

NBER WORKING PAPER SERIES

WHAT DO SELF-REPORTS OF WELLBEING SAY ABOUT LIFE-CYCLE THEORY  
AND POLICY?

Angus Deaton

Working Paper 24369  
<http://www.nber.org/papers/w24369>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2018

I thank Gallup for access to their data, and I acknowledge financial support from them, for whom I am a consulting Senior Scientist. The research was supported by the NIA through grants to the NBER, R01AG040629, R01AG053396, and P01AG005842, to the University of Southern California, R01AG051903, and to Princeton University P30AG024928. Prepared for an issue of the Journal of Public Economics in honor of Tony Atkinson. I am grateful to Marc Fleurbaey, Carol Graham, Ori Heffetz, Danny Kahneman and Arthur Stone for discussions on these issues over many years and to Ori Heffetz, Erzo Luttmer, Andrew Oswald, Anita Pugliese, Julie Ray, Rajesh Srinivasan, and a referee for comments on an earlier version. Most of all, I am grateful to Tony Atkinson for a lifetime of intellectual leadership, support, guidance, extraordinary wisdom and insights that were often many years ahead of their time. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Angus Deaton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What do Self-Reports of Wellbeing Say about Life-Cycle Theory and Policy?

Angus Deaton

NBER Working Paper No. 24369

March 2018

JEL No. E21,H2,H31,H55,I3

**ABSTRACT**

I respond to Atkinson's plea to revive welfare economics, and to considering alternative ethical frameworks when making policy recommendations. I examine a measure of self-reported evaluative wellbeing, the Cantril Ladder, and use data from Gallup to examine wellbeing over the life-cycle. I assess the validity of the measure, and show that it is hard to reconcile with familiar theories of intertemporal choice. I find a worldwide optimism about the future; in spite of repeated evidence to the contrary, people consistently but irrationally predict they will be better off five years from now. The gap between future and current wellbeing diminishes with age, and in rich countries, is negative among the elderly. I also use the measure to think about income transfers by age and sex. Policies that give priority those with low incomes favor the young and the old, while utilitarian policies favor the middle aged, and men over women.

Angus Deaton

Woodrow Wilson School

127 Julis Romo Rabinowitz Building

Princeton University

Princeton, NJ 08544-1013

and NBER

deaton@princeton.edu

## 1. Introduction

In a paper published in 2011 and titled “The restoration of welfare economics,” Atkinson mourned the absence of explicit discussions of welfare economics from the contemporaneous teaching and practice of economics. He noted that economists routinely make welfare claims on the basis of an implicit utilitarianism that neglects other approaches that might lead to different policy recommendations. Traditional welfare economics starts from the assumption that individual welfare is represented by a utility function that is also the basis of decisions. The magnitudes needed for applied welfare economics—cost benefit analyses, price indexes, consumer surplus—can then be inferred from behavior under the assumption that people are well-informed and do what is good for them. By contrast, modern behavioral economics thrives on the supposition that people make mistakes, sometimes systematically so because of widespread behavioral biases, so that we lose one of the supports of the traditional approach. Decision utility and welfare utility need not be identical.

Atkinson (2011) approvingly quotes Samuelson’s (1947, 220) statement that it is “a legitimate concern of economic analysis to examine the consequences of various value judgments,” to which I want to add the words “and alternative measures of welfare.” In this paper I want to examine how different measures of welfare and different value judgments affect policy; my examples are largely illustrative but some of their features will certainly carry through to more comprehensive analyses. In particular, I want to experiment with the use of self-reported wellbeing (SWB) measures as a basis for doing welfare economics, and to see how they interact with alternative ethical frameworks, not just utilitarianism. The policies I use to illustrate are those that affect people differently depending on their age, for example in redistribution from young to old.

Self-report wellbeing measures do not need to be related to behavior. If decision utility differs from welfare utility, and if people sometimes behave against their best interests, the direct measurement of wellbeing might still give an accurate measure, and might even enable people to do better, either through paternalistic government policies, or external nudges, but more simply by providing information on the circumstances and choices that promote wellbeing, what Halpern calls “deshrouding,” see Dunn and Norton (2013, 154) and O’Donnell et al (2014, 71–2) for discussions of the concept. If people have little idea what is good for them, providing that knowledge might be a public service, as in, “don’t become a lawyer because lawyers are typically unhappy with their lives.” We also know that people sometimes choose away from the options they believe will make them happiest, or will give them the best possible life, and that this happens in both hypothetical and in high-stakes real life choices, Benjamin, Heffetz, Kimball and Rees-Jones (2012), (2014).

Direct measures may also capture aspects of welfare beyond real income, which is what economists typically use to proxy utility. Health is the most obvious example; education, civil liberties, civic participation, respect, dignity, and freedom are others. The provision of public goods may or may not affect market behavior, but it may show up in self-reports. The literature reports “happiness regressions” in which self-reported wellbeing is related to health, to public goods, and to other measures that are beyond the reach of market-based welfare measures, including religiosity, race, ethnicity, gender, friendship and unemployment. For Benthamite utilitarians, happiness is *the* measure of wellbeing and its aggregate is the appropriate target of public policy; if self-reported wellbeing (SWB) actually captures Benthamite utility, its measurement is all we need for welfare economics and for policy, see Layard (2005) who argues along these lines.

Legitimate doubts remain. There are many different measures of happiness that capture different aspects of feelings and self-evaluation; they do not correlate perfectly with one another, and they correlate differently with circumstances like income that affect wellbeing, Kahneman and Deaton (2011). It is sometimes not clear what it is that people have in mind when they tell you about their wellbeing, or even whether they understand or can answer the questions. For these and other reasons, economists have traditionally been skeptical of the value and usefulness of self-reports of wellbeing. Such skepticism is reinforced if experimenters can manipulate reported SWB by irrelevant cues, or by the order in which questions are asked, see Schwarz and Strack (1999) though Yap et al (2017) document more positive findings. The stability and reproducibility of self-reports is an important topic that remains unsettled. For example, the effects of an immediately prior question about the state of the country had a large negative effect on the reports about state of the respondent using the Cantril ladder in Deaton (2011) and Deaton and Stone (2012), and those effects are large enough to compromise its use in at least some contexts, such as time-series monitoring of national wellbeing.

In spite of the complexities, there are many passionate defenders of the measures, in both economics and psychology, and there is general agreement even among skeptics that these measures usually behave in a sensible way, and provide insights that would not be available otherwise.

