mass.: National Bureau of Economic Research, August 2007).

2. Norbert Schwarz, *Stimmung als Information: Untersuchungen zum Einfluß von Stimmungen auf die Bewertung des eigenen Lebens* (Heidelberg: Springer Verlag, 1987).

thinking of the well-being of Californians, ignoring the smog, congestion, and daily hassles of life that overwhelm the effect of the weather on satisfaction. On average, Minnesotans are no less happy than Californians when they report their subjective well-being in surveys like the General Social Survey.³ This tendency is often called *focusing illusion*. In a paper I co-wrote with Daniel Kahneman and coauthors, we reported a strong focusing illusion for income.⁴ Specifically, we asked, “Think of someone who makes less than \$20,000 a year, or someone who makes more than \$100,000 a year. How much time do you think they spend in a bad mood?” Respondents on average predicted that the lower-income group would spend 58 percent of their time in a bad mood, whereas other data we collected suggested that people in this income group spent, on average, only 32 percent of their time in a bad mood. Respondents also predicted that the higher-income people would spend 26 percent of their time in a bad mood, whereas our evidence suggested they actually spent 20 percent of their time in a bad mood. Thus, the effect of income on mood was over-predicted by 20 percentage points. In general, people tend to exaggerate the effect of circumstances such as income, fringe benefits, and marriage when they are asked how circumstances affect well-being.

When people are asked about their own life satisfaction or happiness, they may reflect on their economic conditions and partly use that as a handle on providing an answer. The mental exercise that well-off respondents go through is probably something like, “I’m a fortunate person. I have a high-paying job. I live in a big house and I have an expensive car. I should report myself as satisfied with my life. If I don’t, I’m not a very responsible person.” This tendency is less likely to affect people’s moment-to-moment mood or affect. It is probably more than a coincidence that the measures Stevenson and Wolfers examine that are more closely related to how people felt yesterday (that is, their affect), as opposed to measures that reflect an evaluative judgment of how they feel about their lives over all, tend to be less related to income. Measures of well-being that are closer to actual feelings are probably less prone to bias from a focusing illusion.

This leads to a potential concern with some of the international data used in the paper. When respondents are asked to place themselves on a

3. See David Schkade and Daniel Kahneman, “Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction,” *Psychological Science* 9, no. 5 (1998): 340–46.

4. Daniel Kahneman, Alan Krueger, David Schkade, Norbert Schwarz, and Arthur Stone, “Would You Be Happier If You Were Richer? A Focusing Illusion,” *Science* 312 (June 30, 2006): 1908–10.

ladder of the best possible life in a survey that is represented as a world poll, they may be more prone to a focusing illusion that goes something like, "I live in a rich country with many amenities. I should place myself high on the ladder of life, regardless of how I feel in my own life moment to moment." A reasonable concern is that the focusing illusion causes the ladder of life to exaggerate the effects of national economic development on people's self-reported step on the ladder.

Turning to the paper's econometric estimates, I regard as a clear advance the authors' approach of specifying the Easterlin paradox as the nonequivalence between time-series and cross-sectional estimates of the happiness–log income gradient. For far too long, economists have made casual, informal comparisons in this area. The paradox should be rephrased as a flatter (or nonexistent) happiness-income gradient when happiness is related to income using national time-series variation instead of cross-sectional individual or country-level variation. There is some slippage, however, in the way the analysis is implemented because log GDP is an imperfect substitute for the mean of log income. If income is log-normally distributed and labor's share is constant across countries, then the standard deviation of log income belongs in the aggregate equation. I appreciate that there are difficulties with the micro income data, and the authors' reluctance to aggregate the micro data is defensible, but the use of different income concepts and measures is unfortunate.

