National Well-being Policy and a Weighted Approach to Human Feelings

Gus O’Donnell & Andrew J Oswald

October, 2015
Series Number: 1071
ISSN 2059-4283 (online)
ISSN 0083-7350 (print)
Governments are becoming interested in the concept of human well-being and how truly to assess it. As an alternative to traditional economic measures, some nations have begun to collect information on citizens’ happiness, life satisfaction, and other psychological scores. Yet how could such data actually be used? This paper is a cautious attempt to contribute to thinking on that question. It suggests a possible weighting method to calculate first-order changes in society’s well-being, discusses some of the potential principles of democratic ‘well-being policy’, and (as an illustrative example) reports data on how sub-samples of citizens believe feelings might be weighted.

Keywords: Life satisfaction; anxiety; happiness; national well-being; mental health.

JEL codes: I31, I38, Z18.

For helpful discussions, we are grateful to Dan Benjamin, Chris Deeming, Amanda Goodall, Carol Graham, Jan-Emmanuel De Neve, Dan Hamermesh, Peter Hammond (who gave us many penetrating criticisms from an applied-mathematics perspective), John Helliwell, Richard Layard, William Nicholls, Malcolm Oswald (who gave us many penetrating criticisms from a philosopher’s perspective), and Swaran Singh. Neel Sagar provided invaluable research assistance and ideas.
There is concern that traditional economic measures of progress may no longer be suitable for the modern world. This paper discusses the possible use of ‘happiness’ and well-being data as more appropriate measures.

The paper examines a fundamental difficulty in this area. When various measures of human feelings (happiness, satisfaction, stress, worthwhileness of life, and so on) are available, it is not self-evident how such data on different feelings should be treated.

The paper suggests an approach that relies on the democratic idea of asking citizens how much weight they (the citizens) think society should put on each of the possible measures of feelings.

Example numbers for this new kind of weighting approach -- using data on the UK government’s well-being questions -- are discussed.
National Well-being Policy and a Weighted Approach to Human Feelings

The welfare of a nation can ... scarcely be inferred from a measure of national income.

p. 7 Simon Kuznets, 1934.

We may ... reject the old, naive, behaviorism which assumed that it was ... more ‘scientific’ to judge human beings by animal standards. One consequence ... was that the whole notion of purpose and goal was excluded ... simply because one could not ask a white rat about his purposes.

p. 390 Abraham Maslow, 1943.

By mid-twentieth century ... people as a whole were not disease-ridden, and ideas of so-called positive health emerged. This emboldened the WHO to define health in a new way as ‘physical, mental and social well-being, not merely the absence of disease or infirmity’. ... Medicine would then focus on ... moving people toward the favorable end of the health spectrum, as determined subjectively by responses to questions...

p. 347 Lester Breslow, 1972, Dean of the School of Public Health, UCLA.

A unifying theme ... is that the time is ripe ... to shift emphasis from measuring economic production to measuring people’s well-being.


“Resolution 65/309 of July 2011: We invite Member States to pursue the elaboration of additional measures that better capture the importance of the pursuit of happiness and well-being in development with a view to guiding their public policies.”


1. Introduction

Eighty years ago, Simon Kuznets, a pioneer in the measurement of Gross Domestic Product, pointed out that human welfare could not be captured by his measure (Kuznets 1934). Economists have recently begun to take seriously his strictures, and many governments are beginning to attempt to measure ‘human well-being’.

How could such ideas be made practical and how would governments act upon them? This paper is a tentative attempt to explore the principles of what might be called ‘well-being policy’. Such terminology is unconventional. We will think of it as any form of economic and social policy-making in which use is made of survey data on people’s feelings of psychological well-being. It could be thought of, more broadly, as national decision-making that draws upon data on citizens’ reported emotions. As a matter of historical contrast we will have in mind -- although it would be wrong to believe in a rigid dichotomy -- policy-making of a kind more familiar to economists where attention is paid largely to accounting data and objective¹ data (such as statistics on Gross Domestic Product and on unemployment)².

The paper discusses the idea that average well-being scores might be used as a guide to policy. It suggests a way to put weights on different kinds of human feelings.

¹ We are aware, of course, that it is possible to dispute the objectivity of many such kinds of data. Methods for the calculation of the unemployment rate, for instance, are open to the charge that there is no unique way to measure true joblessness. But those issues are not the primary concern of this paper.

² It is sometime thought that the economist Lionel Robbins was against all subjective data, but that is not true, and he was not an extreme behaviourist. “I do not think that it is sensible to restrict our generalizations to observables..” (Robbins 1953, p. 102). He was, however, famously against the idea that people can make legitimate interpersonal comparisons, although his case was based on introspection and not buttressed with formal evidence for such a view.
As a society evolves, it seems natural to allow the concept of success also to evolve. The paper builds on that simple idea. The intention here is not to argue that nations should go over to an exclusive use of ‘happiness’ and subjective well-being data. Nor is its purpose to denigrate the traditionally collected kinds of information. Nor does it argue for an extreme pro-Easterlin or anti-growth position, even though such arguments should, we believe, be taken seriously by economists. The objective is instead to make an effort to say something constructive about how these new kinds of data might be used. Because the terrain is unfamiliar, later ideas should be seen as tentative and exploratory.

Because there is little standard knowledge in this area of policy-making, we attempt below to suggest ideas from first principles -- as they would be conceived, in particular, by an economist. The field is not, of course, the unique preserve of the economist. When the topic is that of human well-being, many research literatures are relevant. Those include thousands of writings by psychologists, epidemiologists, sociologists, psychiatrists, neuroscientists, and others. They also include articles in political science such as important new work by Flavin et al. (2014) in which the authors document evidence of how government intervention can be a source of greater well-being in society.