My context here is the life-cycle, I shall examine how self-reported wellbeing varies with age, and try to tell whether the welfare measure it provides makes sense, what standard concept of welfare (if any) it might correspond to, and what policies might result under alternative ethical approaches if we were to accept the measures in place of those that are

standard in economics, such as real income. In this, I am again following Atkinson, whose attention to the life cycle was a constant theme throughout his work, starting with Atkinson (1971) on the distribution of wealth and the life cycle. As he noted there and many times since, national measures of poverty, income, wealth, and inequality must be interpreted with an understanding of age-patterns of wealth, income, and consumption. Society contains individuals of many ages, whose wellbeing is different by virtue of their age, and who often have competing interests. Important government policies must trade off the welfare of the young, the middle-aged, and the elderly.

I take a consciously schizophrenic approach that reflects my dual purpose. On the one hand, I am examining whether SWB measures makes sense, are internally consistent, and respond sensibly to variables that we think of as affecting wellbeing. On the other hand, I will try to explore the consequences of using the measure on the assumption that it does actually make sense. The second use should obviously be conditioned on the first.

The measure I use is an *evaluative* measure of wellbeing that asks people to report, on an eleven-point scale, from 0 to 10, how their life is going. The question is originally due to Cantril (1965), and is asked in exactly the same way of all individuals sampled by Gallup in their World Poll. The question is *“Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally stand at this time?”* There is no mention of happiness, so the ladder is explicitly *not* a hedonic measure that enquires into momentary mood or feelings. Rather it asks people to assess how their life is going “at this time,” an answer to which requires cognitive effort by the respondent, and which is a more considered assessment of

wellbeing than trying to weight together or average the host of emotions and feelings that make up the evanescent texture of everyday life. In the surveys I use here, the Cantril ladder question is immediately followed by a question identical to the ladder question but with the last sentence replaced by, *“Just your best guess, on which step do you think you will stand in the future, say about five years from now?”* I shall use this question too.

The Gallup World Poll has collected data on individuals ages 15 and above in countries around the world since 2006. I use more than 1.7 million observations from 166 countries from 2006 until the end of 2016. Not all countries are included in every year, and the sample sizes range upwards from a baseline standard of 1,000 observations per year, see the note to Figure 1a for sample sizes for each country. Data are usually collected over a period of a few months, and so do not necessarily capture the means for whole years; this could be important in largely agricultural countries. Data are collected by telephone in countries where telephones are owned by at least 80 percent of the population, and face-to-face elsewhere; in common with other commercial telephone surveys, response rates are low in rich countries, with a median of 14 percent, but much higher in face-to-face settings, where the median is 69 percent. I shall also present material from the Gallup–Sharecare Well-being Index (GSWBI) poll, which is a large-scale telephone survey of 1,000 (most recently 500) Americans every evening. Both surveys provide weights that are used to correct means for the sampling design and for non-response. The response rates and survey timings should be kept in mind when comparing the results with those from other, especially official, surveys that employ questions about life satisfaction that differ from the Cantril ladder.

## **2. Life-cycle patterns of evaluative wellbeing and their interpretation**

Figures 1a and 1b show the global patterns of evaluative wellbeing (the ladder) by sex and by region of the world. Those aged 15–24 comprise the first age group; other age groups, 2 through 6, are for those aged 25–34, through to 65–74. I have deleted individuals aged 75 and over, not through lack of interest, but because there are relatively few such people in the sample in some regions. Within each region, I have counted each country equally irrespective of population so that, within regions, I average over countries without weighting. I have used survey sampling weights within countries. The countries in each region are listed in the notes to Figure 1a together with the numbers of observations in each. The two figures are identical except for choice of scales. Figure 1a uses the same scale for all regions while, in Figure 1b, each region has its own scale. The first allows us a clear perception of differences across regions and their importance compared with differences by age and sex while the second allows a much clearer picture of age and sex within each region. In some regions, such as Africa and the World as a whole, the range is quite small and Figure 1b can tend to exaggerate the effects associated with age and sex.

Immediately noticeable are:

1. Average ladder values vary greatly around the world, from around 4 in Africa, to between 7 and 8 for the rich countries of Europe and the English-speaking world. Earlier work with the World Poll has established an approximately linear relationship across countries between average ladder scores and the logarithm of price-adjusted per capita GDP, Deaton (2008), Stevenson and Wolfers (2008).
2. Differences between men and women within regions are smaller than differences between regions. Women tend to evaluate their lives somewhat more highly than

men, except in Africa, and sometimes among those over 60. The female advantage is largest in the English-speaking countries, in Continental Europe, and in East Asia.

3. Age patterns are apparent, but neither universal, nor very pronounced, at least compared with those associated with international differences in incomes.

There is a large previous literature on age patterns in SWB. Perhaps best known are a series of papers by Blanchflower and Oswald (2008, 2017) that argue that there is a universal mid-life low in wellbeing; others, including Steptoe et al (2015), who use an earlier version of the Gallup data, do not find any consistent pattern. Different authors use different countries and different data sets with different SWB questions, so it is possible that the age patterns in the Gallup data are different from those that come from other questions and different survey protocols; it would be an important (if daunting) task to make systematic comparisons. Another important difference across studies is between those that work with the raw means, as in Figures 1, and those, like much of Blanchflower and Oswald's work, look for age patterns after adjustment for other factors; depending on which controls are used, adjustment often has minor effects, as in Blanchflower and Oswald (2017) and Stone et al (2010).

In this paper, and for these figures, I make no adjustments and make no claims about what these data would show with various possible controls. One minor practical reason is that it is difficult to apply consistent controls to the Gallup data, not because the questions do not exist, but because their meaning varies so much across the globe, with different patterns of education, work, retirement, and health systems. A weightier argument is that many possible and potentially important controls are age dependent, including income and the presence of children but especially health, disability and marital status. If we adjust for

these and find, for example, relatively high SWB among the elderly, we have uncovered the not very interesting fact that people in their 70s would rate their lives highly if they were in prime health, and if their lost friends and spouses were returned to them. The *unadjusted* increase in SWB that is often found among the elderly, as it is if SWB has the U-shape, is an important finding on its own, simply because we so strongly expect the opposite. So is any unconditional midlife dip, given that income peaks in midlife. As for children, whose existence is (largely) chosen by their parents, their “effect” on SWB is ill-posed and they should not be controlled for, any more than we control for choices such as how people spend their budgets or allocate their wealth.