That technical quibble aside, the authors' table 3 seems to me to provide the most relevant new evidence on the Easterlin paradox. This evidence strikes me, however, as more ambiguous than the description "remarkably robust" implies. The estimates from the long differences from the World Values Survey are imprecise. The estimates of the effect of GDP on happiness are noisy and insignificant in both the long and the short first-differences specifications. It is also puzzling that the evidence of an effect of income on satisfaction is notably weaker when the differences are taken over a longer period of time; one would expect there to be greater signal in GDP changes over longer periods. These results seem to me to be consistent with the psychological phenomenon of adaptation or habit formation: over time, people adapt to their circumstances and return to a steady-state level of satisfaction. Nevertheless, the results suggest the difficulty of testing the Easterlin paradox with available happiness data: despite the pooling of time-series data from more than fifty countries, the estimates are not sufficiently precise to rule out effects of roughly the same magnitude found within countries or between countries, or no effect.

I am reluctant to conclude that the Easterlin paradox is just due to noisy data, however. The reason is that the evidence supports the existence of heterogeneous happiness-income gradients across countries. In the United States and China, for example, GDP growth is apparently unrelated to increases in satisfaction, whereas in Japan the new evidence suggests a positive correlation. There is an apparent paradox: why do some countries do a much better job translating income gains into happiness than others?

Readers should also be warned that when there are heterogeneous treatment effects, regression models that constrain the countries to have a homogeneous effect can yield misleading results. To explore this issue further, I estimated the following equation with the World Values Survey data:

$$Y_{jt} = \alpha_j + \delta_t + \beta_j \log(GDP_{jt}),$$

where Y_{jt} is satisfaction in country j in year t .⁵ Notice that this equation includes country fixed effects (α_j) and allows for unrestricted country-specific GDP per capita gradients (β_j). (By comparison, the models in the authors' table 3 constrain β_j to equal a constant β .) I worked with the short first differences and limited the sample to countries with at least two observations.⁶ When I estimate a model with a homogeneous effect (no j subscript on β), I find a coefficient of 0.25 and a standard error of 0.10, which is inconsistent with the Easterlin paradox. But an F-test reveals that this model is overly restrictive: country-log GDP interactions are highly statistically significant. The data prefer separate satisfaction-GDP gradients for each country. When instead I estimate a separate β_j for each country and then take the arithmetic average of those coefficients, the effect of GDP on satisfaction for the average country is -0.14 , with a standard error of 0.10.⁷ Thus, for the average country, GDP growth is not associated with improvements in satisfaction. I do not want to push this finding very far, because I put more faith in the longer differences, and there are not enough time periods to estimate interactive models with the longer differences. Nonetheless, the country-level differences in the satisfaction-GDP gradient, which strike me as more than a statistical artifact, represent a real puzzle that deserves further research.

5. Justin Wolfers kindly provided me with the data used for table 3 in the conference draft of their paper to estimate this less restrictive equation.

6. In addition, my sample differs from the one used in the published version of Stevenson and Wolfers' paper because their sample has evolved over time.

7. One reason why the standard error does not increase is that the country-GDP interactions fit the data better, leading to a lower mean square error.

I have long felt that it is naïve to view self-reported subjective well-being as measuring utility, at least as that concept is conceived in economics. At best, subjective well-being captures a component of utility. That said, life satisfaction, happiness, and the other measures of subjective well-being studied in this paper seem to behave more like one would expect utility to behave, at least as far as cross-sectional differences in income are concerned. So perhaps the findings in this paper should support more research by economists using subjective well-being as an outcome measure. One priority, it would seem to me, is to explain why some countries do a better job of converting income gains into higher subjective well-being than others.

To conclude, I think Stevenson and Wolfers have moved the literature forward by precisely specifying and testing the Easterlin paradox. Their conclusion (and Deaton's) that income has a logarithmic effect on subjective well-being will be hard to reject with available data for some time, although a more definitive test will await consistent panel data for many countries spanning many years. A final issue involves policy. Stevenson and Wolfers motivate their analysis by pointing to some of the more extreme policy proposals that have been justified by reference to the Easterlin paradox. Many of these strike me as farfetched, since income growth has been shown to be associated with improvements in longevity, health, education, knowledge, and other beneficial outcomes for individuals and society, regardless of any effect on subjective well-being. But the paper does not return to the policy implications of a logarithmic effect of GDP on subjective well-being with a semi-elasticity of around 0.2 to 0.4. These estimates imply a rather rapidly diminishing marginal utility of income for developed countries, and perhaps add to the justifications for policies such as progressive income taxation. I suspect most people would be surprised by the rather small implied effect of increases in income on their psychological well-being. The focusing illusion should ensure that this topic will continue to attract attention for years to come.