A particular concern in this paper will be to try to address the following difficult issue:

Consider a world in which survey statistics are being collected by government statisticians on citizens’ feelings of well-being and quality of life. Imagine that these data cover various measures: they record data on different kinds of feelings. If little is known about the form of the social well-being objective function, but it is believed that well-being is an increasing function of certain emotions (say, ‘happiness’) and a declining function of others (say, ‘anxiety’), can anything useful be concluded about how to measure changes in society’s well-being?

This paper suggests that, perhaps surprisingly, something can be concluded – as a first order approximation and if other stringent assumptions can be made. What emerges is one hypothetical approach to the weighting of human feelings. The paper also offers an empirical contribution as well as a conceptual one. It collects and documents evidence on citizens’ views

---

3 Famously, Richard Easterlin (1974, 2013) has provided evidence that nations may not be getting happier as they get richer.

4 For any non-technical reader who finds this mathematical kind of language strange, it boils down to the idea that we might think that happiness is generally a good thing and anxiety is generally a bad thing, and hence we do know something about the ‘direction’ of their effects, even if we know virtually nothing else about a society’s preferences.
about how well-being data might be weighted. We report the results of four small surveys that were conducted for this project. These were of, respectively, (i) economics students, (ii) MBA students, (iii) Amazon Mechanical Turk survey respondents, and (iv) professional non-academic economists. The findings are described in the later figures, Fig 1 – Fig 4. This part of the paper might be seen as building on the spirit of Benjamin et. al. (2014), because here we also, if in a different way, ask citizens for their views on ‘happiness’ data.

The seeds of these ideas in the empirical economics literature go back to articles such as Easterlin (1974) and Oswald (1997), which argue that policy-makers should move away from solely objective measures and incorporate subjective measures of human well-being, and in the psychology literature to a huge corpus of writings, from the 1980s to today, by Edward Diener. In practical terms, a seminal role for policy-makers was played by a report that may be relatively little-known outside the United Kingdom, namely Donovan and Halpern (2002). Much earlier, and again perhaps not widely known, a number of related ideas had been proposed in the field of public health by Breslow (1972).

In this paper, we follow in the same intellectual avenue, and particularly in the footsteps of the now well-known report by the Stiglitz-Sen-Fitoussi Commission. That Commission drew conclusions such as the following: (pp. 10-16):

- Life is now more complex (“The time has come to adapt our system of measurement … to better reflect the structural changes which have characterized the evolution of modern economies.”)

- Services, rather than manufacturing, now dominate the world of work. (“In effect, the growing share of services and the production of increasingly complex products make the measurement of output and economic performance more difficult than in the past.”)

- We, as a society, need to measure well-being per se. (“A… unifying theme of the report, is that the time is ripe for our measurement system to shift emphasis from measuring economic production to measuring people’s well-being.”)

- Inequality itself matters (“Quality-of-life indicators in all the dimensions covered should assess inequalities in a comprehensive way.”)

---

5 It should perhaps be mentioned that one of us was on that Commission, but the group of authors was a large one.
• Official government statistics should blend objective and subjective well-being data. (“Statistical offices should incorporate questions to capture people’s life evaluations, hedonic experiences and priorities in their own survey.”)

• Sustainability must be a criterion (“Sustainability assessment requires a well-identified dashboard of indicators…the components of this dashboard should be … interpretable as variations of some underlying “stocks”. A monetary index of sustainability has its place in such a dashboard.”)

The later analysis fits within a growing research literature written by economists and social scientists (including Benjamin et al. 2012, Booth and van Ours 2008; Carr and Chung 2014; Clark and Oswald 1996, Diener et al. 1999, Di Tella et al. 2001, Easterlin 2003, Frey and Stutzer 2002a, Graham 2011, Layard 2005, Powdthavee 2010, Senik 2004, and Wulfgramm 2014). Although this paper is not designed as a survey of the literature, the appendix briefly provides a guide to further writings in the field. Work at the border between happiness research and policy-making is currently still relatively scarce. While the well-being of people might be believed to matter strongly in itself, it is now thought that it also does so through another channel. This is because there is growing evidence that ‘happier’ workers are more productive (see especially the work of Carol Graham, such as Graham et al. 2004; many papers by the late Michael Argyle and Alice Isen; and research by Bockerman and Ilmakunnas, 2012; Edmans 2012; De Neve and Oswald 2012; and Oswald, Proto and Sgroi 2015). Well-being may therefore also be of indirect interest to policy-makers and societies.

2. Conceptual Issues

By building upon the work of Kuznets, economists have for a long time thought about, and measured empirically, the concept of Gross Domestic Product. All industrialized nations now calculate estimates of their real GDP. Those estimates are regularly publicized in newspapers and other media. One way to think of the definition of real GDP is as a measure that uses current output flows multiplied by historical prices in order to generate an implied national level of income. Might some formally similar method be applicable, perhaps in a different form, to the study of national well-being? This section discusses that possibility (although we wish to stress the fact, as also mentioned to us by a referee, that many of the ideas in the paper will go through without the need for algebra). It lays out a simple way to use current data on human feelings multiplied by historical social-importance weights to generate an implied total well-
being level. As is true of GDP data, much of this paper’s focus will be on how to calculate a rate of change.

It is helpful to begin with a mathematical point. Here we essentially follow the approach of Equation 2 of Benjamin et al. (2014), and indirectly also some of the innovative modeling ideas in Kimball and Willis (2006).

