

The financial crisis and the well-being of Americans

2011 OEP Hicks Lecture*

By Angus Deaton

Center for Health and Well-being, Princeton University;
e-mail: deaton@princeton.edu

I use daily data on self-reported well-being (SWB) to examine how the Great Recession affected the emotional and evaluative lives of the population. In the fall of 2008 and lasting into the spring of 2009, at the bottom of the stock market, Americans reported sharp declines in their life evaluation, sharp increases in worry and stress, and declines in positive affect. By the end of 2010, in spite of continuing high unemployment, these measures had largely recovered. The SWB measures do a better job of monitoring short-run levels of anxiety than the medium-term evolution of the economy. Even very large macroeconomic shocks will cause small and hard to detect effects on SWB. Life evaluation questions are extremely sensitive to question order effects—asking political questions first reduces reported life evaluation by an amount that dwarfs the effects of even the worst of the crisis.

1. Introduction

The financial crisis that began in the summer of 2008 saw a rise in the unemployment rate from 4.8% in April 2008 to 10.6% at its peak in January 2010, a 4.4% drop in employee compensation over five months in 2009–10, large stimulus-associated tax credits and rebates, a 4.7% drop of personal disposable income in May 2008 and 1.7% in May 2009, as well as a collapse and subsequent recovery of the stock market—the S&P 500 Index on 6th March 2009 had fallen to 40% of its all time high of October 2007, and then more than doubled again by end 2010. Through the fall in the market and the fall in the prices of housing and of other assets, 60% of households saw their wealth decline between 2007 and 2009, and 25% lost more than half of their wealth, not including defined benefit pensions (Bricker *et al.*, 2011); these declines were widespread, affecting large shares of households across all age, income, and education groups (Hurd and Rohwedder, 2010; Chakrabarti *et al.*, 2011). Financial losses were associated with reductions in consumption, and many households reduced consumption even without experiencing financial losses (Shapiro, 2010; Christelis *et al.*, 2011).

* This paper is a revised version of the 2011 Hicks Lecture, sponsored by *Oxford Economic Papers*, and delivered at Oxford University on 16 May 2011.

These are large fluctuations in magnitudes that matter to people. Income, wealth, and joblessness are among the measures on which economists have traditionally focused. In the well-being literature too, a host of studies identify income and unemployment as two of the most important and reliable determinants of self-reported subjective well-being (SWB). Unemployment, in particular, typically exerts a larger negative influence than can be accounted for by the associated reduction in income. Although there is less literature on the effects of wealth on well-being, a 60% drop in the market has dramatic effects on expected future incomes, especially for those who are nearing a retirement to be funded out of accumulated saving or defined-contribution pension funds. These events had different economic implications for people of different ages; the elderly are not much affected by unemployment or employee compensation, but some are susceptible to stock market fluctuations. Young people may shrug off falls in wealth when they have not yet accumulated much and have many years before they need it, while their parents, nearing retirement, may see an imminent threat to their future.

The crisis brought harm to many, but it is a boon to researchers on subjective well-being, for whom it provided an unparalleled opportunity to examine how these events affected the standards of living, the emotional experiences, and life evaluations of those who lived through it. Our ability to make this evaluation is made possible by the data from the Gallup Healthways Well-being Index (henceforward GHWBI). Starting on 2nd January 2008, Gallup has run a daily (landline and cell-phone) telephone poll of 1,000 randomly-sampled adult Americans each day who are asked about how their lives are going, whether they are satisfied with their standard of living, and whether they experienced a range of feelings on the previous day. Over the three-year period examined here, from 2nd January 2008 to 29th December 2010, there are around a million observations on self-reported well-being, as well as on demographics, income, occupation, employment status and numerous health measures. These data allow daily tracking, not only of national averages, but of the outcomes of different groups.

In addition to investigating reports of well-being over the crisis, I look at a range of methodological issues. One is the long-standing question of whether variations in self-reported subjective well-being (SWB) over time correspond to what might be expected from cross-section analysis. A three-year period is too short to address the Easterlin paradox, that long-term growth appears not to generate the increase in SWB that would be predicted from the positive effects of income in the cross-section, but I can examine whether the large fluctuations in income and unemployment generate fluctuations in SWB that match the findings from the cross-section. The literature on well-being over the business-cycle—Di Tella *et al.* (2001, 2003), Wolfers (2003), Di Tella and MacCulloch (2009)—has relied on data pooled over several countries, rather than on tracking well-being over time within a single country as here, although Stevenson and Wolfers (2008) show that the output gap in the US predicts well-being. This literature also argues that people dislike inflation and unemployment, even controlling for their own experience, which would drive a wedge between the cross-sectional and time-series effects of

macroeconomic outcomes; see also Clark *et al.* (2010) for discussion and reinterpretation.

I also ask whether the temporal tracking of self-reported well-being measures is useful for economic policy beyond the standard dashboard of measures such as employment, income, and financial market indicators. Many happiness researchers argue that SWB provides a deeper and more comprehensive measure than standard economic indicators and should take priority over them. That SWB should be routinely collected by national statistical agencies, at least as a supplementary measure, was given impetus by the positive recommendations in Stiglitz *et al.* (2009), and many statistical offices in Europe are currently moving in this direction. Yet questions remain. Although SWB measures have led to many important insights in the cross-section—for example about the relative importance of income and unemployment, or of marriage and marital dissolution, or of non-priced amenities—the usefulness of average SWB for macroeconomic monitoring over time is far from established, and investigating that is one of my main concerns in this paper.

The paper is constructed as follows. I start, in Section 2, with a brief discussion of concepts of well-being, including reminders of longstanding concerns about happiness measures in general, and within SWB measures, the differences between hedonic and evaluative measures of well-being; this distinction is important in what follows. Section 3 discusses the behavior of the life evaluation over the crisis, and documents the sensitivity of the measure to questionnaire order effects. It also presents an attempt to repair the series to permit substantive analysis. Section 4 shows what happened to life evaluation and hedonic experience over the crisis, and Section 5 relates that experience to macroeconomic magnitudes such as income, unemployment, and the stock market. Section 6 concludes.

2. Preliminaries: concepts and measures of well-being

Self-reported well-being is my main topic in this paper, but I do not wish to approach it uncritically, but rather to keep in mind longstanding objections to these measures. Sir John Hicks, after whom this lecture is named, played a central role in banishing cardinal utility from economics, replacing it with ‘choice’ utility, an ordinal representation of preferences, together with a welfare economics that eschewed interpersonal comparisons and, at the individual level, emphasized the income required to attain a given standard of living. This long standing skepticism has been eroded in recent years by a literature in economics and psychology that has demonstrated the usefulness of self-reported well-being measures in a number of contexts. Nor has choice utility gone unscathed. Experimental work has also found biases in recollections of events that will sometimes cause choices to deviate from people’s own preferences, Kahneman *et al.* (1997) suggesting that direct measures of momentary feelings, integrated over time, might be more reliable guides to decision than remembered utility.

Yet many of the original concerns remain important. One of the most important is associated with adaptation to circumstances. If people become accustomed to economic misery, so that the response of SWB to such pain is only temporary, the continuing harm is no less real nor demanding of policy attention just because people say that they are used to it. Sen (1985, p.14) notes that ‘a person who is ill-fed, undernourished, unsheltered, and ill can still be high up in the scale of happiness or desire fulfillment if he or she has learned to have “realistic” desires and to take pleasures in small mercies’. By accepting people’s own assessment in such circumstances, ‘the metric of happiness may, therefore, distort the extent of deprivation in a specific, and biased way’, and ‘it would be ethically deeply mistaken to attach a correspondingly small value to the loss of well-being because of this survival strategy’ (Sen, 1987, pp.45–6). I accept these arguments, and believe that we should not base policy on a measure that is subject to hedonic adaptation. Yet the extent to which any particular measure of SWB is actually subject to the adaptation critique is a question that can be investigated empirically, so that it is possible that Sen’s concern is hypothetical, or is hypothetical for some measures but real for others. Note also that Sen does not deny the goodness of happiness in and of itself, only that it is an unreliable indicator of overall well-being. Put differently, ‘[t]he central issue is not the significance of happiness, but the alleged insignificance of everything else, on which many advocates of the happiness perspective seem to insist’ (Sen, 2009, p.173). Beyond that, not everyone accepts that questions about pleasure, or about how life is going, reveal anything about those aspects of life that are the most important or most meaningful (Nussbaum, 2007).

