The U-shape without controls: A response to Glenn[☆]

David G. Blanchflower [b,c], Andrew J. Oswald [a,*]

[a] Warwick University, Economics, Coventry CV4 7AL, UK
[b] Dartmouth College, USA
[c] University of Stirling, UK

ARTICLE INFO

ABSTRACT

This paper is a continuation of results reported in our article "Is well-being U-shaped over the life cycle?" (Blanchflower & Oswald, 2008). It provides new evidence that well-being follows a curve through life. We use data on half a million randomly sampled individuals across eight major European nations. Importantly, we show that in this set of countries there is a U-shape even in unadjusted data, that is, without the inclusion of control variables. But we also advise against a focus on elementary bivariate associations.

© 2009 Elsevier Ltd. All rights reserved.

Introduction

Our article "Is well-being U-shaped over the life cycle?" (Blanchflower & Oswald, 2008) provided evidence that human well-being follows an approximately U-shaped path through life. We argued that it starts high among the young; then it flattens out to a minimum around middle age; then it rises quite strongly up into a person's 70s. This is for the average person, and for the typical country in our data set, and it holds ceteris paribus in the sense that it emerges from a regression equation with other personal variables held constant.

In his critical commentary on our work, Glenn (2009) has made a valuable contribution. In the commentary, various points are raised. But his key point can be summed up, in Glenn's words, as the claim that: "… the appearance of this U-shaped curve of well-being is the result of the use of inappropriate and questionable control variables".

First, we explain in this short response that Glenn's claim is incorrect. In many countries, the U-shape can be found without any control variables. We believe Glenn's mistake has been to focus too heavily on the United States. It is known by researchers that the United States does not have a well-being U-shape in age if control variables are omitted.[1] This was shown by Easterlin (2006) and was discussed in detail in early versions of work by Blanchflower (2009) and Blanchflower and Oswald (2004). Strangely, Glenn looks only at data from a single nation, whereas the object of our original article

was to find a relationship common across large numbers of nations (Blanchflower & Oswald, 2008).

Second, we disagree with Glenn's methodological position, which seems to be that social scientists should not hold constant other factors when they study the relationship between well-being and age. We take up these two issues in turn.

The U-shape in unadjusted data without controls

A straightforward way to dispose of the claim in Glenn's quote given above is to show a U-shape in a setting without control variables. Fig. 1 does so. It takes a random sample of Europeans aged between 18 and 80 from: Belgium; Denmark; France; Germany; Italy; Ireland; Netherlands; and the United Kingdom. The overall sample size is approximately half a million people (more exactly, 546,038), and the data come from the 1973–2006 Eurobarometer Surveys. The dependent variable is taken from answers to a life-satisfaction question. For the sake of transparency, we estimate this regression equation with Ordinary Least Squares (OLS). Technically, this is not quite the right form of statistical estimation method, but the substantive results are the same if we use instead an ordered probit equation or equivalent.

We deliberately estimate the equation behind Fig. 1 with none of the supposedly 'inappropriate' control variables we used in our 2008 (Blanchflower and Oswald) paper. Here we use only year dummies, country dummies, and a dummy for gender. Moreover, to show that the broad mathematical shape in the data is not driven by the imposition of a quadratic estimation form, we do not impose any parametric structure. Instead, we do something simple. We enter

---

[1] However, in GSS data for the United States, well-being is U-shaped in age for the non-married subsample.
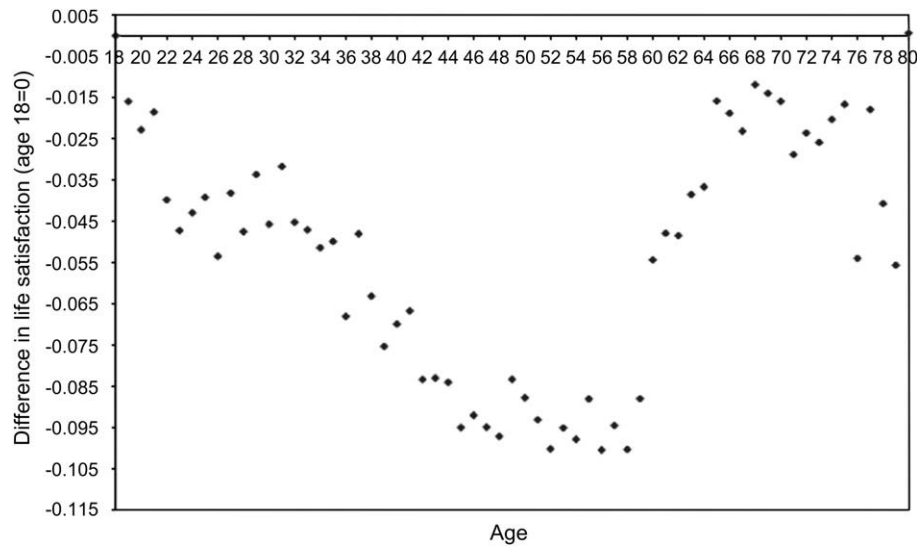
Fig. 1. Unrestricted U-shape in life-satisfaction across age for eight European countries, 1973–2006. No personal controls.

a separate dummy variable for every two years of age (so we have a separate dummy for people aged either 38 or 39, another dummy for people aged 40 or 41, another for those of age 42 or 43, and so on). We then plot the pattern in these age dummies and see what emerges.

Fig. 1 depicts the key result. Well-being is on the vertical axis and age is on the horizontal axis. To our eyes, this is remarkably close to a U-shape. Yet we have included <u>none</u> of the time-varying variables, such as a dummy variable for marital status, to which Glenn in the commentary objects. Nor have we imposed any functional form on the data to force the numbers to take a particular geometric form (which is why Fig. 1 has a ragged pattern of dots).