In Figures 1, the (unconditional) U-shape appears in the English-speaking countries (U.K., U.S., Canada, Ireland, New Zealand and Australia), to a lesser extent in East and in South Asia and perhaps in Latin America and the Caribbean—though only in the last age group, and in Europe—more for men than women—but not elsewhere. The World as a whole shows the U-shape, but note the very limited range of the ladder with age, from 5.3 to 5.5. In the two poorest regions, Africa and South Asia, life evaluation is low throughout life and, in Africa, it falls with age. In the ex-Communist countries of Asia and Eastern Europe, life-evaluation is markedly lower among the elderly. This is consistent with the fall of communism opening up new opportunities for the young, while destroying the hopes and beliefs of many of their grandparents who, in some cases, also lost pensions and healthcare. When I split the observations from these countries into two periods, 2006–2011, and 2012–2016, life evaluation is markedly higher in the second period, so that life is getting better, but the downward age slope is the same, so the premium of youth over age has not diminished. Even so, we can hope that this pattern will eventually fade, and that the events

of the transition have obscured whatever age pattern in wellbeing would be seen in a more stable environment. At the same time, the plausibility of the pattern in these countries, and the contrast with other regions, provides some face-validity to the Cantril measure.

Other evidence of time effects come from the United States where the examination of General Social Survey data on happiness in Case and Deaton (2017a) provides (weak) evidence that the mid-life dip in wellbeing in the U.S. may be recent, perhaps connected with the increase in mortality and morbidity in mid-life in the 21<sup>st</sup> century, especially among those without a university degree, Case and Deaton (2015, 2017b), see also Graham (2017). The stagnation of working-class wages since the 1970s may also be important here. Those in midlife now have suffered this stagnation throughout their working lives, and about a half of them are worse off than their parents were at the same age, Chetty et al (2017). With rising inequality in wages, mean wages have risen more rapidly than have median wages, and Social Security payments are linked to the former, not the latter. In consequence, the current American elderly, who receive Social Security and whose healthcare is (partially) provided through Medicare, have incomes indexed to average wages, and are currently doing better than those a decade or so younger, most of whom have seen little or no increase in wages. Further evidence in favor of this interpretation comes from the fact that the midlife dip in the ladder is confined to those with less than a college degree.

To what extent are these age patterns consistent with traditional life-cycle models in economics? Do they suggest some reinterpretation either of the theory or of the measures? To think about this, it is helpful to distinguish current from lifetime utility. One standard formulation is to write

$$U = \sum_{a=0}^T v_a \quad (1)$$

where  $a$  indexes age from birth, 0, to death  $T$ . Life time utility is  $U$ , while utility at age  $a$ , sometimes called period utility or “felicity,” is  $v_a$ . Note that the additive form is special, and perhaps implausible, because it rules out potentially important linkages over age, such as habits or even memories. I assume that (1) holds, irrespective of an individual’s current age, so that  $U$  summarizes utility of the *whole life*, no matter whether it is viewed from youth, middle, or old age. Lifetime utility is built, year by year, and the past, present, and future are always counted at all ages. Of course, as someone looks forward, future values of  $v_a$  will be expected values, but will become certain as the individual ages. In this sense,  $U$  is consistent with the Aristotelean idea that it is impossible to tell whether or not an individual has had a happy life until after death when uncertainty is resolved.

If people make rational expectations and so use information efficiently, lifetime utility  $U$  should not change with age if there is no uncertainty—people are just filling in the pre-planned boxes as they go. In reality, with uncertainty, lifetime utility will change but only in the face of new information. As a result,  $U$  will follow a random walk or, more precisely, a martingale difference. If surprises cancel out over individuals, this view of utility predicts a flat profile over life, close to what we see in Africa, or South Asia. Declining health at the end of life is consistent with a flat profile for lifetime utility. Poor health is anticipated as part of the normal process of aging and death, and while some people have worse health than they anticipated, others have better, so that, on average, the age profiles are flat. With economic growth, or temporal declines in mortality and morbidity, so that there are cohort effects, the young will have higher lifetime utility than the old, and the

cross-sectional profiles will slope down, as in the Former Communist countries, the Middle East, East Asia, Latin America, and Europe, with smaller effects in Africa and South Asia. The downward slope should be larger the higher is the rate of economic growth and the faster the improvement in health. These accounts ought to give the two fastest improving regions, East Asia and South Asia, the largest negative slope, which is not apparent in Figures 1.

Consider now an interpretation of the Cantril Ladder as period utility, as  $v_a$  in (1) above, as is perhaps suggested by the inclusion in the question of “at this time.” For a life-cycling individual who controls his or her labor supply and faces the highest wage in mid-life, it is optimal to work hardest and take less leisure in midlife, to earn enough then to finance good times before and after. If so, it is optimal to have lower period utility or felicity in mid-life, which would go some way to generating the U-shape. If this is the explanation, the U-shape ought to be less pronounced for women than for men who have the stronger attachment to the labor force, but this is not apparent in Figure 1b. Standard life-cycle theory is also consistent with either declining period utility over life, if people are impatient and can borrow so that consumption falls over life, or increasing period utility if people are patient, and save to support a rising consumption path. Cohort effects in growing economies will lead to declining cross-sectional age patterns of period utility, just as they do to declining cross-sectional age-patterns of integrated lifetime utility.

In what follows, I shall adopt the period interpretation of the ladder to think about policy, but accept that this can at best explain only a part of SWB. We have good evidence that SWB is higher when people are in good health, or are married, just to mention two of many other factors that play no part in the standard economic story. These other factors

have age patterns of their own that must be taken into account, so that it may not be reasonable to expect the SWB numbers to match the consumption or labor supply accounts of utility in life cycle models. Indeed, the most puzzling aspect of the U-shape of wellbeing, where it exists, is when SWB rises *after* middle-age, when people are losing their spouses, and when both morbidity and mortality are rising. Other components of psychological wellbeing may improve with age, there is less anger, and less stress, Stone et al (2010), and the negative side-effects (e.g. physical pain) of work evaporate with retirement.

The Gallup World Poll follows the Cantril ladder question with a question about the future, asking people on which step of the ladder they expect to be five-years ahead. This question has recently been analyzed by Graham (2017); there is also an important contribution by Schwandt (2016), who looks at a similar question in the German Socio-Economic Panel, and who anticipates many of the results here. Figure 2 shows the results for the same regions of the world as in Figure 1, but with men and women combined. The green line is for five years in the future, while the broken navy-blue line is for the present, and represents the combination of the male and female data in Figure 1. Of particular interest is the comparison of the future with the present, the gap between the two lines.

Among the young, expectations for future wellbeing run far ahead of reported wellbeing today. The gap diminishes with age, and in the rich countries, the lines cross around age 65 after which the future is expected to be worse than the present. Except for this, people appear to be perpetually optimistic about their futures even though this optimism is perpetually frustrated by actual outcomes. It is clear from Figure 2 that, given the relative flatness of the ladder with age, and given the very large differences between the two lines,

especially early in life, that the green lines overestimate, not only current wellbeing, but any reasonable expectation of wellbeing five years ahead.