A second concern, documented in the psychology literature, focuses on the measurement of SWB, particularly in response to questions about the evaluation of life as a whole. The answers to these questions are often treated as if ‘global feelings of well-being . . . remain relatively constant over extended periods, and that people can describe them with candor and accuracy’ (Campbell, 1981, quoted in Schwarz and Strack, 1999). But as Schwarz and Strack’s review makes clear, actual reports ‘do not reflect a stable inner state of well-being’ that is always available, but rather judgments that are formed in response to the question, which makes them subject to context effects (Schwarz and Strack, 1999, p.61). That these warnings must be taken seriously will be amply documented in the results presented below.

One key distinction is between ‘living life’ and ‘thinking about life’ (Kahneman and Riis, 2005). On the one hand, there is a range of feelings that comprise emotional well-being (or hedonic well-being or experienced happiness); these feelings, such as enjoyment, sadness, happiness, anger, stress, or worry, are mostly evanescent and rapidly forgotten, but make up the hedonic texture of life as we live it. Because they are forgotten, they can play only a limited role in choice. On the other hand, there is the assessment or evaluation of life as a whole that is prompted by questions about life satisfaction or about how life is going overall.

The gold-standard for measuring emotional well-being is experience sampling, by which people are randomly prompted to record their current activities and

feelings; a less demanding alternative is the Day Reconstruction Method (DRM) (Kahneman *et al.*, 2004), by which people are asked to reconstruct their activities over the previous day, and to associate feelings experienced during each episode. Experience sampling and the DRM both allow the construction of hedonic well-being measures by summing over episodes. The American Time Use Survey contains a DRM module. In the Gallup data, a DRM is not feasible, but respondents are presented with a list of emotions, such as stress, worry, and happiness, and are asked to say yes or no to questions about whether they experienced a lot of each on the previous day. In this paper, I shall look at a range of these feelings, including happiness, smiling, enjoyment, sadness, worry, stress, and anger, as well as a similar question about physical pain.

Life evaluation is most often measured by a question that asks people to report on a several point scale how satisfied they are with their life as a whole. In the Gallup surveys, the question is Cantril's (1965) Self-Anchoring Scale, a ladder question which asks respondents on which rung they stand of an 11-rung ladder where zero is the 'worst possible life for you' and 10 is the 'best possible life for you'. It is possible that this ladder question, analysed here, is a purer question of life evaluation than is life satisfaction, where the use of the word 'satisfaction' invites contamination by the respondent's current hedonic state. The 11-point scale, with equally spaced 'rungs,' invites respondents to think of equal intervals as having the same interpretation, so that the difference between one and two is the same as the difference between eight and nine, which permits analysis in terms of means rather than, for example, ordered probits, though differences in results are typically not important.

That hedonic and evaluative well-being behave differently has been reported in a number of studies. Kahneman *et al.* (2004) find, contrary to their initial hypothesis, that measures of affect are more subject to adaptation and less tightly linked to long-term life circumstances than are measures of life evaluation. Kahneman and Deaton (2010) find that reports of life evaluation (Cantril's ladder), positive affect, negative affect, and stress are all better at higher incomes. However, while the effects of income on affect and stress satiate at an annual income of around \$75,000, higher incomes continue to be associated with higher life evaluation beyond \$75,000, with the average ladder approximately linear in the logarithm of household income. In a similar vein, life evaluation, but not affect, is better among more educated people.

Figure 1 explores the distinction in another way and shows the average reports for life evaluation and for affect for each day of the week. The average reported ladder, in the top left panel, shows essentially no variation over the week, while positive affect, negative affect, and stress (here coded so that higher is always better) all show superior hedonic outcomes at weekends than in the week, with Fridays intermediate between week and weekend. At this daily frequency, affect responds to circumstance, while life evaluation does not. (Note that Americans generally evaluate their lives very highly, around seven on a scale of zero to 10, and that most people experience a great deal of positive emotion, little negative emotion,

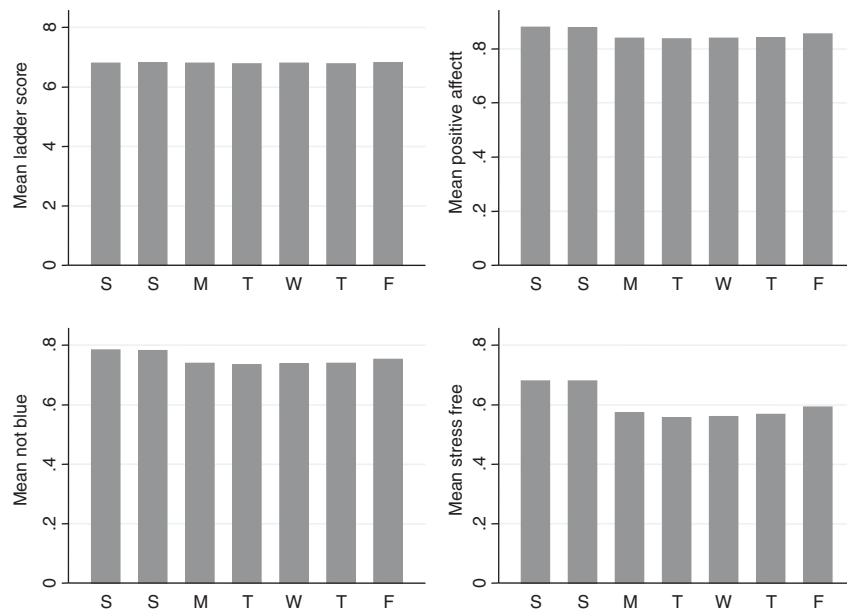


Fig. 1 Ladder scores, positive affect, negative affect, and stress over days of the week.

Note: Averages are taken over all respondents from 2nd January 2008 to 29th December 2010. Positive affect is the average of the fractions of respondents who said that they experienced a lot of smiling, enjoyment, and happiness in the previous day. Not blue is one minus the average of respondents who reported a lot of worry or sadness in the previous day. Stress free is the fraction of respondents who did not report a lot of stress on the previous day.

and no stress.) There are also marked differences in patterns of life evaluation and of affect over the life-cycle. Stone *et al.* (2010) replicate the now familiar finding that life evaluation falls with age until middle age, rising mildly thereafter, but also a remarkably large decline in reported stress, worry, and anger with age. This is consistent with socioemotional selectivity theory (Carstensen *et al.*, 1999), according to which people, as they age, acquire emotional capital, a set of skills that allows the ever-more successful avoidance of negative emotional experiences. These benefits of age, together with more modest increases in positive emotional experience, perhaps offset the increase in physical pain and may help account for the increase in overall well-being with age in spite of deteriorating health.

That the ladder responds to long-term circumstances and emotions to short-term circumstances is further supported by international evidence from Gallup's World Poll, which asks similar questions in more than 150 countries around the world. In Deaton (2008), I show that, across countries, average ladder scores are linearly related to the logarithm of *per capita* GDP. There are large differences across countries, from Togo, Benin, and Chad, with average ladder

scores between three and four, to Denmark, with an average ladder score around eight. An increase of one in the log of *per capita* GDP is associated with an increase in the average ladder score of 0.84, and the simple correlation is 0.83 (Deaton, 2008, Table 1, Fig. 1). For future reference, it is important to note that international differences in income are very large compared to what is seen in time-series; a difference in log GDP of one corresponds to a GDP that is 2.7 times as large or about 20 years of rapid economic growth, while Togo's *per capita* income is barely 2% of that of the United States, a difference that took a quarter of a millennium to establish.

In contrast to life evaluation, the average hedonic experiences of countries are only weakly related to *per capita* GDP. The fraction of the population that reports a lot of happiness yesterday is only mildly related to national income, essentially because of a few outliers such as, at the bottom, Togo, which is notably unhappy and, at the top, the US, where the pursuit of happiness is constitutionally guaranteed. Otherwise, there are happy and unhappy countries at all levels of GDP *per capita*. Hence, as far as self-reported happiness (the affect) is concerned, the data support Sen's argument that even the poorest people in the world are often happy, although when we look at life evaluation, poor people generally recognize that their lives are going badly. That the hedonic and evaluative components of well-being have such different correlates implies, not only that they are different concepts that reflect different parts of human experience, but that we must consider each separately in assessing what happened over the financial crisis. A single broad measure of 'happiness' will not do. From now on, I shall use the term 'happiness' for happiness proper, referring to the hedonic experience of being happy, and I shall keep life evaluation for the judgment of life as a whole, as in the Cantril ladder. The evidence cited in this section shows that the emotional measures adapt to life circumstances, at least over the long term, and that this is not true, or at least less completely true, for measures of life evaluation.