GENERAL DISCUSSION Robert Gordon noted the paradox of real wages stagnating at the same time that Americans are enjoying increased material amenities. Median real wages over the last forty years have barely increased, yet Americans have accumulated an enormous quantity of goods: the typical house is larger, many families own multiple cars, and approximately two-thirds of houses have air conditioning. He suggested that economists have not adequately measured the value of the many new

inventions over the last thirty years that people now take for granted, and that this may create a wedge between real income and happiness.

Gordon also commented on the stunning differences in the level and growth of longevity between the top and the bottom 10 percent of the income distribution. He wondered how health and happiness interplay in the utility function. Edward Glaeser added that happiness may be considered similar to mental health, which, like physical health, can enter as an argument in a utility function. He suggested that it should be possible to use happiness data to uncover market failures, in much the same way that health data are used to measure whether or not particulates in the air make children sick. Perhaps there is some negative externality associated with getting rich that manifests itself as a negative social multiplier as one moves from individual to aggregate data.

Robert Barro wondered how happiness differs conceptually from utility, and what new knowledge or understanding is gained from examining happiness data. The paper's findings that happiness and income levels are correlated make sense, unlike much of the previous literature. However, if the results indicated otherwise, he would conclude that either the data or the methods were flawed in some way, not that there is no relationship between happiness and income.

Christopher Carroll proposed that the results in this paper could be reconciled with the Easterlin paradox by invoking habit formation. He noted that the relationship between happiness and short-term GDP growth appears stronger than that between happiness and long-term GDP growth, and that happiness is strongly correlated with the output gap in the United States, which seems to suggest that happiness is affected by economic growth (or lack thereof) that deviates from expectations or habits. Gary Becker added that a habit model can produce a much weaker effect of income growth over time than it can cross-sectionally: given people's habits, those who experience an increase in income will report greater happiness in a given period—hence the cross-sectional effect; over time, this new income level is built into habits. He added that this model has very different welfare implications than a model of interpersonal income comparisons.

Benjamin Friedman argued that the Easterlin paradox can partly be explained by the fact that there are clearly diminishing returns to higher incomes in terms of measurable benefits such as longevity and literacy. He noted that the Easterlin paradox had been falsified in international data years ago, for example in the Eurobarometer data from the late 1970s and early 1980s.

Daron Acemoglu observed that an Easterlin paradox should necessarily exist in these data, given that the only benchmark a respondent can use at a given point in time is the happiness of other people. If happiness is scored on a scale from 1 to 10, and happiness increases with income, everyone with growing income should eventually report happiness levels of 10, which is then meaningless. He found it very impressive that the authors were able to find any correlation between happiness and income growth over time, given this limitation.

Carol Graham commented on how the framing of survey questions affects results across countries. For example, if asked, “What is the best possible life?” a Kenyan might respond that it is life in the United States rather than in Kenya, and this could yield a relationship between happiness and economic growth in the data. However, when questions about happiness are noncomparative, a different pattern may emerge. Nigerians report unusually high happiness levels, for example. So it is worth examining how questions are phrased and against whom the respondents are comparing themselves when answering. Graham was not surprised that rapid-growth outliers had a large effect on the results in the paper. Rapid growth can be unsettling, and it creates much inequality. Such an environment is not comparable with that in countries with stable growth.

Andrei Illarionov added that cross-cultural differences in attitudes about happiness and the words used to describe it complicate international comparisons of the income-happiness relationship. Differences in political systems can also complicate these comparisons; for example, people living in a dictatorship might report greater happiness than those living in a liberal democracy, but the former should not necessarily be taken at face value.