Imagine a person who cares about just two kinds of feelings (or ‘well-being variables’). Call them p and s. Assume that there is some form of utility or objective function that depends on those two variables. Let that person’s utility be described by a function

\[ W = W(p, s). \] (1)

Imagine also that these values of p and s change through time. In that case, although it might be considered unconventional by economists, the approximate improvement in the person’s overall well-being over time could then be thought of as:

\[ \Delta W = W_p(p_0, s_0) \Delta p + W_s(p_0, s_0) \Delta s + \text{higher order terms} \] (2)

The change in the person’s well-being = (the weight on emotion p multiplied by the change in emotion p) plus (the weight on emotion s multiplied by the change in emotion s) + higher order terms

where the delta symbol, \( \Delta \), measures the size of a change, and subscripts denote partial derivatives. Intuitively, any curve can be approximated locally by a linear gradient. That approximation will be a good one for small alterations in variables p and s; it will usually be a poor one for large alterations. Under certain conditions, this elementary mathematical idea can be used empirically. To do so, however, it will be necessary for real-world statisticians and policy-makers that:

(i) it is feasible to measure the changes in p and s;

(ii) it is possible independently to approximate -- in some way to be determined -- the size of the partial derivatives of \( W(..) \), which could be thought of as the marginal importance weights on p and s respectively;

(iii) we are willing to accept that measurement is over periods of time appropriately short enough that a linear approximation is sufficient, and thus that it is reasonable that higher order terms can be neglected (this is because the methodology rests on a first-order approximation).

\[ \text{If a clarifying example is needed, think of them as pleasure and stress.} \]
This approach can be generalized. Consider a country, such as the UK, in which four measures of well-being are being collected by government statisticians; this is for illustration (the method, of course, will not require four). Give these four variables the names ‘happiness’, ‘satisfaction’, ‘anxiety’, and ‘worthwhileness-of-life’, and denote them by the symbols h, s, a, and w. These are, self-evidently, reported levels of emotions or feelings.

What, then, might be done with such survey data? Assume it is possible to think of a country as being like a single representative citizen. This is not an attractive assumption (though we will return to the issue briefly later). It is particularly unattractive in ethical matters. But it allows us to make a start on how to think about well-being data as gathered in surveys. Moreover, most of the traditional measures of policy success have also routinely focused on the means of variables – the level of GDP per capita, the amount of unemployment, the mortality rate, and so on. Hence the implicit assumption of a representative person might be argued to lie behind much of the economics and politics of public discussion. Although there is a case for making the unit of analysis the family, in our later empirical work we will work, for simplicity, at the level of the individual.

Imagine that the representative citizen in the economy has well-defined objectives in life. He or she chooses various kinds of actions, z. These can include economic actions, like purchasing decisions. They also include social actions, like decisions about whether to marry and with whom to be friends. Assume that the citizen makes these decisions, z, in some way that he or she believes to be optimal. The choices thus help to determine the person’s happiness, satisfaction, anxiety, and worthwhileness-of-life. Assume people know their four values; in other words, they know how they feel. But when asked by an interviewer, they cannot report the literal numbers in the way they could their age or income. Instead, they must attempt to report those values on scales, and they do this using what might be called a reporting function. Some assumption has to be made about this reporting function.

Since 2011, four questions of this type have been asked annually by the UK’s Office of National Statistics. There is the issue of the economist’s concern over interpersonal comparisons. However, in an interesting paper, Binmore (2008) has argued: “I think John Harsanyi’s theory of interpersonal comparisons of utility defuses all these concerns by providing a clear and relatively uncontroversial approach to the subject.” Harsanyi’s view is, approximately, that the assumption of empathetic preferences can solve the famous problem: people can put themselves in the skins of others. See also Hammond’s (1991) arguments.

Later in the paper, we will look at an average of changes rather than the change in the average, but as a first-order approach these two are approximately the same.

Benjamin et al. (2014) make a similar point.

To use the term suggested in Oswald (2008). That paper offers some evidence that in the case of people’s height the reporting function is approximately linear.
Let us assume that when a human being has an emotion, the person experiences that as a genuine feeling inside themselves. If asked how much of the feeling they have, however, it is not straightforward (perhaps because there are no natural units they can use) for a person to communicate in an exact way. Assume, nevertheless, that the person can report some kind of number, and that the reported amount of emotion increases when the underlying actual emotion increases. When someone is happier they tend to say they are happier.

Then it is possible to make analytical progress. Because the reporting function is monotonic, it is possible to define its inverse. Let that be a function \( i(.) \). Assume that the same reporting function is used for every emotion. It is not easy to know intuitively whether, for human beings, that is an extreme assumption; it might be. Actual emotion can now implicitly be read off from reported emotions. Hence in the case of reported happiness \( h \), reported satisfaction \( s \), reported anxiety \( a \), and reported worthwhileness \( w \), this implies that the underlying numbers for these four emotions are \( i(h) \), \( i(s) \), \( i(a) \), and \( i(w) \). These are the implied or ‘true’ levels of happiness, satisfaction, anxiety, and worthwhileness. Assume that all functions are smooth. This is extreme. In particular, most well-being survey questions are framed as a set of discrete choices (very satisfied, fairly satisfied, etc), which precludes differentiability. Nevertheless, this is the assumption we make. An alternative approach -- we owe this point to Peter Hammond -- might be to give up the idea that we are measuring the well-being of any one representative individual. Instead, it might be assumed that the mean well-being of a population of individuals is a smooth function of the population mean levels of the observed values of \( h \), \( s \), \( a \), \( w \), rather like a conditional mean in statistics.

Consider a starting point called time period 0. Government statisticians initially collect data on well-being: in surveys they ask for the four reported values \( h_0 \), \( s_0 \), \( a_0 \) and \( w_0 \). The statistical offices then follow the economy through time. They are unable to measure well-being itself, because the full form of \( W(\ldots) \) is not known. But assume it is possible in continuing surveys to measure changes in the component parts, namely, in each of the values of \( h \), \( s \), \( a \), and \( w \). This is done by collecting the future answers to each of the four questions.

Well-being can then be written out, where \( i(.) \) is the inverse of the reporting function, as a function

\[
W = f(i(h), i(s), i(a), i(w)).
\]
Even when there is limited knowledge about the \( f(\ldots) \) function, some conclusions can be drawn. The reason is that, by standard calculus, a local linear approximation is, at \( W(h, s, a, w) \), as follows:

\[
L(h, s, a, w) = f(i(h_0), i(s_0), i(a_0), i(w_0)) + f_h(i(h_0), i(s_0), i(a_0), i(w_0))i'(h)(h - h_0) \\
+ f_s(i(h_0), i(s_0), i(a_0), i(w_0))i'(s)(s - s_0) \\
- f_a(i(h_0), i(s_0), i(a_0), i(w_0))i'(a)(a - a_0) + f_w(i(h_0), i(s_0), i(a_0), i(w_0))i'(w)(w - w_0)
\]

(4)

where the anxiety component of the equation has been written with a negative sign because anxiety is taken here to be a bad rather than a good. To first order, the \( i(.) \) function is linear, so for simplicity we will denote its gradient as \( K \) and take it near to the front of the long expression above.

