Social Science & Medicine 69 (2009) 481–485

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Is the apparent U-shape of well-being over the life course a result of inappropriate use of control variables? A commentary on Blanchflower and Oswald (66: 8, 2008, 1733–1749)

Norval Glenn*

Department of Sociology, University of Texas at Austin, 1 University Station, Austin, TX 78712, USA

ARTICLE INFO

Article history:
Available online 17 June 2009

Keywords
USA
Aging
Happiness
Age-period-cohort conundrum
Life course
Logic of causal inference

ABSTRACT

In their article in this journal "Is well-being U-shaped over the life cycle?" Blanchflower and Oswald (Blanchflower, D.G., & Oswald, A.J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine, 66*, 1733–1749) report the results of an ambitious cross-national study of age and well-being and conclude that in most societies studied the well-being of adults tends to be high in young adulthood and old age and lowest around age 40. I argue that the appearance of this U-shaped curve of well-being is the result of the use of inappropriate and questionable control variables. The most clearly inappropriate control variable is marital status, the control of which to a large extent accounts for the U-shaped curve. Most researchers who have studied the relationship between being married and being happy believe that happiness affects marital status (happier people are more likely to marry and stay married), and of course a variable that is affected by the dependent variable should not be included as a control variable in a simple recursive model. Such control variables as income and education are suspect because the relationship between them and well-being is likely to be partially spurious, and such variables as race and whether or not persons lived with both parents at age 16 should not be controlled, because they cannot affect or be affected by age. Finally, year of survey should not be controlled because of the age-period-cohort conundrum, which makes including age, period, and cohort all as predictor variables in a regression inappropriate (and impossible if the three variables are measured precisely and comparably). The only clearly appropriate control variable is birth cohort, and when only it is controlled, the regression data become estimates of how the well-being of persons has actually changed as they have gone through the life course. I argue that such estimates are much more useful than the counterfactual abstractions provided by Blanchflower and Oswald (Blanchflower, D.G., & Oswald, A.J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine, 66*, 1733–1749), and I conclude that those authors (or someone else) could make a very important contribution by redoing their analyses with birth cohort as the only control variable. I do that with the American happiness data and find that the results do not come close to the U-shaped pattern.

Critique

Blanchflower and Oswald (2008), in an article in this journal, report an ambitious cross-national study designed to answer the question, "is well-being U-shaped over the life cycle?" (Among sociologists and social psychologists, the term "life cycle" has been replaced by "life course.") Their answer to that question is generally affirmative, though they report some apparent exceptions in a few of the many countries included in their study.

The Blanchflower and Oswald study is potentially one of the most important studies of age and well-being that has ever been conducted, but I argue, and present evidence to the effect, that its main finding is largely the result of the inclusion of inappropriate control variables in the analyses, at least in the case of happiness in the United States. Therefore, I urge the authors to redo their analyses with the questionable control variables removed so that their study can reach its potential to be a very important contribution to our understanding of how aging tends to be experienced.

The most clearly inappropriate control variable in the research is marital status, which Blanchflower and Oswald enter as a set of dummy variables that distinguish among the different kinds of unmarried persons. Marital status cannot affect age, and thus the

* Tel.: +1 512 232 6320.
E-mail address: ndglenn@austin.utexas.edu

0277-9536/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2009.05.038

implicit assumption on which the decision to include marital status in the simple recursive model is based must be that it is an intervening, or mediating, variable between age and well-being. The full set of implicit assumptions includes:

a. age affects, but is not affected by, marital status;
b. marital status affects, but is not affected by, well-being;
c. the association of marital status with well-being, net of the effects of other variables in the model, is not spurious.

The first assumption is clearly correct, the "not affected by" part of the second is almost certainly incorrect, and the third is suspect.

There is an extensive literature on the relationship between marital status and well-being (e.g., Glenn, 2001; Horwitz & White, 1991; Kessler & Essex, 1982; Marks & Lambert, 1998; Nock, 1998; Stack & Eshleman, 1998; Waite, 1995; Waite & Gallagher, 2000), including considerable attention specifically to personal happiness, one of Blanchflower and Oswald's well-being variables. The relevant research, which has been conducted with data from numerous countries, has consistently shown a moderate to strong relationship between being married and being happy. Most authors who have discussed this relationship *do* believe that it results partly from effects of marriage on happiness, but almost all of them also acknowledge that it is partly a selection effect, that is, an effect of happier persons being more likely than others to marry and stay married. At least one study provides evidence that the relationship is to a large extent the result of such selection (Stutzer & Frey, 2006). Furthermore, the relationship of happiness to marital status may be partially spurious due to the two variables being commonly affected by one or more other variables that are not measured and included in the research, such as, one might suspect, general competence. One study used analysis of residuals to assess the probable common effects of unmeasured variables on marital status and happiness and found evidence for such effects (Blackman, Clayton, Glenn, Malone-Colon, & Roberts, 2005). None of the research provides definitive evidence of the nature of the causal relationship between marital status and happiness, but it does indicate that the relationship is complex and is not just a simple matter of the former affecting the latter. Clearly, specifying a causal model in which marital status simply mediates effects of age on happiness in not warranted and is likely to result in substantially erroneous estimates of effects.