3. Life evaluation in the crisis

Figure 2 shows the daily average of the ladder from the beginning of 2008 to the end of 2010; I have shown it as a 21 day moving average in order to remove the day to day sampling variance which otherwise obscures the trends. In order to minimize jumps, I use a unit sum triangular weighting scheme with maximum weight (1/11) on the central observation, with weights declining by equal amounts (1/121) for ten days on either side. The broken line is for those aged 60 and over, the solid line for the whole population. As has often been found in the literature—see Stone *et al.* (2010) for these data—the elderly generally do better on life evaluation measures. Although there are features of these plots that make sense—for example, the sharp drop in well-being around the time of the collapse of Lehman Brothers—they are often very hard to explain. The oddest feature is that the ladder should be higher at the end of the period than at the beginning; indeed, according to this figure the general population had more than recovered from the

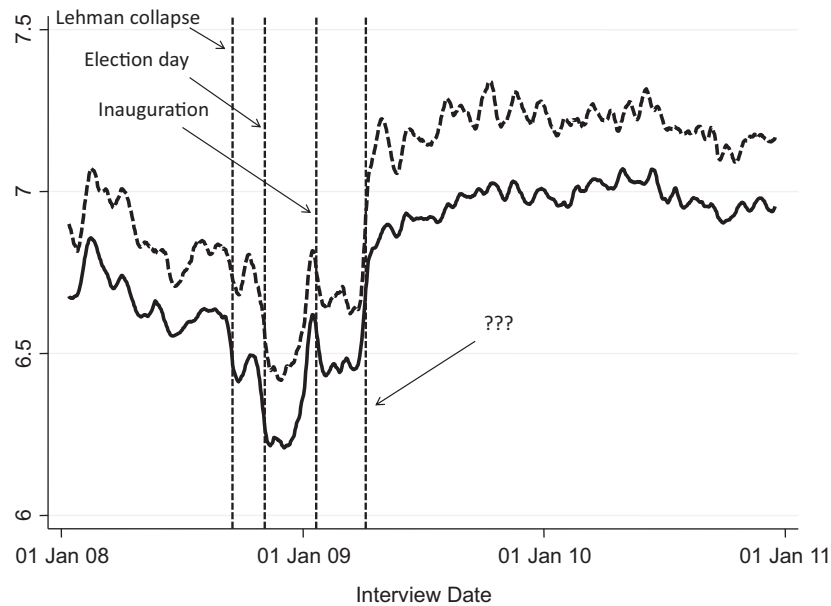


Fig. 2 Average ladder scores for general population and for those 60 and over.
Notes: The broken line is for the population aged 60 and above, the solid line for the whole population. The underlying data are daily averages, and the figure shows 21-day triangularly weighted moving averages.

crisis by the spring of 2009, even though unemployment was still rising, and even though the stock-market, although past its low in March, was still far below its values before the crash. The graphs also show a sharp drop in well-being on Election Day, and a somewhat more credible spike on the day that President Obama was inaugurated. The huge increase on 6th April 2009, is substantively inexplicable; although the moving average smooths over days, the key dates are easily checked by going back to the raw data. The boost to the ladder would require a more than doubling of *per capita* GDP, and the main events of the day—the earthquake in L’Aquila in Italy, Robert Gates unveiling of the US defense budget, or even the winning of a country music award by American Idol winner Carrie Underwood—are surely insufficient to explain an increase in well-being that is the single largest change over this otherwise very eventful period.

In fact, the largest changes in the ladder are driven, not by real events, but by changes in question order in the GHWBI survey, the context effects discussed in Section 2. Gallup’s own analysis is presented in Newport (2011), and in more detail in Agrawal and Harter (2011), an analysis that was carried out in parallel with this paper, and is consistent with the results that I present. I confine myself here to a summary of the main points, and an outline of the methods used to correct the data.

Gallup uses its daily poll to collect information on political preferences, including whether the respondent plans to vote, if so for whom, who is preferred in a primary, whether the respondent approves of the President's performance, or whether or not the country is headed in the right direction. Because 2008 was an election year, and because these polling questions are important to Gallup whose statisticians are well aware of context effects, they were placed first in the survey, immediately prior to the ladder question. The precise questions were changed frequently depending on the election calendar so that, for example, voter preference questions were dropped on the eve of the Presidential Election, and the Bush approval question was dropped when he left office. Changes in ladder scores at the time of these events prompted Gallup to randomly split the sample, with half the respondents being asked the political questions as usual, and half the respondents being asked none, so that the first question they were asked was the ladder of life evaluation. This randomized controlled trial enables us to assess the effect of asking the political questions as opposed to not asking them, and it turns out that these questions cause a large negative effect on reported well-being. People appear to dislike politics and politicians so much that prompting them to think about them has a very large downward effect on their assessment of their own lives; over the 111th Congress (2009–2010), only 25% of the population approved of Congress, then one of the lowest numbers in history. Indeed, as we shall see, the effect of asking the political questions on well-being is only a little less than the effect of someone becoming unemployed, so that to get the same effect on average well-being, three-quarters of the population would have to lose their jobs. Not everyone becomes unemployed, but either half or all of the respondents are asked the political questions.

Figure 3 shows the 21-day moving averages of mean ladder scores, split by the two randomized forms of the survey. The two forms were identical until 5th January 2009, after which the political questions were dropped from Form 2; this is shown as the beginning of the shaded area in the figure. During the shaded period, which ends on 5th April, Form 1 respondents were asked political questions (which changed occasionally) while Form 2 respondents were asked none. The difference between the solid and the broken line, about 0.6 of a rung on the ladder, is the negative effect of the political questions on life evaluation; it is about the same size as the decline in the ladder from the beginning to the end of 2008. The Form 1 spike is on inauguration day; on this day, and this day only, there were no political questions on either form; the difference between the two means on that day was 0.15, and appears much larger in the figure because of the moving average.

After the end of the shaded period, on 6th April 2009, Gallup added a 'transition question' between the political questions and the ladder question. After asking the political question to Form 1 respondents, the interviewer asked, 'Now thinking about your personal life, are you satisfied with your personal life today?' This is intended to remove the effects of the political questions by refocusing the respondent's attention, and the figure shows that it is remarkably, if not totally, successful

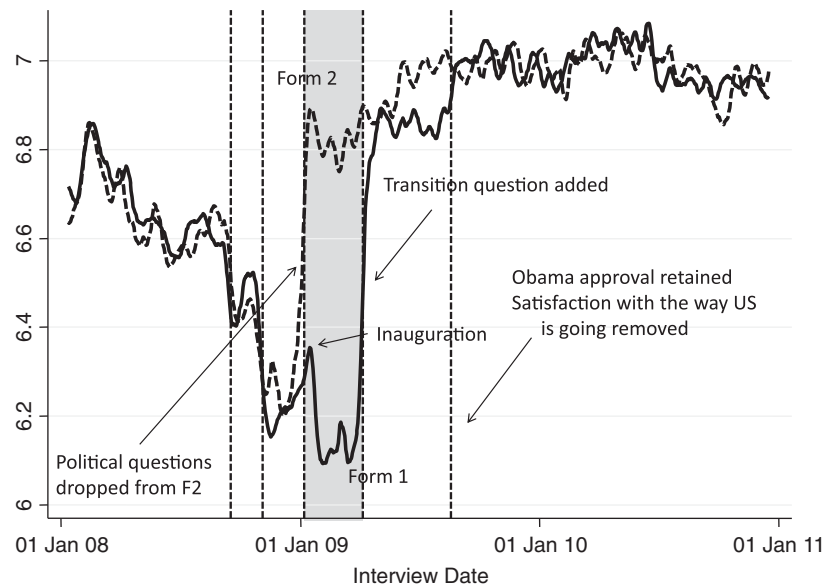


Fig. 3 Average ladder score by questionnaire type.

Notes: 21-day triangular weighted averages of averaged daily data for the population as a whole. Forms 1 and 2 are two versions of the questionnaire, identical until the beginning of the shaded area, after which the political questions at the beginning of the questionnaire were removed from Form 2. The end of the shaded period shows the date on which a 'transition' question about life satisfaction was added to the questionnaire after the political forms and before the ladder question.

in doing so at least for this measure; see Schwarz and Schuman (1997) for discussion of such buffer questions to offset order effects. From 6th April on, the difference falls from 0.6 of a rung to less than 0.1 of a rung. Starting 18th August, the political questions were reduced from two to one; the question about satisfaction with the way that things are going in the US was dropped, retaining only a question about approval of President Obama. Once that is done, the Form 1 and Form 2 means are indistinguishable.

For the period from 6th January 2009, we have available a 'clean' measure of the ladder from the half of the sample who were asked no prior political questions. This is useful but leaves unanswered the main question, which is the evolution of well-being through 2008, for which it is necessary to make some corrections. There is no assumption-free or non-parametric way of doing this, but there is a very strong correlation between the ladder and the responses to questions about the respondent's standard of living, questions that are asked later in the survey. There are two relevant questions: 'Are you satisfied with your standard of living, all the things you can buy and do?' with a Yes/No answer, and 'Right now, do you feel your standard of living is getting better or getting worse?' with answers 'Getting

better, staying the same, or getting worse'. The answers to both of these questions predict the ladder with high significance. When I use the 'clean' data, from Form 2 for 2009 on or after 6th January, to regress the ladder on the standard of living answers, and collapse the predictions to get a daily time series, the correlation between this prediction and the actual ladder is 0.54. When I do the same for the period before 6th January, and during which there are much larger fluctuations in the series, the correlation is 0.80.