By rearranging the above equation, the first-order change in well-being can be written in a less daunting way. Converting the equation to a change, by taking the first term on the right hand side of equation (4) over to the left hand side, and with simplification of terms, it takes the more compressed form:

\[
The change in society’s well-being \equiv K\{(h - h_0)f_h + (s - s_0)f_s - (a - a_0)f_a + (w - w_0)f_w\}
\]

(5)

where the partial derivatives of \( f(\ldots) \) are to be evaluated at their starting values (namely, the variables’ 0 values).

Although this formula is abstract, and will be unappealing to some philosophers, including Nussbaum and Sen (1993), it is, in two senses, meant to be utilitarian. The weighting formula potentially allows a useful empirical application, because, given this structure, we might be able to use various kinds of survey data to fill in the different elements in the equation. Written out in English, the change in society’s well-being is:

\[
\equiv K (\text{the increase in measured happiness multiplied by the marginal social weight on happiness}) \\
+ (\text{the increase in measured satisfaction multiplied by the marginal social weight on satisfaction}) \\
- (\text{the increase in measured anxiety multiplied by the marginal social weight on anxiety}) \\
+ (\text{the increase in measured worthwhileness multiplied by the marginal social value of worthwhileness}).
\]
If viewed from a Benthamite utilitarian tradition, the above form might be seen as a straightforward and perhaps even obvious one. Each positive feeling is weighted; each negative feeling is weighted. The outcomes are then summed.

3. How Could Marginal Social Weights on ‘Feelings’ be Calculated?

The next question is a practical one. How could the weighting values required in the above italics equation ever be calculated? In other words, how might this equation -- rather than the data collection itself -- be implemented by national statisticians? Given the representative-citizen assumption, the difficulty is principally that of deciding how the social-importance weights on each ‘feeling’ are to be decided.

There is no orthodox way to put marginal social values on feelings such as happiness, life satisfaction, anxiety, and the worthwhileness of life. Yet it seems particularly unattractive to make a mechanical arithmetical assumption that all four should be weighted evenly. Our approach to the issue, therefore, is to build toward a way of measuring the social weights by starting with a number of principles that might be viewed as reasonable. It should be emphasized that we shall have to maintain an essentially utilitarian philosophical position.

We consider four ideas.

Idea 1. People could be allowed to say how they are feeling, and have their emotions and preferences be taken seriously.

The conception here, which underlies the rationale for the ONS survey data that are already being collected, is that individuals might be seen as entitled to give their own judgments, and assess their own reported feelings, about their lives.\(^{12}\) The citizens may be an appropriate, and perhaps even the best, judge of themselves. Hence people’s answers in well-being surveys to questions such as ‘how satisfied are you?’ might be viewed as offering serious data for use by government statisticians and policy-makers. Such survey answers would then provide valuable raw material for social decision-making. For the survey answers to be reliable in the way required, it will be necessary that people answer as they currently feel rather than as they wish to feel. The wording of the questions gives some reason to be optimistic about that.

Idea 2. Citizens could be allowed -- possibly after discussion in ‘citizen juries’ -- to choose their own desired weights on the elements in a government maximand.

\(^{12}\) A fair question would be: how about individuals with extreme mental ill-health such as schizophrenia sufferers? It will be necessary to assume that they are not so numerous as greatly to distort the mean values.
In the design of policies, ethical choices also have to be made. Idea 2 captures the idea that those choices might be put in the hands of individuals in the society. One alternative would be to allow politicians to make the choices; a third possibility would be to let experts make the choices (see Anderson 2010 for a discussion of actual policy-making). A softened ground, from this perspective, would be explicitly to use citizen juries. Standing back, Idea 2 could be seen as in the spirit of a modern society. We are conscious that some writers, as explained as Fischer (2000), would object to any simple version of Idea 2 on the grounds that citizens are not qualified or able to choose optimally.

Idea 3. Issues could be decided by numerical strength of voting.

It is likely that in any society there is an array of opinions about the desirable social weights on objectives of happiness, life satisfaction, anxiety, and so on. To choose among them, one approach would be to rely on the notion of democratic decision-making (though we recognize that only rare nations, like Switzerland, use this extensively on individual political decisions). Idea 3 could be viewed as the notion of one-person-one-vote.

Idea 4. When people’s preferences alter, government objectives could be allowed to alter.

This might seem obvious. However, a simple GDP objective would not satisfy this fourth idea. One purpose in analytical public policy could be said to be to find a social objective that is more flexible than that of, for example, the maximization of Gross Domestic Product. As an industrialized society develops it is perhaps natural to allow the concept of success also to develop. The statement in Idea 4 is redolent of the famous paper by Maslow (1943) about the so-called hierarchy of human needs, which, remarkably, was published at the height of the Second World War when self-actualization itself was probably not on the mind of many of the world’s citizens.

4. How Could Data on Well-being Weights and Reported Feelings Be Combined?


14 It and Idea 3 are also reminiscent of Dahl (1971), page 1: “A key characteristic of a democracy is the continuing responsiveness of the government to the preferences of its citizens, considered as political equals.”
The ideas from the previous two sections could in principle be blended. It would be necessary to obtain proxies for the terms in the (earlier) equation for the size of the increase in well-being

\[
\text{Change in well-being} \equiv K\{(h - h_0)f_h + (s - s_0)f_s - (a - a_0)f_a + (w - w_0)f_w\} \quad (6)
\]

which means that we now have to find empirical measures for each of the social weight terms (the f terms) and for each of the changes in the well-being feelings such as the alteration in happiness \((h - h_0)\) and the alteration in satisfaction \((s - s_0)\). The negative on the anxiety term is because it can be considered, in the economist’s jargon, a bad rather than a good.  