This (unjustified) optimism seems to happen everywhere in the world, and has at least as strong a claim to be a biological regularity and a general part of the human condition as does the U-shape in wellbeing. Indeed, Kahneman (2011, Chapter 24) argues that a pervasive optimistic bias is “the engine of capitalism,” in that without it, people would not start new businesses. Optimism bias may be part of the normal healthy brain and Sharot et al (2011) study the neural mechanisms that allow the brain to persist in its delusions in spite of reality. Self-reports about the future may also reflect what people *hope* to see for themselves, not what they *expect* to see. People in the Midwest, who are as happy as people in California, think they would be happier were they to live in California because, when asked the question, they focus on the climate, one aspect of life in California that comes immediately to mind, Schkade and Kahneman (1998). Similarly, people who think about the future may focus on the positive outcomes that the future makes possible, and ignore the reality that the future will be a lot like the present.

Schwandt (2016) has used longitudinal data on life satisfaction from the German Socioeconomic Panel, and found a pattern like that for Europe (and the English-speaking countries) in Figure 2. He suggests that the persistent disappointment of the future failing to live up to expectations, and the resulting regret, is an explanation for the midlife dip in wellbeing, and he provides empirical evidence that is consistent with the hypothesis. Yet Figure 2 documents optimism bias even in regions where there is no obvious sign of the midlife dip.

In spite of the familiarity to psychologists of undue optimism, the most obvious feature of Figure 2 is that it makes no sense, at least if we interpret the expectation of future wellbeing as the mathematical expectation of what wellbeing will be five years from now, a point also made by Schwandt (2016). The results defy rationality. Whether current wellbeing is lifetime utility or period utility, the future reports that we see in the data, whatever they signify, cannot be what people rationally expect. Admittedly, the Figure shows cross-sectional results, so that it is mathematically possible that at all ages up to 65, and in all countries, people are constantly experiencing unanticipated very large negative macro shocks, year after year, but such an account stretches credulity, and is disproved for Germany by Schwandt who uses longitudinal data to track individuals over time. And if the future ladder is not, in fact, the expectation of future wellbeing, either lifetime or period, then perhaps the current ladder is not, in fact, a measure of current wellbeing. While noting this possibility, I continue on the supposition that the current ladder is a useful measure of period wellbeing, and that, for the reasons given in the previous paragraph, the future ladder is not an unbiased expectation of future wellbeing.

That said, such an account poses considerable problems for standard life-cycle models, in which people are supposed to equate the marginal utility of money across periods in order to decide on their current consumption, saving, and labor supply. If they have such difficulty estimating their future utilities, it is hard to see how this standard account of behavior could possibly be true. If it is, then the Cantril measures are not obviously related to the utility concepts that appear in models of intertemporal choice.

## 2. Redistributive policies across age groups

Some of the most important and contested public policy decisions involve age- and sex-related transfers of money; pensions are the most obvious, especially in PAYGO schemes where the working population supports the non-working population, but also healthcare, nearly half of which is incurred among those aged 65 and over, with greater spending on women because they live longer, Alemayehu and Warner (2004).

Most welfare discussions of redistribution by age are conducted in terms of income. The GSWBI data for the United States shows that income for the household in which the individual lives is highest when the individual is in midlife, between ages 40 and 55. Gallup asks only a single question about income, with answers given in bins, which I have converted to a continuous measure with reference to the lognormal distribution. Not surprisingly, the mean of Gallup's measure is substantially lower—by about 25 percent—than that from the American Community Survey, a comprehensive source that asks many questions about income, but the age patterns are similar. If we look instead at income per capita, the peak is later, around age 60, because children leaving home increase income per capita. Both income and per capita income are highest in middle or late middle-age, household income is lowest among elderly adults, and per capita household income is lowest among young adults.

If we judge wellbeing by income, or per capita income, and if we are “income prioritarians” defined as believing that the social marginal utility of money declines with income, as in Atkinson's famous 1970 paper, either elderly adults or young adults are those in the greatest need. Note that this ignores the value of Medicare, which is not included in survey

income. It also takes no account of variations in needs by age, for example for health expenditures or fuel among the elderly, or for housing and child-related expenses for young parents. Income prioritarians do not need information on individual wellbeing, only on income and on their ethical judgment that those with lower incomes should have the highest priority for transfers.

Benthamite utilitarians, who care about aggregate experienced utility irrespective of its distribution, will prioritize income transfers for those whose *marginal* utility of income is highest. Figure 3 is an attempt, using the GSWBI data, to estimate the pattern of marginal utility by age. These graphs show the coefficients on income in a regression of the (current) ladder on income, separately for each age group, with and without controls for the number of children in the household, race and ethnicity, state of residence, marital status, calendar year, and smoking; the controls are not interacted with age-groups. For marginal utility, as opposed to total utility, we need the controls because the concept of interest is the partial derivative of experienced utility with respect to income; transfers are transfers of income, not of everything correlated with income.

The red lines are for women and the navy-blue lines are for men; solid lines are unadjusted, broken lines adjusted. These patterns are the *opposite* of those for income (or per capita) income prioritarianism, with those in their 50s providing the greatest contribution to aggregate utility from additional income. The age patterns for men and women are similar though men have somewhat higher marginal utility than women, particularly when we control for other factors. Contrary to what is often recommended, transfers to men will do more to improve social wellbeing than transfers to women, at least when social wellbeing is taken to be total utility.

Income prioritarianism and Benthamite utilitarianism do not exhaust the possibilities. Layard and O'Donnell (2015) have argued that we should modify utilitarianism to give additional priority to those with low levels of wellbeing—this is prioritarianism as originally labeled by Parfit (1997)—while, in similar vein, Kahneman and Krueger (2011) suggest that policy should target those who spend a high fraction of their time being unhappy. The first of these would work by targeting marginal utility with additional weight for those with low ladder scores. Figure 4 shows the ladder scores for men and women, with and without controls, exactly paralleling the information in Figure 3 but now including in the (controlled) regression the logarithm of income without age group interactions. In this case, I would argue for *not* controlling, because when we weight marginal utilities, we need to know who has low wellbeing, and not worry about why they are there. The age pattern of wellbeing as measured by the ladder is the opposite of the age-pattern of marginal utility, so that modification of utilitarianism to prioritize those with the lowest ladder will only enhance the original priority for midlife adults. Their marginal utility is highest, and their priority weight is highest because they have the lowest wellbeing. The same compounding works for men over women; they have higher marginal utility of income, and lower utility, so transfers from men to women are desirable.