My corrections are as follows. Using the 'clean' data, I fit an ordered probit to the 11 ladder values (0 to 10) using the standard of living questions as predictors, and then use the results to predict the probabilities of each rung for each household in the period up to 5th January 2009. These predicted probabilities are then used to give an expected ladder value for each household. Figure 4 shows the daily mean of these predicted values as the lower (light broken) of the two upper lines; for the period after 6th January, I use the actual ladder from Form 2, dropping Form 1. The bottom line is the original, contaminated, ladder score, and even this first correction makes a dramatic difference to the behavior of well-being over the crisis. Most importantly, and in contradiction to the original graph, the corrected series estimates that, on average, well-being had not recovered to its January 2008 level by the end of 2010.

This correction likely still understates well-being at the beginning of the period. The reason is that the values used for the adjustment, the answers to the standard of living question, are themselves biased downward by the political questions, even though there are 16 or so (depending on skip patterns) (mostly health related) questions between the ladder and the standard of living question; see Bishop (1987) for an example of an order effect that persists in spite of the insertion of up to 101 intervening questions. That this happens here can be established from an analysis of the period from 6th January (when the political questions were removed from Form 2 respondents) to 5th April 2009, when the 'transition' question was added to Form 1. From 21st January, after the inauguration, the political questions (approval of Obama, satisfaction with way things are going in the US) did not change, so there is clean randomized treatment and control design. Respondents who were given the political questions over this period were 3.6 percentage points less likely to report satisfaction with their standard of living than those who were not asked; the *t*-value on this is greater than 10. For the future standard of living, where there are three answers, an ordered probit shows a similarly significant downward shift. Since political questions in one form or another were asked throughout 2008, it seems likely that the standard of living answers were biased downward throughout that period, which means that my first correction is probably insufficient.

The political questions changed several times over the election period, so that I cannot immediately apply the results from the randomization period to the earlier period. However, a detailed analysis shows that changes in the political questions had only a small effect compared with the effect of there being political questions at all; with 1,000 observations every day, regression discontinuity methods can be used

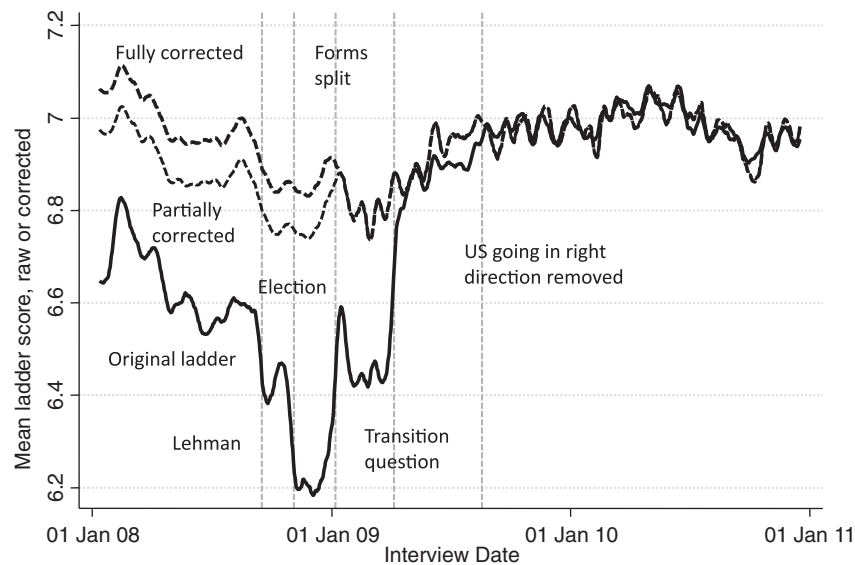


Fig. 4 Correcting the ladder for context effects.

Notes: The bottom line is the original, uncorrected ladder. The middle line uses only Form 2 after the political questions were removed, and imputes the ladder prior to the split from the answers to two standard of living questions. The top line makes a further correction for bias in the answers to the standard of living questions. See text for details. All plots are 21-day triangularly weighted moving averages of population means.

to investigate what happens as the questions change. So it seems that any political question will depress the answers to the ladder and that the effects persist through the questionnaire into the standard of living and some other questions.

I have therefore made a second adjustment to all of the data prior to 6th January 2009 based on the use of the pure randomization period discussed above. The method is as follows. From the ordered probit calculated for the first adjustment and estimated on the full 'clean' period, I calculate the predicted $x'b$, and then compare the values of this index between experimentals and controls over the randomization period. This gives me an estimate of the downward bias from the political questions, which I add back in, to give $x'b + \delta$, say, which I then use in place of $x'b$ with the originally estimated cut points, to calculate a predicted set of probabilities for each rung for each household. From these, I calculate a predicted ladder for each household. The moving average of the daily mean of this series is the top line in Fig. 4. This second adjustment is modest compared with the first—the standard of living responses are not nearly as badly affected as is the ladder by the political questions—but it further shifts up the early well-being measures, suggesting more strongly still that Americans are indeed still some way from recovering from the crisis.

4. Riding the waves: life evaluation and hedonics over the crisis

Stock market fluctuations have a direct effect on those who own assets, and are thus important for linking the financial crisis to the well-being of individuals. Ideally, the data would allow us to separate people with direct, indirect, or no involvement in the market, but this information was not collected. One imperfect way to look at this issue is to look at effects across the age and income distributions. Rather than use the corrected ladder itself, I look directly at the fraction of people who report that they are satisfied with their standard of living. This series tracks the ladder closely, which is why it works for the correction, and although it is also contaminated by the political questions, it is much less so than is the case for the ladder. Judgments about standard of living are of interest in their own right, and can readily be decomposed by age or income. Figure 5 shows the results disaggregated by age group, showing the working age population, 18–59, as well as 60–69, 70–79, and 80 and over. As is the case for the ladder itself, the fraction of people satisfied with their standard of living rises with age, even though the level of income itself falls with age; among other things, this result demonstrates that neither the ladder nor the standard of living question is a recode of income. All groups showed some decline in their standard of living from spring 2008 until late in the year, and there was some recovery through 2009. But the severity of the effects diminishes with age, and the crisis had a barely perceptible effect on those aged 80 and above. Interestingly, the 60–69 age group was affected as severely as those under 60, possibly because of their greater dependence on the market through defined-contribution pensions. The oldest group, whose pension income is unlikely to depend on the market, are the least affected by the crisis.

Figure 6 shows the corresponding information by household income group; this is somewhat less satisfactory, because about a quarter of respondents either refuse to answer the income question or report that they don't know, and are therefore dropped from the analysis. That people who live in higher income households are more likely to be satisfied with their standard of living is no surprise. More difficult to explain, at least through the wealth channel, is that the impact of the financial crisis on perceived standards of living is much larger—about twice as large—for those whose incomes are below \$4,000 a month than for those whose incomes are more than \$4,000 a month; compare the top two graphs in Fig. 6 with the bottom two, especially towards the end of 2008. An obvious hypothesis is that the rich were affected by the wealth shock and the poor by unemployment but, as we shall see, the decline in standard of living perceptions happened before unemployment began to rise.

The effects of the financial crisis can also be seen in the reports of hedonic well-being. In Fig. 7 I have grouped together the various experiences, averaging and changing signs as necessary, so that higher numbers always mean worse outcomes. The underlying positive emotions (smiling, enjoyment, happiness) appear to be unaffected by the presence or absence of the political questions at

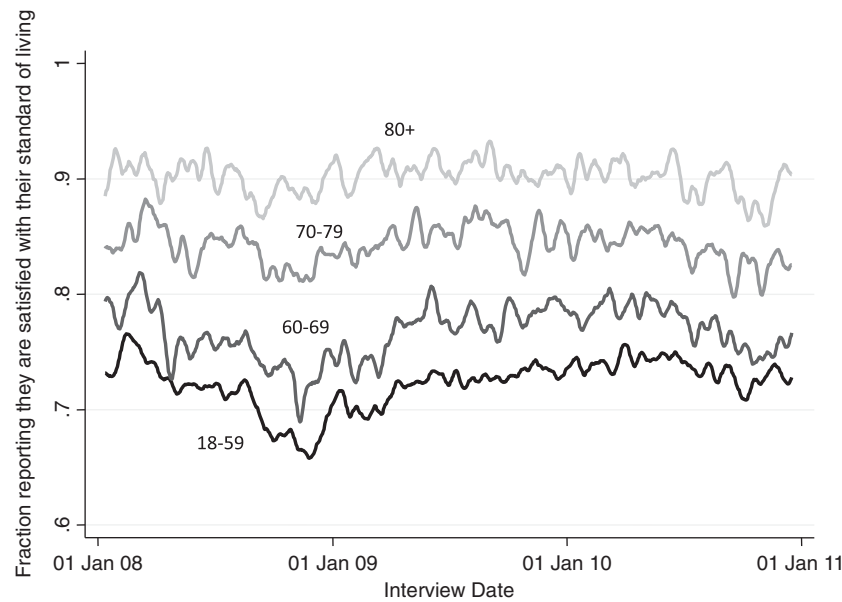


Fig. 5 Fractions of the population satisfied with their standard of living, by age group.