The changes in the four well-being elements \((h, s, a, w)\) can in principle be measured in regular population surveys. This is currently being done each year, for example, in the United Kingdom. In the Annual Population Surveys, the UK Office of National Statistics asks:

- Overall, how satisfied are you with your life nowadays? On a 11 point scale where 0 is not at all satisfied and 10 is completely satisfied.
- Overall, to what extent do you feel that the things you do in your life are worthwhile? On a 11 point scale where 0 is not at all worthwhile and 10 is completely worthwhile.
- Overall, how happy did you feel yesterday? On an 11 point scale where 0 is not at all happy and 10 is completely happy.
- On a scale where 0 is not at all anxious and 10 is completely anxious, overall, how anxious did you feel yesterday?

By looking at the mean values of these, measured through time, it is possible to calculate the four changes defined, respectively, by the terms \(h - h_0, s - s_0, w - w_0,\) and \(a - a_0\). These measurements provide some of the necessary input to the overall change in well-being equation.

The marginal weights are less easy to calculate. They are

- \(f_h\) = the marginal weight on happiness
- \(f_s\) = the marginal weight on satisfaction
- \(f_w\) = the marginal weight on worthwhileness
- \(f_a\) = the marginal weight on anxiety.

One approach to the difficult problem of how to assess these is suggested by the ideas outlined in the previous section. We could ask people for their individual views on the appropriate weights (on happiness, on satisfaction, and so on), and then take the mean weight on

\[15\] Dan Hamermesh has reminded us that in principle this would have to come about through some form of voting process.
each of the four variables. As an empirical illustration, this is what we have done for a number of groups of individuals.

As explained more fully in Table 1, we asked citizen samples the following questions:

“We are interested in people’s opinions on the quality of a society.

The UK government is collecting information on the four well-being questions lower down the page. These measure happiness, satisfaction with life, how worthwhile life is, and people’s anxiety. We would like to know your view on the relative importance of these for assessing how well a society is doing.

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures below? [for example, if you believe all four are equally important, you would allocate 25% to each of the four measures]”

Our samples of individuals were not meant to be statistically representative of the national population. The data given below are merely designed to illustrate the method and to provide some idea of how certain groups choose to answer. First, we asked economics students. These were a sample of 76 students who attended an economics summer school in 2014. Second, we asked MBA students. These were 206 students at INSEAD Business School in France (the questions were given in English). Third, using web-based methods, we asked a wider group of citizens. These were 306 people in an Amazon Mechanical Turk survey. Fourth, we asked professional non-academic economists, all of whom work in the UK. The sample size here was 52 people.

Our sample size across these categories is approximately 650. To encourage a high response rate, we did not ask questions about people’s personal characteristics.

The answers are portrayed in the histograms of Figures 1, 2, 3, and 4. Perhaps the most interesting finding to emerge from these is that all of the variables are given a considerable share of the possible total of 100 points to be allocated. Nevertheless, in each case, the anxiety question garnered the smallest mean weight, that is, the smallest value of $f_a$. The social weight put on this variable did not exceed 20% in any of our four small surveys. The marginal value of life satisfaction, $f_a$, gets the highest weight in each of the three samples of economists, economics students, and business students. However, the findings in the MTurk survey produce a pattern that is not as similar to the others. Happiness is given the greatest weight, followed by life satisfaction. Because data are not available on the survey respondent’s detailed characteristics, and because the numbers are only illustrative of the method, formal statistical tests are not
presented. Given sufficient resources, however, it would be straightforward to collect a large and statistically representative sample of opinions on the four marginal-importance weights.

5. Discussion

The above ideas will look unusual to an economist raised on a combination of utility theory and the notion that one should study people’s actions. Nevertheless, we think it may have been a mistake for researchers and economic policy-makers to have so eschewed the use of data on reported feelings. In government policy, there is a limited budget. One day it may come to be formally accepted in industrialized society that £X spent on a policy to achieve a well-being gain of Y is an inefficient policy if the same gain can be achieved by a different policy costing a fraction of £X. If this approach were taken to its logical conclusion, the next public spending round in a nation like the UK would consist of the different government departments presenting their sets of policies with estimated costs and well-being benefits and then choices being made in an attempt to maximize well-being. Society is currently some way from such an outcome. How would one measure the impact of extra defence spending in well-being? In practice, departments could be asked to submit budgets for “must do” activities and then list a set of discretionary spending with estimates of their well-being impacts. Ideally, departments could join up to present a suggested policy e.g. improved education in prison, which would have well-being and financial impacts in many departments.

In the world of affairs, rather than in an economist’s equations, what might be the key obstacles to making such a process work?

First, agreement would be needed on a definition of well-being. In the policy field, the estimated impact on GDP is regularly employed as a success measure (even though that is to ignore the strictures of researchers such as Richard Easterlin who have provided evidence that humans care about relative rather than absolute income). Many understand that GDP is a flawed measure; but it is still used pervasively. Having one summary measure is powerful and allows compatibility with other, possibly competing, policy options. Our earlier weighted-well-being approach -- though meant here only as an illustration -- offers one style of alternative. Others are likely to be suggested by future researchers.

Part of the problem is presentational: GDP seems to policy-makers like a robust, objective measure and it has been measured for a long time in many different countries, so cross-section and time series comparisons can be made. Can the economics and social science professions
agree on a definition of well-being that is equivalently robust, albeit subjective, and a measure that can be calculated consistently across time and countries? Is it intrinsically harder to get acceptance of a subjective measure like well-being than an ‘objective’ one like GDP? Do cultural and linguistic factors matter more for well-being than measures like GDP and therefore make it hard to make cross-country comparisons? More knowledge on these questions is needed.