This analysis is little more than an illustrative exercise, and I have paid little attention to econometric issues, especially when I calculate marginal utilities of money. I also make the strong and not easily defended assumption that the ladder is actual (period) utility and not an ordinal indicator (monotone increasing transform) of it, an assumption that is required for the calculation of marginal utility, though not for the age pattern of wellbeing itself. Beyond that, I have explored only the benefit side of the redistribution, and not

thought at all about the costs, about how the redistribution could be done, or taken into account behavioral responses by those receiving or paying for the transfers.

#### **4. Conclusions**

This paper takes up Atkinson's challenge to bring back welfare into economics, and to do so in a way that does not have to deny the findings of modern behavioral economics. It also thinks about how policy varies with different ethical assumptions. It uses Gallup's data on the Cantril Ladder measure of life evaluation as a measure of (period) utility, and examines the implications for welfare variations by age. I find a midlife dip in wellbeing in rich countries, but less or no evidence of it elsewhere.

Overall, only some features of the ladder are consistent with life cycle theory if it is interpreted as a period, or age-specific wellbeing measure; it is much harder to interpret it as a lifetime measure. People's expectations of the ladder five years ahead are grossly inconsistent with those reports being the mathematical expectations of future period utility; throughout the world, people are generally over-optimistic about the future, except for the elderly in rich countries, who are generally over-*pessimistic* about the future. These findings are of considerable interest, but cast doubt on the idea that people arrange their consumption and labor supply to maximize the expected integral of age-specific utilities, at least if utilities are well measured by the Cantril Ladder as recorded in Gallup's surveys. This does not, by itself, invalidate the ladder itself as a measure of period age-specific wellbeing.

Finally, I illustrate what the ladder measure would imply for distributing income by age, and how the choice of the ladder versus income interacts with different ethical frameworks. In the United States, income prioritarrians will tend to favor redistribution towards elderly or young adults, whose incomes or per capita incomes are lowest. By contrast, utilitarians will favor those for whom marginal utility is highest, which turns out to be those in midlife, and especially men over women, a prescription that would be enhanced by giving additional priority to those whose wellbeing is lowest, also those in midlife. Both approaches would favor men over women. I make no prescriptions here, but am concerned only to show the possibilities that arise from having a direct measure of wellbeing. I also am responding to Atkinson's challenge by emphasizing and illustrating that policy prescriptions depend, not just on measurement, but on ethics, a point that is not always given the weight that it should be when economists discuss policy.

As in the happiness literature in general, my findings are mixed; the Cantril Ladder is clearly useful, is not obviously inconsistent with standard notions of period utility, and allows us to say things about welfare and distribution that we could not say using standard revealed preference methods. It does not assume that people do not make mistakes. Reported expectations of future wellbeing are consistent with the widespread finding of optimism bias in psychology, but inconsistent with the rational expectations formulation that is standard in much of economics. Whether this is an advantage or a disadvantage will surely differ according to preference and disciplinary background.

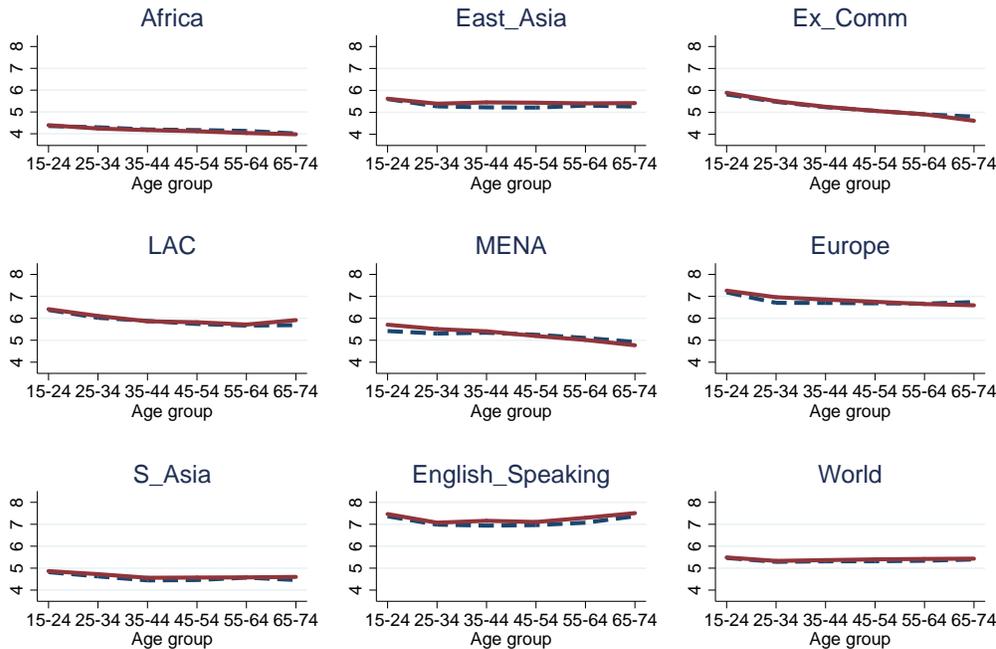
## **5. Citations**

Alemayehu, Berhanu, and Kenneth E. Warner, 2004, "The lifetime distribution of health care costs," *Health Services Research*, 39:3, 627–42