Notes: 21-day moving averages of daily averages by age group.

the beginning of the survey; there is no significant difference in outcomes by form over the randomization period. These questions are deeper in the questionnaire than even the standard of living questions. However, people are somewhat more likely to express negative emotions when they have been asked the political questions at the beginning of the survey. Over the randomization period, asking the political questions increases reported anger by 0.006, stress by 0.017, and worry by 0.006; the effect on worry is barely significant, that on anger has a t -value of 2.5, and that on stress a value of 5.0. These results mean that the top and bottom lines in Fig. 7 overstate worry, stress, and anger at the beginning of the period, and although even the largest effect, for stress, is not very large, it is enough to suggest that the average of stress and worry is still somewhat higher at the end of 2010 than it was in early 2008.

The results in Fig. 7 are an important part of the overall story. Hedonic experience, particularly worry and stress, but also physical pain, deteriorated during the crisis, becoming rapidly worse during the summer and fall of 2008, recovering briefly during the year-end holidays, only to reach their worst values around the time that the stock market was at its lowest. There is a (small) increase in hedonic affect in all of these measures on St Valentine's Day, and a much larger one around the Christmas holidays. As the stock market revived, negative affect fell. By mid-2010, there is very little trace of the crisis in these measures—though admittedly it is hard to detect small trends among the variance—even though

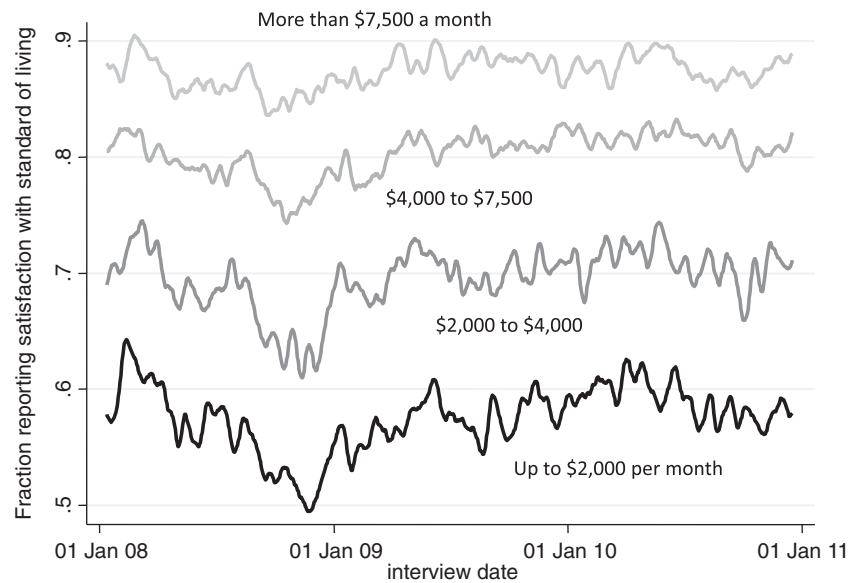


Fig. 6 Satisfaction with standard of living by household monthly income level.
Notes: Question is about whether you are satisfied with your standard of living plotted as 21-day moving averages by household income groups.

the crisis continued in terms of lower incomes, employment, and home and stock prices. These results are consistent with hedonic adaptation, especially in the positive measures (happiness, enjoyment, smiling, not being sad). Worry and stress (which behave similarly to one another) are particularly sensitive to the crisis, at least in terms of the increase in the fraction of the population reporting them. Hedonic adaptation is somewhat less clear for worry and stress than for the positive emotions. Although these series came close to full recovery by mid-2010, they never quite reached their original values, especially if we recognize that there is some upward bias in the series at the start of the period.

5. Correlates of changes in well-being over the crisis

The literature on the correlates of SWB has identified many factors that consistently show up as important for one or more measure. Some of these factors, like health, education, marital status, or religiosity, are also important in the GHWBI data, but they do not change quickly enough at the population level to be candidates for explaining the fluctuations that I am considering here. This is obviously not true of income and of employment, both of which saw rapid changes over the crisis; the literature finds that both contribute to well-being, and that becoming unemployed has a very large negative effect on well-being. Two early studies are Clark and Oswald (1994), who find that being unemployed is worse than being divorced or

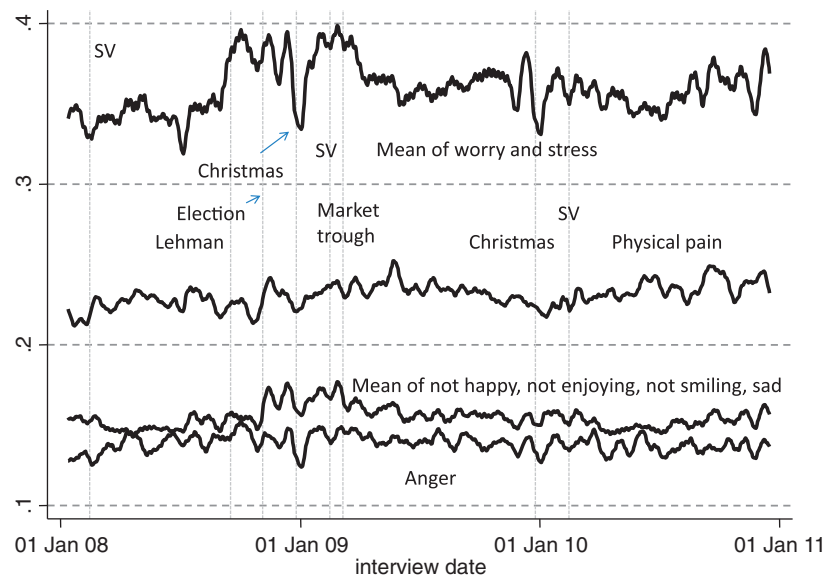


Fig. 7 Negative hedonic experience, 2008 to 2010.

Notes: 21-day moving averages of underlying daily averages over the whole population. The top line is the average of the fractions reporting worry or stress on the previous day. The second line is the fraction reporting physical pain on the previous day. The third line is the average of the fractions who reported sadness, or did not report happiness, enjoyment, or smiling on the previous day. The bottom line is the fraction of those reporting anger on the previous day. The vertical broken lines represent, from left to right, St Valentine's Day 2008, Lehman collapse, Election Day 2008, Christmas 2008, St Valentine's Day 2009, the market trough, Christmas 2009, and St Valentine's Day 2009.

separated, and Winkelman and Winkelman (1998), who use the German panel data to show that the effect of unemployment operates in addition to the effects of lost income. The relationship between stock market prices and well-being seems to have been less researched, perhaps because it is not easily addressed in cross-sectional data; while it is true that individuals are differentially affected according to their exposure to a market shock, the effects of the shock are identified from whatever determines stock holding, not from the shock itself. Exploring the link between stock prices and well-being requires relatively high frequency time-series data.

To tie the well-being outcomes in the GHWBI to the crisis-related behavior of income and unemployment requires, as a preliminary, matching of the Gallup data to official series on income and unemployment. This is not straightforward, given that one is daily, seasonally-unadjusted while the other is typically monthly and seasonally adjusted. Beyond that, the GHWBI asks a single question on monthly household income, and the respondent is asked to choose from a set of income intervals; we have no way of knowing how the respondent interprets this question, nor whether the respondent is well-informed about his or her household's income.

Single questions about household income are far from perfect, but are still informative, see Micklewright and Schnepf (2010). Gallup has also experimented over time with several sets of questions about work status, so the data do not contain a consistent unemployment measure over the whole three-year period.

I have calculated an estimate of log income for each respondent by filling in mid-points assuming that income is log-normally distributed; this is possible for the three-quarters of all respondents who answered the income question. This is clearly at best a rough and ready measure of income, but averaged over congressional districts, there is a correlation of 0.96 with log median income from the American Community Survey (Kahneman and Deaton, 2010). For the current comparison, I have averaged log income over all respondents for each day, and over each day in each month, and compared the results with BEA data on *per capita* disposable income and on total employee compensation, both of which are available only on a seasonally adjusted basis.