The ultimate desire is for a policymaker to be able to calibrate different policy options in terms of their likely impact on human well-being – and to do so in both the short and long run. This will require models which explain the determinants of well-being. Once the determinants have been established there will be a desire for new ideas to explore how well-being can be sustainably enhanced and for experiments to provide robust evidence of the effectiveness of different interventions.

A continuing theme is likely to be how to put monetary valuations on factors that influence well-being but that do not come with easily read financial values. A method for doing this was developed by Clark and Oswald (2002)\textsuperscript{16}, and has since been used in the well-being research literature, especially in the environmental economics and health economics literatures (see Appendix 1). It is based on the estimation of ‘happiness’ equations in which income enters positively and other variables of interest enter statistically significantly. That approach offers a way to calculate a form of marginal rate of substitution. The implicit values of those variables can then, in principle, be calculated. To date, the methodology for doing this remains worryingly crude. Four concerns stand out. First, in most studies, there is no attempt to adjust for the endogeneity of income, so the size of the estimated coefficient on income in the well-being equation is suspect. Second, too little distinction has so far been paid to the obvious fact that marginal values and average values are routinely not the same. Third, in some cases the independent variables of interest (like marriage) are endogenous while others (like certain illnesses or the death of a child) can reasonably be taken as close to exogenous, and the implications of that distinction remain to be understood. Fourth, as in Powdthavee and van den Berg (2011), there is some indication that the size of implied happiness values varies with the exact measure chosen as the dependent variable in the well-being regression equation. These

\textsuperscript{16} This method for putting monetary values on the factors that influence happiness is actually older. It was first proposed in papers by Blanchflower, Clark and Oswald at a 1993 conference on the economics of happiness at the London School of Economics (a conference to which few people came). One of those papers, which was eventually published 13 years later as Blanchflower and Oswald (2004), was originally coauthored with the psychologist Peter Warr of Sheffield University, although Peter Warr voluntarily withdrew from the project. That paper estimated the happiness value of marriage, for example, at approximately 100,000 US dollars per year.
four problems are not trivial. This research area is thus likely to be one that sees much activity in
the future.

There also exist technical and philosophical obstacles to implementing a well-being
approach in policy. One is the choice of discount rate. How should we trade off well-being today
against the same amount tomorrow? If, as some evidence suggests (for example, Blanchflower
and Oswald 2004), the average well-being level in a country such as the UK is broadly constant
over many decades, should we use a zero discount rate? Next, should a nation care about the
total level of well-being or well-being per capita? Many economists, and in the UK a Lords
Committee on the impact of migration said so, would go for a per-capita formulation. But, in
country A, do we really not care about well-being levels of those currently outside A? What
about former A residents who are currently living aboard? Is an aid programme really only about
raising A residents’ well-being? Third, do we care about the distribution of well-being?
Arguably a policy that raises the average level of well-being by X spread evenly across the
population is not as good as one that achieves the same average results but has all its effect on
those with below average initial well-being. [This also has spatial implications. For the UK, it is
known that GDP per capita is lower in Wales than England, although, as Richard Layard has
pointed out to us, Wales scores well on a number of well-being measures. If we were equally
sure that well-being was distributed in the same way, would governments be prepared to adopt
policies that reduced well-being inequalities in the two countries? Could we be reasonably
confident that policies would deliver this objective, as for many years policies have attempted to
reduce geographical income inequalities with mixed results?]

How do we handle the questions of different aspirations? If region A has lower well-being
but higher average income because the inhabitants of A are less concerned about money and
have traded off money for, say, lower working hours or commuting times, should governments
get involved? Many economists might say no, but this perhaps implies living with large income
inequalities. And should the goal be to equalize well-being? Some researchers claim that people
may choose consciously or unconsciously not to maximize their well-being. It is possible to
think of reasons why this might occur: inertia, feelings of obligations, for example remaining
close to ageing parents, and people who believe their purpose in life is something different e.g.
certain religious groups. Benjamin et al. (2014) provides some evidence of the goal of family
happiness as a driving force in individuals’ decision-making.
Finally, can we address the issue of hedonic adaptation? On average, people who lose limbs eventually regain some of their prior level of well-being (for example, Lucas 2005; Oswald and Powdthavee 2008 shows that disability adaptation is not 100% and the extent of the recovery depends on the seriousness of disability). Does this mean that policy only needs to focus on minimising the well-being loss during the transition period? Should we be attempting to influence what constitutes higher well-being by trying to adapt preferences – for example making it fashionable to drive more energy efficient cars? Governments frequently take ‘tastes’ as given, but actually through the education system, via the setting of cultural norms and through legislation, they play a role in determining these tastes. At what point does this process become improper?

Since governments have already influenced through early education how people perceive higher well-being, it may make sense to allow adults to determine for governments how they should measure it. If the public believe happiness is what matters, having been educated in a system greatly influenced by government, then there may be a lot to be said for governments taking this weighting seriously when defining well-being. This notion is at the centre of the weighting proposals in the paper’s earlier section. Yet much remains to be worked out.

Two other difficult issues, pointed out to us by Dan Benjamin, include the following. Suppose the UK ONS survey questions on “happiness” and “life satisfaction” partly measure the same thing. Then constructing a well-being index using both of those questions risks double-counting the component of well-being that is measured by both. Such a potential weakness presumably applies to all emotions solicited in well-being surveys, so that is an open unsolved problem. Moreover, it might be wondered whether we are justified in thinking of the importance weights in our empirical section as actual marginal utilities. The justification, in so far as there is one, would have to be that in the survey the respondents are asked where they believe effort should be put.

6. Conclusion
Human beings are creatures of emotion and feeling. To understand the quality of their lives, therefore, it might be thought natural to try to measure not only their actions, as economists have traditionally done, but also their feelings. Such an instinct may lie behind recent academic interest in the economics of well-being and happiness. Another potential explanation for that interest may be growing belief in the ideas of Richard Easterlin, who, as early as 1974, questioned the value of measuring human progress by using data on economic growth. Another may be the influence -- after a century of remarkable economic growth -- of diminishing marginal utility of income. Maslow (1943) might have predicted that human beings today would care more, at the margin, about non-pecuniary variables. Another might be rising concerns about global warming and the likely consequences of ever-increasing carbon emissions.