- Atkinson, Anthony B., 1970, "On the measurement of inequality," *Journal of Economic Theory*, 2:3, 244–63.
- Atkinson, Anthony B., 1971, "The distribution of wealth and the individual life cycle," *Oxford Economic Papers*, 23:2, 239–54.
- Atkinson, Anthony B., 2011, "The restoration of welfare economics," *American Economic Review*, 101:3, 157–61.
- Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Alex Rees-Jones, 2012, "What do you think would make you happier? What do you think you would choose?" *American Economic Review*, 102(5), 2083–110.
- Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Alex Rees-Jones, 2014, "Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices," *American Economic Review*, 104(11), 3498–528.
- Blanchflower, David G. and Andrew J. Oswald, 2008, "Is well-being U-shaped over the life-cycle?" *Social Science and Medicine*, 66: 1733–49.
- Blanchflower, David G., and Andrew J. Oswald, 2017, "Do humans suffer a psychological low in midlife? Two approaches (with and without controls) in seven data sets," Cambridge, MA. NBER Working Paper 23724 (August).
- Cantril, Hadley, 1965, *The pattern of human concerns*, New Brunswick, NJ. Rutgers University Press.
- Case, Anne, and Angus Deaton, 2015, "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21<sup>st</sup> century," *PNAS*, 112:49, 15078–83.
- Case, Anne, and Angus Deaton, 2017a, "Suicide, age, and wellbeing: an empirical investigation," in David A Wide, ed., *Insights in the economics of aging*, Chapter 10, 307–34.
- Case, Anne and Angus Deaton, 2017b, "Mortality and morbidity in the 20<sup>th</sup> century," *Brookings Papers on Economic Activity*, Spring, 397–443.
- Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang, 2017, "The fading American dream: trends in absolute income mobility since 1940," *Science*, 356, 398–406.
- Deaton, Angus, 2008, "Income, health, and wellbeing around the world: evidence from the Gallup World Poll," *Journal of Economic Perspectives*, 22:2, 53–72.
- Deaton, Angus and Arthur A. Stone, 2016, "Understanding context effects for a measure of life evaluation: how responses matter," *Oxford Economic Papers*, 68:4, 861–70
- Dunn, Elizabeth and Michael Norton, 2013, *Happy money: the science of smarter spending*, New York. Simon and Schuster
- Graham, Carol, 2017, *Happiness for all? Unequal hopes and lives in pursuit of the American Dream*, Princeton, NJ. Princeton.
- Kahneman, Daniel, 2011, *Thinking fast and slow*, New York. Farrar, Strauss, and Giroux.
- Kahneman, Daniel and Angus Deaton, 2011, "High income improves evaluation of life but not emotional wellbeing," *PNAS*, 107(38), 16489–93.
- Kahneman, Daniel and Alan B. Krueger, 2006, "Developments in the measurement of subjective well-being," *Journal of Economic Perspectives*, 20:1, 3–24.
- Layard, Richard, 2005, *Happiness: lessons from a new science*, Penguin Press.
- Layard, Richard, and Gus O'Donnell, 2015, "How to make policy when happiness is the goal," Chapter 4 in John Helliwell, Richard Layard, and Jeffrey Sachs, eds., *World*

- Happiness Report 2015*, 76–87. <http://worldhappiness.report/wp-content/uploads/sites/2/2015/04/WHR15.pdf>
- O'Donnell, Gus, Angus Deaton, Martine Durand, David Halpern, and Richard Layard, 2014, *Wellbeing and policy*, London. Legatum Institute.
- Parfit, Derek, 1997, "Equality and priority," *Ratio*, 10(3), 202–21.
- Samuelson, Paul A., 1947, *Foundations of economic analysis*, Cambridge, MA. Harvard.
- Schkade, David A., and Daniel Kahneman, 1998, "Does living in California make people happy? A focusing illusion in judgments of life satisfaction," *Psychological Science*, 9(5), 340–6.
- Schwandt, Hannes, 2016, "Unmet aspirations as an explanation for the age U-shape in well-being", *Journal of Economic Behavior and Organization*, 122, 75–87.
- Schwarz, Norbert, and Fritz Strack, 1999, "Reports of subjective well-being: judgmental processes and their methodological implications," in Kahneman, Daniel, Ed Diener, and Norbert Schwarz, eds., *Well-being: the foundations of hedonic psychology*, New York. Russell-Sage.
- Sharot, Tali, Christoph W. Korn, and Raymond J. Dolan, 2011, "How unrealistic optimism is maintained in the face of reality," *Nature Neuroscience*, 14(11), 1475–9.
- Stephens, Andrew, Angus Deaton, and Arthur A. Stone, 2015, "Subjective wellbeing, health, and ageing," *The Lancet*, 385:9968, 640–8.
- Stevenson, Betsey, and Justin Wolfers, 2008, "Economic growth and subjective well-being: reassessing the Easterlin paradox," *Brookings Papers on Economic Activity*, Spring, 1–87.
- Stone, Arthur A., Joseph E. Schwarz, Joan E. Broderick and Angus Deaton, 2010, "A snapshot of the age distribution of psychological wellbeing in the United States," *PNAS*, 107(22), 9985–90.
- Yap, Stevie C. Y., Jessica Wortman, Ivana Anusic, S. Glenn Baker, Laura D. Scherer, M. Brent Donnellan, and Richard E. Lucas, 2017, "The effect of mood on judgments of subjective well-being: nine tests of the judgment model," *Journal of Personality and Social Psychology*, 113(6), 939–61.