Figure 8 shows the means of the logarithms of reported income by age group. For the typically working age-group 18 to 59, there is a sharp reduction in reported incomes from late 2008 through to the middle of 2009. This roughly corresponds to a drop in employee compensation in the BEA statistics. However, this reduction in income is either much less or altogether absent among the other age groups. From the summer of 2009 onwards, the reported incomes of the working-age group are not distinguishable from those of the age group 60 to 69, who are typically retired. The incomes of the two older groups, although lower, actually increased after early 2009. It is notable that the two stimulus-related payments, one in April to June 2008 (tax rebates) and May and June 2009 (one time \$250 transfers sent to those receiving social security, SSI, railroad retirement benefits, and veteran benefits) make no appearance in this figure, although a graph of the official series is dominated by those two spikes. This is an interesting finding in its own right. Either people did not recall these transfers or, more likely, did not think of them as income.

Unemployment series are shown in Fig. 9. For the periods where they overlap, the two series are quite close, reflecting the considerable amount of work that Gallup has done to match the data from the BLS. The timing of unemployment is important, and is quite different from the timing of the stock market. Although unemployment inched up during the events of the fall of 2008, the big rise was between November 2008 and March 2009—a period when the market was also falling—but it then continued to rise, albeit more slowly, reaching a (seasonally unadjusted) peak of 10.6% in January 2010. The unemployment consequences of the financial crisis came much later than the financial crisis itself.

The ordinary least squares regressions in Table 1 show that, in the cross-section, unemployment and income play the expected roles in conditioning both life evaluation and hedonic experience. People who report themselves unemployed are 0.80 of a ladder rung below those who are in full-time employment. Compensating for this would require an increase of 1.89 in log income, or more than a six-fold increase in income itself; this compensation is close to, but

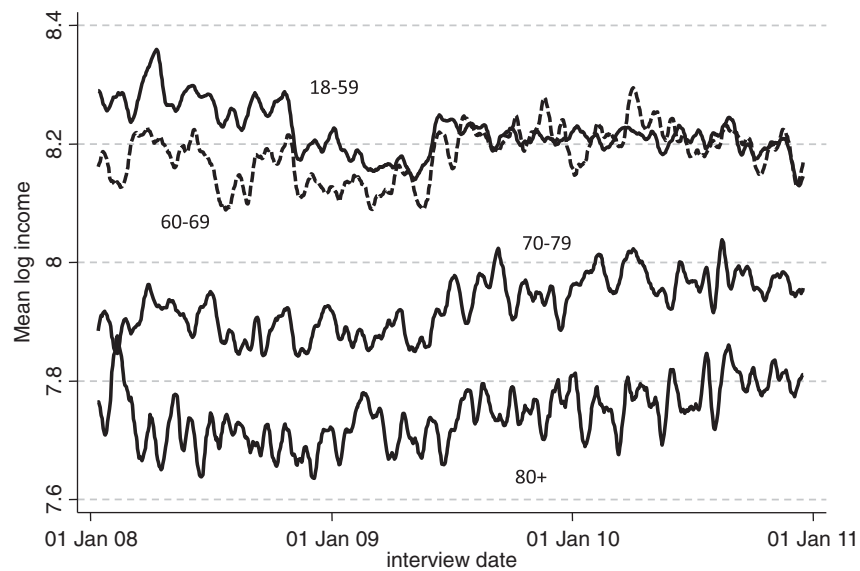


Fig. 8 Mean log income by age group.

Notes: 21-day weighted moving averages of raw means of reported log incomes. Note that 25% of the sample either refused to answer or said they did not know.

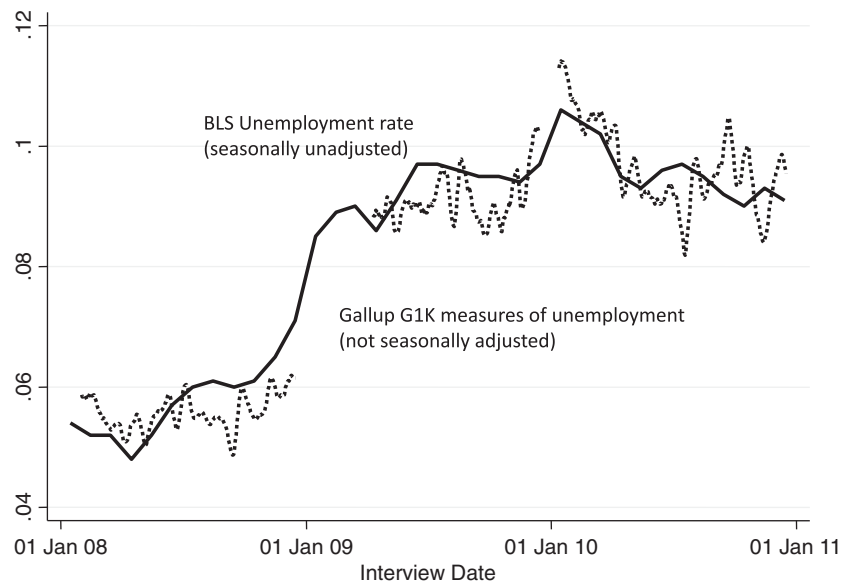


Fig. 9 Unemployment rates from BLS and Gallup Healthways Wellbeing Index Poll.

Notes: Both series are seasonally unadjusted.

Table 1 Regression coefficients of unemployment and of log income

| | Unemployed | t-value | Log income | t-value |
|--------|------------|---------|------------|---------|
| Ladder | −0.821 | (42.4) | 0.412 | (75.1) |
| Smile | −0.057 | (14.3) | 0.027 | (23.9) |
| Enjoy | −0.035 | (9.2) | 0.038 | (35.4) |
| Happy | −0.047 | (13.9) | 0.033 | (34.5) |
| Anger | 0.057 | (15.4) | −0.024 | (23.2) |
| Worry | 0.179 | (36.9) | −0.050 | (36.5) |
| Stress | 0.052 | (19.1) | −0.031 | (21.8) |
| Sad | 0.114 | (29.2) | −0.048 | (43.9) |
| Pain | 0.048 | (11.2) | −0.058 | (47.7) |

Notes: Estimated using approximately 163,000 observations from January 2nd 2010 to August 31st 2010 for which there are data on both unemployment and income using the final definitions of the former. For the ladder, the left-hand side variable runs from 0 to 10; for the other experienced emotions, the left-hand side variable is dichotomous. Also included in the regressions were dummies for other kinds of employment status, marital status, age group, sex, Hispanic status, and race. The comparison group for unemployment is full time employees.

somewhat smaller than that calculated by Winkelmann and Winkelmann (1998) using German panel data. Income is also good, and unemployment bad, for both positive and negative emotions, as well as for physical pain. (Note that these estimates are likely inflated by reverse causality; disabilities cause pain and make people less likely to be in the labor force, and people with more positive emotion and less negative emotions may have better outcomes as a result; see also Graham *et al.*, 2004.) Unemployed people are particularly more likely to be worried and sad, and these are also the two cases (along with stress) where the ratio of the unemployment coefficient to the log income coefficient is the highest. As repeatedly found in the literature, unemployment is associated with poor life evaluation and poor hedonic outcomes, and is so independently of (and in addition to) the effects of (current) lost earnings.

Based on these cross-sectional results, we might expect variations in income and unemployment to be the main drivers of well-being over the financial crisis, with the role of the stock market an open question. But it should be already clear from the timing in the figures that unemployment cannot provide a good account of the time-series, at least on its own. To look at the correlates more formally, I start with the daily averages and with the relation between income and the ladder.