Table 1 gives more information on the same point. Here we look separately at well-being OLS regression equations for France, Germany and the UK. We give a comparison with the US, from the General Social Survey. In the first column, there are no personal control variables (apart from gender). In the second, we bring in marital status as an extra control. It can be seen that with no controls there is a U-shape in each of France, Germany, and the United Kingdom. The turning points of the U-shape, i.e. the minimum in well-being, are at 47.0 years in France, at 44.0 years in Germany, and at 40.7 years in the UK. Yet, as Glenn correctly points out, and as has long been known, there is no minimum in raw American data. When marital status is included, however, the US has a well-determined U-

shape in well-being through the life cycle, with a well-being minimum here at age 38.2. As can be seen from the second column of the lower half of Table 1, the four countries then look rather similar.

Even on Glenn's own terms, therefore, his general claim is incorrect. What Glenn can legitimately say is that, unusually, the United States has no U-shape in unadjusted data. He could have suggested that researchers need to spend more time trying to evaluate why the USA is different from so many other industrialized countries. This may, as a conjecture, be because of different trends in marriage, but the issue remains open.

Should control variables be included? Consider smoking and disease

Glenn takes an extreme methodological position – one with which we disagree. He argues that the social scientist's standard personal controls (marital status, income, education, etc) should be omitted from these kinds of well-being equations, and goes on to argue that "Blanchflower and Oswald should…provide estimates of what has really happened to well-being through the life-course in the countries studied" (Glenn, 2009). We hope Fig. 1 goes some way to the latter goal.

Table 1
The life-satisfaction U-shape in age in Europe and the United States.

	Age	Age^2	Age	Age^2
a) Regression equation coefficients on age and its square (first without controls; then with marital status as a control)				
France	−0.01968	0.0002092	−0.02470	0.0002664
Germany	−0.00516	0.0000587	−0.01036	0.0001128
UK	−0.00612	0.0000753	−0.01313	0.0001526
All 8 European nations	−0.00800	0.0000815	−0.01551	0.0001628
USA (a maximum; a minimum)	+0.00495	−0.0000375	−0.00600	0.0000786
		(1)		(2)
		Without controls		Plus marital status
b) The life-satisfaction minima in Europe and the USA				
France ($n = 63,221$)		Age 47.0		Age 46.4
Germany ($n = 92,815$)		Age 44.0		Age 45.9
UK ($n = 81,970$)		Age 40.7		Age 43.0
All eight countries ($n = 546,038$)		Age 49.1		Age 47.6
USA ($n = 46,513$)		No Minimum		Age 38.2

Source: Eurobarometer Surveys and General Social Surveys of the United States.

Overall, however, we are doubtful of the sensibleness of Glenn's argument. It is a matter of methodological judgment, admittedly, so unlike the points we make above – where we think Glenn is wrong – there is scope on this for reasonable people to disagree reasonably.

But think of smoking and cancer. Would Glenn really want medical statisticians in their published articles to report the bivariate association between disease probability and the number of cigarettes smoked? For good reason, medical journals do not favour that. We know that there are two ways to look at the correlation between disease and smoking:

$$Disease\ probability = f(smoking,\ diet,\ education,\ income,\ exercise) \quad (1)$$

$$Disease\ probability = g(smoking) \quad (2)$$

Both of these relationships $f(.)$ and $g(.)$ can be thought of as valid and true, in the sense that the data can be presented in either way. But equation (2) is known to be a misleading way to give epidemiological information to the public. The bivariate association between risk of cancer and the number of cigarettes from Equation (2) is not the same as the ceteris paribus adjusted risk from Equation (1). When compared to non-smokers, people who are smokers tend to have worse diets, less education, lower income, and to take less exercise. The gradient of Equation (2) will then routinely overstate the medical benefit from cutting back on cigarettes.

Ultimately in social science, the control variables that are included in multiple regression equations have to be chosen with an eye on the intellectual or policy question being answered. If the aim is to describe the data, it is reasonable to leave out most or all control variables. "Smokers die at rate Z" is an acceptable statement to make. But that is not the same as "smoking changes your risk by Z". It would be an error to use Equation (2) to tell the public what smoking does to their health. Yet this error is what, in parts of his commentary, Glenn seems to be recommending. Such a view puzzles us. If the aim is to understand relationships, it will rarely be desirable to stop at bivariate patterns.

Conclusion

As will be clear, we do not greatly agree with this commentary. Nevertheless, we are grateful for a chance to discuss the concerns raised. Human well-being is important. Further statistical research on this area will be needed. The difference, in particular, between the (apparently somewhat unrepresentative) United States and many other industrialized countries is a topic on which more work is required.[2]

We hope we have made it clear that the U-shape exists powerfully in raw unadjusted data in many European nations, and that there are strong methodological grounds for well-being researchers to present more than bivariate statistical associations.

References

Blanchflower, D. G. (2009). International evidence on well-being. In Alan B. Krueger (Ed.), *National time accounting and subjective well-being.* Chicago: NBER and University of Chicago Press.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of Public Economics, 88,* 1359–1386.

Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine, 66,* 1733–1749.

Easterlin, R. (2006). Life cycle happiness and its sources. Intersections of psychology, economics, and demography. *Journal of Economic Psychology, 27,* 463–482.

Glenn, N. (2009). Is the apparent U-shape of well-being over the life course a results of inappropriate use of control variables? A commentary on Blanchflower and Oswald (66:8, 2008, 1733–1749). *Social Science & Medicine,* 69(4), 481–485.

Van Landeghem, B. (2008). Human well-being over the life cycle: Longitudinal evidence from a 20-year panel, University of Leuven, unpublished paper.

---

[2] It is also desirable that future analysis try to use true longitudinal data (as in Van Landeghem, 2008).