Currently it is not known how best to combine objective and subjective data. In the research avenue that has come to be known as the economics of happiness (or of well-being), random samples of individuals are asked in surveys how they rate the quality of their lives. Researchers take the answers and correlate those with the observable features of those people’s lives – their income, their marital status, their age, their gender, and so on. Such research, whatever its limitations, has one strength that may not be completely recognized by all economists. People are not asked how much one thing makes them happy compared to another. Deeply complicated cause-and-effect survey questions are thereby largely eschewed, and that is an advantage. In this way, it is possible for researchers to work out the patterns in a single variable like reported happiness, and how they relate statistically to influences such as earnings and divorce and bereavement, without relying on the sampled individuals to tell survey teams about what those individuals think makes them happy and by exactly how much. This paper does not have that advantage. Instead, it works with a form of nationally weighted feelings. The reason is that the approach in the paper -- intentionally of course -- has to allow for multiple

---

17 It is possible that the discipline of economics can make progress by using some blend of the two kinds of data. If anyone reads this article in 50 years from now, we will be well pleased if their reaction is to think this sentiment banal or obvious. As a matter of historical record, or for any such hypothetical future reader, the two authors of this paper have grown up (one immersed for nearly 40 years in economic policy-making, one immersed for nearly 40 years in academic economic research) in an environment where such sentiments would be thought somewhere between mildly muddle-headed and plain bad.

18 As one marker of that interest, an electronic search on the Web of Science indicates that since 2008 the single most-cited paper in mainstream economics journals is Clark et al (2008). That is out of the approximately 120,000 journal articles in all of economics that have been published since 2008.

19 The literature on the economics of well-being does not ask “how much unhappiness did you suffer from being made unemployed?” The literature might instead be seen as methodologically akin to parts of epidemiology. Rather than ask people whether they believe that smoking makes them less healthy, and fruit and vegetables makes them healthier, epidemiologists study ex post patterns and calculate the implied linkages.
measures of people’s emotions and for normative concerns. The simplicity of the happiness-economics approach is then impossible. Somehow or other, a mixture of human feelings must then be weighted in government policy-making.

This paper has discussed the conceptual and empirical foundations of well-being policy. It has also reported data on four surveys -- of a small and deliberately illustrative kind -- that we carried out for the project. In those surveys, people were asked how they would wish to weight the importance of the four well-being questions currently being asked by UK government’s Office of National Statistics. To the best of our knowledge, these are the first empirical results of their kind.

Three main conclusions have emerged. First, we have argued that it is possible to think of a potentially workable empirical approach to the measurement of well-being improvements in a society. To do so, we have had to make a number of strong assumptions, but given those assumptions the change in well-being can be written as the change in well-being \( \equiv K \times \sum \text{marginal social weights on subjective feelings } \times \text{measured changes in subjective feelings} \) where, first, the weights are decided by asking in separate surveys how citizens believe different feelings should be weighted, and, second, the changes in subjective feelings are assessed by regular random-sample surveys of the population. A weighted average of this type has informal appeal and some readers may feel it is not necessary to think algebraically about the problem in the way done in earlier sections. Yet it is easy to think of imperfections in the approach. Nevertheless, perhaps this paper might act as a spur to better methods. Here the emphasis has been on the four questions being asked by the UK government, but the first-order method could be applied to any nation in which any form of well-being or feelings questions are being asked. Second, our four small surveys offer an illustration of what actually happens when citizens are asked how they wish to weight answers to different well-being questions. Figures 1 to 4 reveal that all four of the ONS questions are given a reasonably-sized amount of importance. To implement the approach properly, it would be necessary to have larger random samples of the population rather than our heavily selective samples. Third, we have reviewed some of the

---

20 Our weighting method could be applied to many other kinds of well-being survey questions, including, for example, those for ‘positive affect’; we thank John Helliwell for helpful discussions on this issue. The ONS questions are used here as one illustration of the approach.

21 It should perhaps also be mentioned that the nation of Bhutan has for many decades promoted such ideas.

22 To summarize some of the difficulties with our weighted approach above: we have assumed (i) a representative citizen, (ii) smooth functions, (iii) the same reporting function for each kind of emotion, and (iv) that changes are measured suitably to first-order only. All these assumptions are violated, or probably violated, in reality. How large a net error is produced by such imperfections is currently unknown.
difficulties with, and current research applications of, well-being policy, and said something about the unsolved questions. Perhaps it should also be recorded that we are acutely aware that real-world politics is moulded partly by ideology, power and constituencies; the current analysis necessarily contributes nothing on those matters.

Where might well-being policy go in the future? The demand from politicians for policies that raise well-being is high and is likely to stay high during the so-called period of “austerity” and into the future. Governments of all persuasions want to keep up, or even improve, public satisfaction with public services while reducing the money spent on them. Traditional solutions focus on reducing costs, for example, by lowering procurement costs. More innovative approaches focus on trying to raise public well-being, for example by preventing problems before they start to cost money. In the health service, the adoption of behavioural approaches to encourage healthier living is in its infancy. Such programmes have tiny funding compared to the amounts spent on drugs and hospitals. A well-being focus might be likely to shift resources from physical to mental health as the well-being “bang per buck” is thought to be higher in the latter. The use of quality adjusted life years (QALYs) by the UK’s National Institute for Clinical Excellence (NICE) was an early example of the use of a well-being measure to guide policy and provides useful lessons. For example, patience will be necessary: the time gap between acceptance of QALYs as the unit of measurements and their use in policy was many years. Part of the acceptance of QALYs was their approval by the medical professionals. All now accept that QALYs are not perfect – and one might prefer to use well-being adjusted life years or WELBYs – but they are better than purely fiscal alternatives. Another difficult and here unresolved issue is how, as mentioned by Graham (2013), to treat inequality; like those who discuss the pursuit of GDP, we have said almost nothing beyond the study of mean values of well-being.