The correlation between mean log income (from the Gallup data) and the mean corrected ladder is 0.32, and the regression coefficient is 0.71 ($t=11.0$) which is substantially larger than the cross-sectional coefficient in Table 1. This difference is in the opposite direction to that predicted by the Easterlin paradox. Both the correlation and the regression coefficients are biased down by the fact that both the daily ladder and daily mean log income contain substantial day to day sampling error. The daily corrected ladder is correlated with the S&P 500 at 0.46, and the

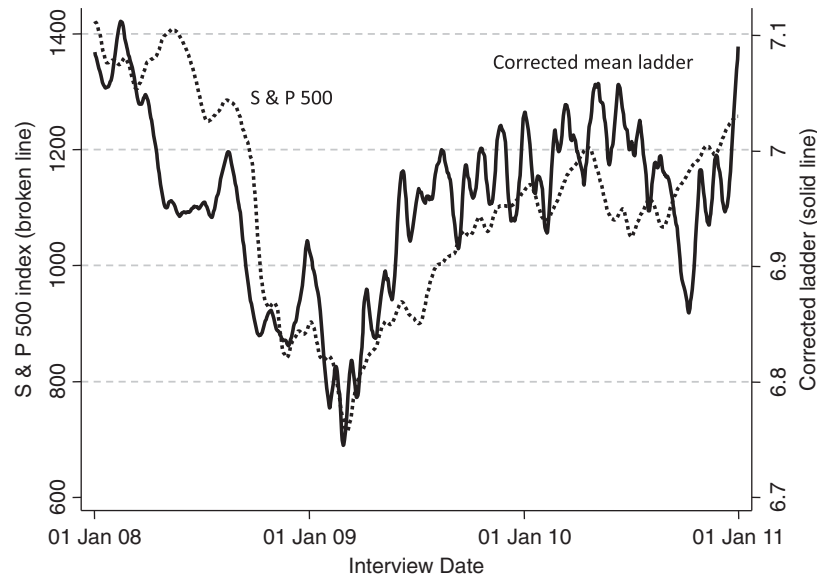


Fig. 10 Corrected mean ladder and S&P 500 index.

regression of the corrected daily ladder on mean log income and the log of the S&P 500 index reduces the coefficient on the former to 0.33 ($t=4.2$) while the coefficient on the log of the stock index is 0.31 ($t=11.4$). The high correlation between the corrected ladder and the market is shown in Figure 10; although it is clear that most of the correlation comes at the (low) frequency of the crisis itself, some higher frequency correlation is visible too.

Moving to monthly data, so as to use the unemployment data from the BLS at the price of reducing the number of observations to 36, the ladder remains correlated (0.43) with the S&P 500, and in a regression of the corrected ladder on the log of the S&P 500, unemployment, and the log of employee compensation, unemployment is significantly positively related to the ladder, the log S&P 500 has a t -value of 7.2, and employee compensation is insignificant.

Averages of emotions and physical pain are also significantly correlated with average log income on a daily basis; as is the case for the ladder and once again contrary to the Easterlin paradox, the coefficients are all (absolutely) larger than the cross-section coefficients in Table 1. On monthly data, with only 36 monthly data points, fluctuations in the S&P 500 are correlated with all of the hedonics except physical pain; neither the unemployment rate nor employee compensation have any effect conditional on the market. For physical pain, only the unemployment rate has predictive power in the monthly data. That the market is more closely related to hedonic outcomes and to the ladder in monthly than in daily data is consistent with Fig. 10 (and similar figures for hedonics) in which there is little high frequency correlation.

In addition to these regression results, it is informative to check whether the changes in SWB over the whole three year period are what might be expected from the cross-sectional results. We know that unemployment is twice as high at the end of the period, and that some measures of income also fell, and we need to know if these changes—which are very large in business cycle terms—show up in the well-being data. For the ladder, the corrected average value in January 2008 was 7.07, which had fallen to an average of 6.99 in December 2010. To make the point, I take the largest possible changes in income and unemployment that are consistent with the data, a decline of 5% in the former, and an increase of five percentage points in the latter. From the cross-section regressions in Table 1, these changes should cause the ladder to fall by 0.06, somewhat less than calculated, but surely within any reasonable margin of error given the corrections for the order effects. Similar calculations for the hedonic experiences show that the positive emotions do not respond as they should. Predicted from unemployment and income, the fractions reporting smiling, enjoyment, and happiness should all decrease by 0.004. In fact, enjoyment is actually higher (by 0.014) at the end of the period, and happiness and smiling are exactly the same at the end of 2010 as they were at the beginning of 2008. By contrast, the negative emotions are all higher at the end of the period, and the increases are larger than predicted from the cross-section regressions and the changes in income and unemployment, by twice or more, especially if we allow for some elevation at the beginning of the period from the presence of the political questions.

We might conjecture that the positive emotions are more subject to hedonic adaptation than are the negative emotions, or that the negative emotions respond, not only to personal circumstances, but to the macroeconomic environment in general. But there is a more important point. Even in the face of some of the largest macroeconomic disruptions since the great depression, both the actual and predicted changes in these well-being measures are very small, possibly smaller than our ability to detect given the difficulties of measuring well-being. Even with 30,000 observations each month as here, the standard errors of the monthly averages for the hedonics range from 0.002 to 0.003, which would mean that we would just be able to detect the predicted effects of a major macroeconomic upheaval with a five percentage point increase in the rate of unemployment and a 5% fall in average income.

It is true that, for an individual, unemployment is associated with very large declines in reported well-being. It is also true that the increase in unemployment after the financial crisis was very large by business-cycle standards. Yet that very large increase affects only one twentieth of the population. In consequence, even major changes in unemployment predict only small changes in average well-being. Detecting such changes presents both statistical and conceptual challenges. A decline in the ladder of 0.06 is (a little) smaller than the standard error in the ladder on the daily sample of 1,000 respondents, and is only three times larger than the standard error from a monthly sample of 30,000 respondents. But 0.06 is an order of magnitude smaller than the mean effect on the ladder of adding political

questions to the beginning of the survey. In a sense, cross-sections of individuals are much more like cross-sections of countries than they are like a time-series for a single country. The differences in incomes across countries are many times larger than the differences in income to be expected over a few years or even over a few decades; the differences in income between the US and China or the US and Togo took hundreds of years to come about. So the fact that the ladder is highly correlated with incomes across countries, or with income and unemployment across individuals, tells us very little about what will happen if we use the measure to track national well-being over time.

6. Conclusions

Perhaps my most surprising substantive finding is how closely well-being tracked the stock market over the 1,000 days of 2008, 2009, and 2010. Yet it is surely wrong that we can abandon income, unemployment, and well-being measures in favor of a stock market index. Most Americans have no direct or indirect financial interest in the stock market, so whatever is the mechanism it cannot work through personal portfolios, or at least not entirely so. My guess is that for much of the financial crisis, the stock market became the most watched indicator, not only of the present, but also of the future. As suggested by Dan McFadden in comments on this paper, both SWB and the stock market were likely responding to the same underlying stream of news, assessing its implications for the future. During the early spring of 2009, with a newly installed administration, and a stock market that was falling even beyond the already precipitous drop at the end of 2008, it was reasonable enough to have dire fears about the future—indeed the stock market itself was reflecting that fear—and those fears were heavily reinforced by media coverage, sending highly correlated interpretations of the events to a large segment of the population. From mid-2008 to mid-2009, the stock market was *the* indicator of the state of the economy, something that would not necessarily be true in more normal times. The pattern in Fig. 10 may simply reflect the peculiarities of this particular recession. In any case, the correlation with the market reinforces that even at best, the well-being measures are not measuring current well-being, in the sense of today's level of real income, but also picking up the fear of the future associated with evolving economic news. For those who own assets to support future consumption, the effect is immediate. For lower income households, who are less likely to own market assets, and whose well-being was if anything more affected by the market, market movements may generate changes in expectations about unemployment. In standard intertemporal models, today's calculation of lifetime utility reflects not only what is happening to today's instantaneous utility, but also the expected utility of future outcomes, predictions of which are changed by today's news. Such effects would be amplified if 'news' utility is important in and of itself, as in Köszegi and Rabin (2009) or Kimball and Willis (2006).

How do measures of self-reported well-being survive the stress-testing provided by the financial crisis? They perform well in that they respond to income,

unemployment, and the stock market in the directions that we would expect. Yet they are more successful as gauges of short-term fear or hope for the future than of the current state of well-being in terms of income or employment. The worry and stress (see Fig. 7) that came with these events is surely real enough, and worth measuring and taking into account in policy.

Even so, there are serious problems in using well-being measures for tracking the performance of the economy over time. They cannot be expected to change much in response to even historically large changes in macroeconomic activity—the predicted (and actual) effects are just too small. Detecting these changes in real time will require very large standing samples and even then, risk being swamped by the much larger short-term effects of day to day events that are sometimes clearly important—a major stock market crash, or events such as 9/11, Metcalfe *et al.* (2011)—but sometimes have only dubious implications for well-being. We also know from the effects of the political questions in the Gallup surveys that the measures are sensitive to the context in which they are asked. In one sense, this is a technical issue that could be dealt with by suitable positioning of the questions, or by the use of buffer questions, though we do not currently know how to do so in a way that would ensure comparability over time or over countries. But the sensitivity induced by question context extends to sensitivity to news and events widely portrayed in the media—for example the large one-day effect of the fall of Lehman Brothers. While it is conceivable that, as is sometimes argued for the stock market, the SWB measures are giving an accurate take on expected future well-being, it seems more plausible that, like the stock market, they have actually very little to do with well-being. As McFadden suggests in comments on this paper, we may be looking at ‘cognitive bubbles’ that are essentially irrelevant for any concept of well-being that we care about. It is an important future task to explore whether or not SWB is excessively sensitive to changes in the market and to other news.