## Cantril's Ladder

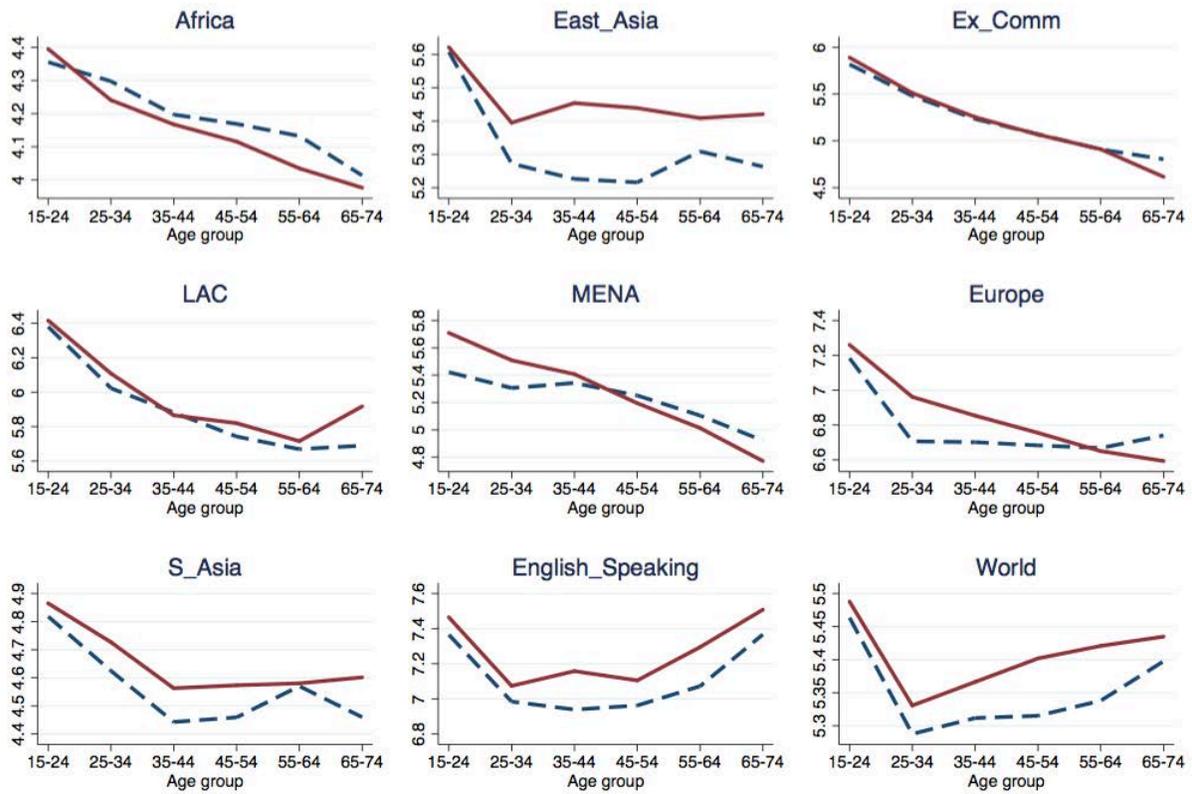


**Figure 1a: Cantril ladders by sex and region of the world: all regions to same scale**

Notes: Solid red lines are women, broken navy lines are men. See text for definition of ladder question. Data from the Gallup World Poll, 2006 to 2016; not all countries are sampled in every year. Within regions, each country is weighted equally; within countries, survey sampling weights are used. The plots are averages within age groups, without adjustment for any covariates. Sample sizes given in brackets. **Africa (323,799)** comprises Nigeria (12,002), Kenya (12,200), Tanzania (11,016), Ghana (11,008), Uganda (11,000), Benin (8,000), Madagascar (8,016), Malawi (9,000), South Africa (12,001), Angola (4,000), Botswana (9,000), Ethiopia (6,004), Mali (10,000), Mauritania (12,992), Mozambique (5,000), Niger (11,016), Rwanda (9,584), Senegal (11,000), Zambia (10,001), Burkina Faso (10,008), Cameroon (11,200), Sierra Leone (9,016), Zimbabwe (11,000), Burundi (4,000), Central African Republic (4,000), Chad (11,000), Comoros (7,000), Democratic Republic of Congo (7,000), Congo (7,000), Gabon (6,016), Guinea (6,008), Côte d'Ivoire (5,008), Liberia (8,000), Namibia (2,800), Somalia (3,191), Sudan (7,592), Togo (6,000), Somaliland (7,000), and South Sudan (3,000). **East Asia (201,526)** comprises Indonesia (13,390), Hong Kong (8,118), Singapore (11,652), Japan (15,153), China (51,914), Philippines (12,200), South Korea (13,101), Taiwan (10,003), Malaysia (11,266), and Mongolia (9,000). Ex-communist (**Ex\_Comm**) countries (**312,248**) comprise Poland (11,029), Hungary (10,091), Czech Republic (10,169), Romania (10,030), Belarus (11,587), Georgia (11,080), Kazakhstan (11,000), Kyrgyzstan (11,000), Moldova (11,000), Russia (26,021), Ukraine (11,323), Albania (10,069), Armenia (11,000), Azerbaijan (11,000), Bosnia and Herzegovina (11,038), Bulgaria (9,009), Croatia (10,068), Estonia (9,234), Latvia (9,074), Lithuania (10,029), Macedonia (10,177), Montenegro (9,837), Serbia (10,618), Slovakia (8,048), Slovenia (9,535), Tajikistan (11,000), Turkmenistan (7,000), Uzbekistan (10,000), Kosovo (10,182) and Nagorno-Karabakh Republic (1000). Latin America and Caribbean (**LAC, 213,432**) comprises Venezuela (11,000), Brazil (12,235), Mexico (12,054), Costa Rica (11,010), Argentina (11,000), Belize (1,006), Bolivia (11,003), Chile (11,247), Colombia (11,000), Cuba (1,000), Dominican Republic (11,000), Ecuador (11,135), El Salvador (11,008), Guatemala (11,050), Guyana (501), Haiti (4,533), Honduras (11,004), Jamaica (2,057), Nicaragua (11,016), Panama (11,028), Paraguay (11,001), Peru (11,000), Puerto Rico (1,000), Suriname (504), Trinidad and Tobago (2,018), and Uruguay (11,022). Middle East and North Africa (**MENA, 232,447**) comprises Egypt (20,924), Morocco (9,072), Lebanon (15,053), Saudi Arabia (16,432), Jordan (15,039), Syria (11,452), Turkey (13,006), Iran (12,864), Palestinian Territories (15,014), Algeria (7,031), Bahrain (12,191), Djibouti (5,000), Iraq (14,024), Kuwait (13,027), Libya (3,009), Qatar (7,060), Tunisia (11,288), United Arab

Emirates (15,945) and Yemen (13,000). **Europe (221,940)** comprises France (12,989), Germany (37,394), Netherlands (10,757), Belgium (11,080), Spain (13,031), Italy (13,039), Sweden (11,762), Greece (10,005), Denmark (11,780), Israel (11,004), Austria (11,010), Cyprus (8,547), Finland (9,766), Iceland (3,131), Luxembourg (8,003), Malta (8,050), Norway (8,010), Portugal (11,062), Switzerland (6,514) and Northern Cyprus (5,006). South Asia (**S\_Asia**) comprises Pakistan (17,151), Bangladesh (13,248), India (44,434), Sri Lanka (11,248), Afghanistan (11,010), Bhutan (3,040), and Nepal (12,107). **English speaking** comprises United States (13,444), United Kingdom (33,651), Canada (13,477), Australia (11,232), New Zealand (9,798) and Ireland (10,502).