More economics research will be needed on the foundations of well-being policy. We are conscious that it would be desirable to understand, and be able to say more about, the distribution of well-being (though see Stevenson and Wolfers 2008a for an innovative early approach, and Clark et al. 2012 for the recent and intriguing finding that happiness inequality is narrowing). The ideas in this paper will eventually somehow have to be fused, with those from other researchers, into a more coherent whole. That is likely to require a cross-disciplinary approach and perhaps larger research teams. Whatever form such a research effort takes, we believe these issues deserve future attention.
References


National Academies Press, Washington DC.


Table 1: The Questionnaire we Used

WHERE SHOULD SOCIETY FOCUS ITS EFFORTS TO IMPROVE WELL-BEING?

We are interested in people’s opinions on the quality of a society.

The UK government is collecting information on the four well-being questions lower down the page. These measure happiness, satisfaction with life, how worthwhile life is, and people’s anxiety. We would like to know your view on the relative importance of these for assessing how well a society is doing.

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures below? [for example, if you believe all four are equally important, you would allocate 25% to each of the four measures]

Happiness – “Overall, how happy did you feel yesterday?”: Personally I would allocate ...... % of my efforts to improving this. Please write above the dots or circle a number

0% 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100%

Satisfaction – “Overall, how satisfied are you with your life nowadays?”: Personally I would allocate ...... % of my efforts to improving this. Please write above the dots or circle a number

0% 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100%

Worthwhile – “Overall, to what extent do you feel that your life is worthwhile?”: Personally I would allocate ...... % of my efforts to improving this. Please write above the dots or circle a number

0% 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100%

Anxiety – “On a scale where nought is “Not at all anxious” and ten is “Completely anxious”. Personally I would allocate ...... % of my efforts to improving this. Please write above the dots or circle a number

0% 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100%

PLEASE REMEMBER THAT YOUR FOUR CHOSEN NUMBERS SHOULD ADD UP TO 100%.

THANK YOU FOR YOUR VIEWS.
Figure 1: The Means of the Importance Weights Chosen by the Respondents in Survey 1 (of economics students)

Note on how to interpret the figure: The vertical bars give the means of the proportions of survey respondents who favoured happiness, satisfied, worthwhileness of life, and anxiety. Thus in this sample the respondents believed that the greatest weight should be put on the objective of life satisfaction. On average, 31% of their possible 100% of points were allocated to that. By contrast, 19% of points were allocated to the goal of (low) anxiety. WESS stands for Warwick Economics Summer School.

Sample size N=76

QUESTION “The UK government is collecting information on… happiness, satisfaction with life, how worthwhile life is, and people’s anxiety. We would like to know your view on the relative importance of these for assessing how well a society is doing.

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures...”
Sample size N=206

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures …
Figure 3: The Means of the Importance Weights Chosen by Respondents in Survey 3 (of Amazon Mechanical Turk respondents)

Sample size N=306

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures...
Figure 4: The Means of the Importance Weights Chosen by Respondents in Survey 4 (of professional economists)

Sample size N=52

We would like you to imagine that you have 100 points to allocate as an indication of the importance of measures of well-being. How would you personally allocate the 100 points across the four measures...
Appendix 1: A guide to the modern literature

Space precludes a proper description of the burgeoning literature. As a guide to further reading, however, a number of modern reviews of the literature have been published. They include Blanchflower and Oswald (2011), Clark et al. (2008), Diener (2013), Di Tella and MacCulloch (2006), Dolan and Metcalfe (2012), Fleurbaey (2009), Graham (2010, 2011), Helliwell and Wang (2013), Jorm and Ryan (2014), Layard (2005), OECD (2011), Powdthavee (2010), and Stone and Mackie (2013). The paper by Jorm and Ryan is likely to be especially useful to those from outside the discipline of economics. Bache and Reardon (2013) discusses the politics of well-being.

Innovative recent papers on public policy approaches using well-being data include Deeming (2013), Delbosc (2012), Dolan et al. (2010), Hamermesh et al. (2014), Hicks et al. (2013), O’Donnell et al. (2014), and Oishi et al. (2012). The first of these contains a detailed discussion of modern UK well-being data.

An especially interesting literature has now sprung up at the border between environmental economics and the economics of well-being. This uses ‘happiness equations’ in various creative ways to put values on green variables; in some way this literature has led the way on possible policy uses for the economics of happiness. The writings include Abdallah et al. (2008), Ferreira et al. (2013), Ferrer-i-Carbonell and Gowdy (2007), Frey et al. (2010), Gowdy (2005), Kubiszewski et al. (2013), Levinson (2012), Luechinger (2009), MacKerron and Mourato (2009), Vemuri and Costanza (2006), Welsch (2009), Welsch and Kuehling (2009), and White et al. (2013).

Good discussions of methodological issues in this field, which are plentiful and important, are available in Exton (2014) and ONS (2011). On so-called vignettes, Kapteyn et al. (2013) is particularly interesting.

Some of the latest research has constructed RCT randomized controlled trials. These examine the well-being effects of factors such as health insurance cover (Baicker et al. 2013), work incentives (Dorsett and Oswald 2013), and mobility vouchers for disadvantaged families (Ludwig et al. 2012). Although expensive, such experiments may be a marker of research trends. They seem likely to be important in the future.


Modern research on the distinction between different kinds of well-being questions includes Benjamin et al. (2012), Clark and Senik (2011), De Neve and Oswald (2012), Huppert et al. (2009), Kahneman and Deaton (2010), Powdthavee and van den Berg (2011), and Stone and Mackie (2013).

Appendix 2: Background data

Year 2011 distributions of ONS well-being answers in the UK
(source: Christopher Deeming, 2013)