It is possible that the ladder question used in the Gallup surveys is more sensitive to context effects than are other evaluative questions such as questions about satisfaction with life or with one’s standard of living. But this seems most unlikely; questions about the standard of living and about emotion are also affected in the Gallup data, even though they are asked long after the political questions that make people feel worse about their lives.

There are many unresolved challenges before well-being measures become a standard part of macroeconomic monitoring, however useful such measures are in and of themselves. The measures have proved themselves in the cross-section across different groups, for example for looking at the effects of life circumstances, such as ill-health, divorce, or unemployment. They still have a long way to go in establishing themselves as good time-series monitors for the aggregate economy. In a world of bread and circuses, measures like happiness that are sensitive to short-term ephemera, and that are affected more by the arrival of St Valentine’s Day than to a doubling of unemployment, are measures that pick up the circuses but miss the bread.

Acknowledgements

I am grateful to the Gallup Organization for access to their data, to Jim Harter of Gallup for his extensive help in understanding the design of the Gallup Healthways Wellbeing Index survey, and to Sangeeta Agrawal, Jim Harter, and Frank Newport for leading Gallup's work on the adjustments discussed here. I thank Daniel Kahneman, Daniel McFadden, and Norbert Schwarz for extensive help and comments. I also acknowledge helpful comments on an earlier version of the paper from Tony Atkinson, Andrew Clark, David Cutler, Paul Dolan, Marc Fleurbaey, Ori Heffetz, John Helliwell, Scott Kostyshak, David Laibson, Richard Layard, John Muellbauer, Matthew Rabin, and Arthur Stone.

Funding

Gallup Organization; National Institute on Aging of National Institutes of Health (P30 AG024928–06, to Princeton, P30 AG024361, to Princeton, P01 AG05842–14 to the National Bureau of Economic Research, R01 AG049629–01, to National Bureau of Economic Research.)

References

- Agrawal, S. and Harter, J.K. (2011) Context effect on life evaluation items: forming new estimates for data prior to Jan. 6, 2009, Gallup Technical Report, Washington, DC.
- Bishop, G.F. (1987) Context effects on self-perceptions of interest in government and public affairs, Chapter 10, in H-J. Hippler, N. Schwarz, and S. Sudman (eds) *Social Information Processing and Survey Methodology*, Springer-Verlag, New York, 179–99.
- Bricker, J., Bucks, B.K., Kennickell, A., Mach, T.L., and Moore, K. (2011) Drowning or weathering the storm? Changes in family finances from 2007 to 2009, NBER Working Paper No. 16985, NBER, Cambridge, MA.
- Campbell, A. (1981) *The Sense of Well-Being in America*, McGraw Hill, New York.
- Cantril, H. (1965) *The Pattern of Human Concern*, Rutgers University Press, New Brunswick, NJ.
- Carstensen, L., Isaacowitz, D.M., and Charles, S.T. (1999) Taking time seriously: a theory of socioemotional selectivity, *American Psychologist*, 54, 165–81.
- Chakrabarti, R., Lee, D., van der Klaauw, W., and Zafar, B. (2011) Household debt and saving during the 2007 recession, NBER Working Paper No. 16999, NBER, Cambridge, MA.
- Christelis, D., Georgarakos, D., and Jappelli, T. (2011) Wealth shocks, unemployment shocks and consumption in the wake of the great recession, <http://ssrn.com/abstract=17929881>.
- Clark, A.E. and Oswald, A.J. (1994) Unhappiness and unemployment, *Economic Journal*, 104, 648–59.
- Clark, A.E., Knabel, A., and Rätzl, S. (2010) Boon or bane? Others unemployment, well-being and job insecurity, *Labour Economics*, 17, 52–61.
- Deaton, A. (2008) Income, health, and wellbeing around the world: evidence from the Gallup World Poll, *Journal of Economic Perspectives*, 22, 53–72.
- Di Tella, R., MacCulloch, R.J., and Oswald, A.J. (2001) Preferences over inflation and unemployment: evidence from surveys of happiness, *American Economic Review*, 91, 335–41.

- Di Tella, R., MacCulloch, R.J., and Oswald, A.J. (2003) The macroeconomics of happiness, *Review of Economics and Statistics*, 85, 809–27.
- Di Tella, R. and MacCulloch, R. (2009) Happiness, contentment, and other emotions for central banks, Ch. 6, in C.L. Foote, L. Goette, and S. Meier (eds) *Policymaking Insights from Behavioral Economics*, Federal Reserve Bank of Boston, 311–66.
- Graham, C., Eggers, A., and Sukhtankar, S. (2004) Does happiness pay? An exploration based on panel data from Russia, *Journal of Economic Behavior and Organization*, 55, 319–42.
- Hurd, M.D. and Rohwedder, S. (2010) Effects of the financial crisis and Great Recession on American households, NBER Working Paper No. 16407, NBER, Cambridge, MA.
- Kahneman, D. and Deaton, A. (2010) High income improves evaluation of life but not emotional well-being, *PNAS*, 107, 16489–93.
- Kahneman, D., Krueger, A.I., Schkade, D.A., Schwarz, N., and Stone, A.A. (2004) A survey method for characterizing daily life experience: a day reconstruction method, *Science*, 306, 1776–80.
- Kahneman, D. and Riis, J. (2005) Living and thinking about it: two perspectives on life, in F.A. Huppert, N. Baylis, and B. Keverne (eds) *The Science of Well-Being*, Oxford University Press, Oxford, 285–304.
- Kahneman, D., Wakker, P.P., and Sarin, R. (1997) Back to Bentham? Explorations of experienced utility, *Quarterly Journal of Economics*, 112, 376–405.
- Kimball, M. and Willis, R. (2006) Utility and happiness, Department of Economics Working Paper, March 3, University of Michigan, Ann Arbor, MI.
- Kőszegi, B. and Rabin, M. (2009) Reference dependent consumption plans, *American Economic Review*, 99, 909–36.
- Metcalfe, D., Powdthavee, N., and Dolan, P. (2011) Destruction and distress: using a quasi-experiment to show the effects of the September 11 attacks on mental well-being in the United Kingdom, *Economic Journal*, 121, F81–103.
- Micklewright, J. and Schnepf, S.V. (2010) How reliable are income data collected with a single question? *Journal of the Royal Statistical Society Series A*, 173, 409–29.
- Newport, F. (2011) Adjustments to the Gallup-Healthways life evaluation scores, September 8, <http://www.gallup.com/poll/149006/Adjustments-Gallup-Healthways-Life-Evaluation-Scores.aspx>.
- Nussbaum, M. (2008) Who is the happy warrior? Philosophy poses questions to psychology, *The Journal of Legal Studies*, 37, S81–113.
- Schwarz, N. and Strack, F. (1999) Reports of subjective well-being: judgmental processes and their methodological implications, Ch. 4, in D. Kahneman, E. Diener, and N. Schwarz (eds) *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage, New York, 61–84.
- Schwarz, N. and Schuman, H. (1997) Political knowledge, attribution, and inferred interest in politics: the operation of buffer items, *International Journal of Public Opinion Research*, 9, 191–5.
- Sen, A.K. (1985) *Commodities and Capabilities*, Elsevier, Amsterdam.
- Sen, A.K. (1987) *On Ethics and Economics*, Blackwell, Oxford.
- Sen, A.K. (2009) *The Idea of Justice*, Belknap, Cambridge, MA.

Shapiro, M. (2010) The effects of the financial crisis on the well-being of older Americans: evidence from the cognitive economics study University of Michigan, <http://www-personal.umich.edu/~shapiro/papers/CogFinCrisis.pdf>.

Stevenson, B. and Wolfers, J. (2008) Economic growth and subjective well-being: reassessing the Easterlin paradox, *Brookings Papers on Economic Activity*, Spring, 1–87.

Stiglitz, J., Sen, A.K., and Fitoussi, J.-P. (2009) *Report of the Commission on the Measurement of Economic Performance and Social Progress*, available at <http://www.stiglitz-sen-fitoussi.fr/en/index.htm>.

Stone, A.A., Schwartz, J.E., Broderick, J.E., and Deaton, A. (2010) A snapshot of the age distribution of psychological well-being in the United States, *PNAS*, 107, 9985–90.

Winkelman, L. and Winkelman, R. (1998) Why are the unemployed so unhappy? Evidence from panel data, *Economica*, 65, 1–15.

Wolfers, J. (2003) Is business cycle volatility costly? Evidence from surveys of subjective well-being, *International Finance*, 6, 1–26.