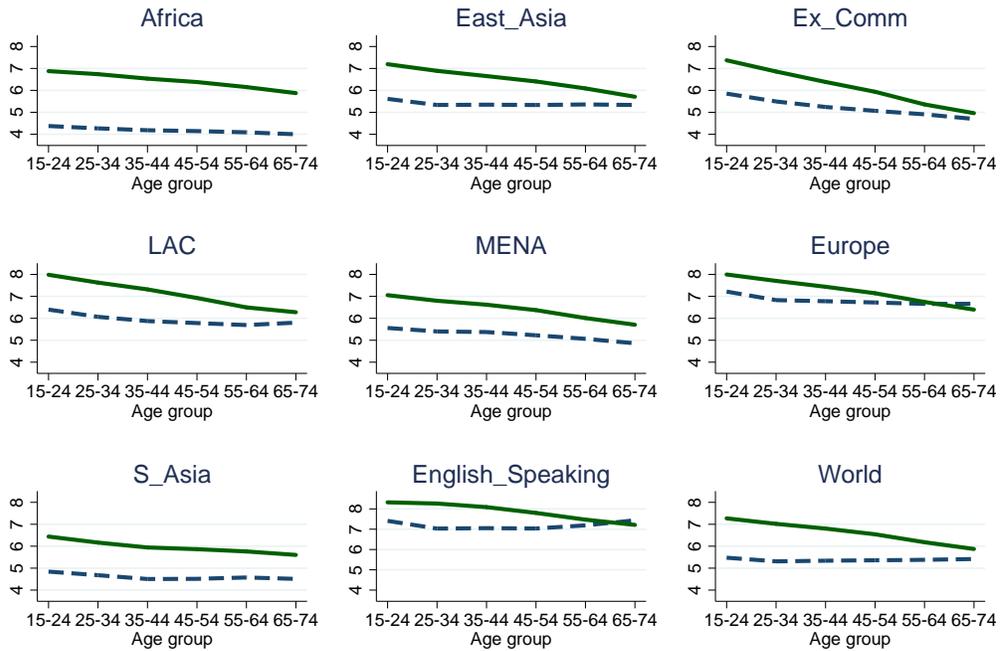
## Cantril's Ladder



**Figure 1b: Cantril ladder by sex and region of the world: scales differ by region**

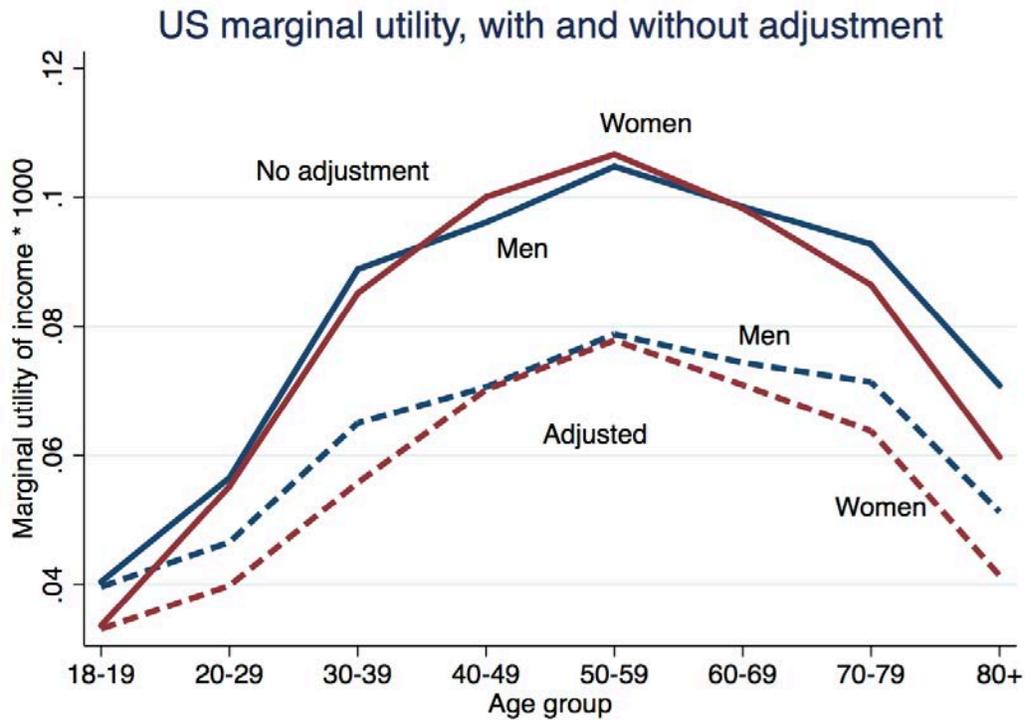
Notes: For definitions of regions and abbreviations, see Figure 1a.

## Current v Future Ladder



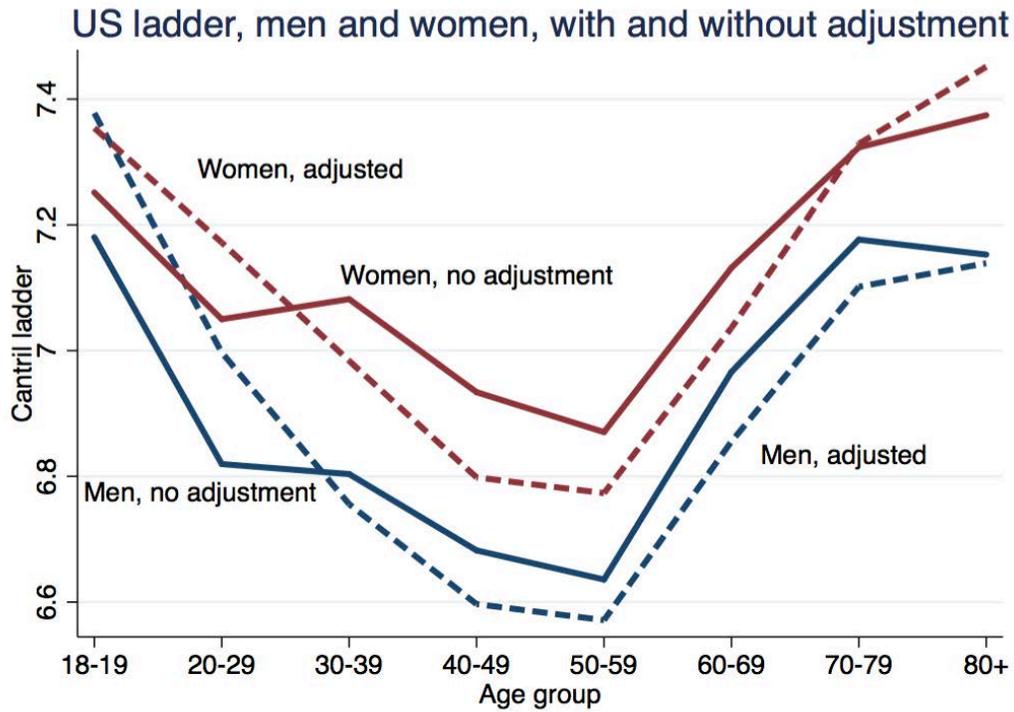
**Figure 2: Ladder present and future, men and women combined**

Notes: Solid green line is the expected value of the ladder five years from now, and the broken navy line is the ladder today. See Figure 1 for list of countries.



**Figure 3: Estimated marginal utility of income by age, with and without controls**

Notes: There are four regressions, two for men, and two for women. The “no adjustment” lines comes from a regression of the ladder on dummies for each age group, and for age group dummies interacted with income, and the estimated marginal utilities are the coefficients on each age group in the interactions. The “adjusted” lines come from the same regression but with controls as follows: dummies for the number of children, dummies for race and ethnicity, dummies for marital status, for state of residence, for education level, for year, and for whether or not the respondent is a smoker.



**Figure 4: Mean ladder scores by gender and with and without adjustment for controls**

Notes: For controls, see the notes to figure 3. Here, in the regression with the controls, income is entered as log income, and there are no interactions between age-group and income.