Just how using marital status as a control variable in an age and happiness study can distort estimates of the effects of the former on the latter is illustrated by a simple exercise with hypothetical data based on some extreme assumptions. Suppose that a sample of 100 young adults born in the same year is studied each year as they age from 21 to 29 years old. Assume further that:

a. growing older has no effect on happiness;
b. no one's happiness changes for any reason during the period covered;
c. when happiness is measured on a 10 point scale (0–9), 10 persons fall on each point on the scale;
d. at age 21, only the 10 happiest persons are married;
e. each subsequent year, the 10 happiest of the remaining unmarried persons marry, so that by age 29, only the 10 least happy persons have not married;
f. no one divorces during the course of the study.

Taking the data set based on these assumptions and regressing happiness on age with marital status controlled creates the strong negative relationship between age and happiness shown by the lower line in Fig. 1. Even though age does not affect happiness, controlling marital status creates the impression that it does.

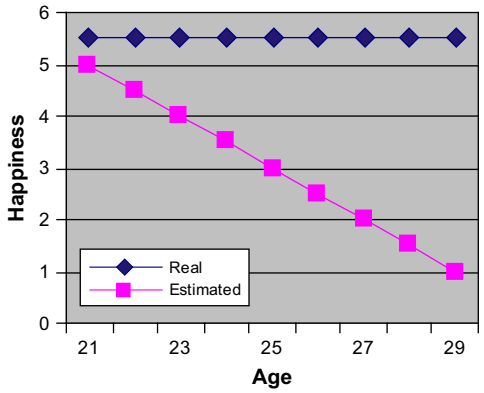

Fig. 1. Real (lack of) effects of age on happiness vs. regression estimates with marital status controlled, hypothetical data.

Of course, no one believes that such an extreme and simple selection effect exists in the real world. However, if any substantial proportion of the relationship between marital status and happiness reflects a selection effect, the distortion with real data will be similar to, though less extreme than, that with the hypothetical data. That is, given the fact that the proportion of persons married increases from youngest adulthood into the thirties and forties, controlling marital status will depress the indicated happiness of early middle-aged persons relative to that of younger adults. In other words, it will bend the data toward the U-shaped pattern found by Blanchflower and Oswald, and much of the change will not reflect an age effect.

Controlling marital status accounts substantially for the U-shaped curve in Blanchflower and Oswald's happiness findings, at least in the United States and very likely in other countries as well given the general similarity of the relationship of marital status with happiness across a large number of countries (Stack & Eshleman, 1998). In Figs. 2 and 3, I show the effects of controlling marital status in the case of the United States General Social Survey data, for men and women, respectively. For these analyses, I used ordinary least squares regression, treated the three-point happiness scale as though it were interval, used the same marital status dummy variables as Blanchflower and Oswald, and used ages 35–39 as the reference category for the age dummy variables because that is the age category for which the mean raw happiness score is the same for males and females.

The male data in Fig. 2 are not distinctly U-shaped, but the lowest indicated level of happiness is at ages 40–44, while the female data in Fig. 3 are more clearly U-shaped with the lowest happiness also at ages 40–44. The most interesting aspect of the difference made by controlling marital status is the increase in indicated happiness at the extremes of the adult ages, a change very substantial for males at the youngest ages and for females at both extremes. What makes this especially important is that the

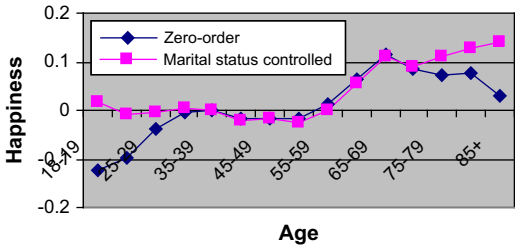

Fig. 2. Regression (unstandardized) of happiness score on age dummy variables, males, combined data from 1972 to 2006 American General Social Surveys.

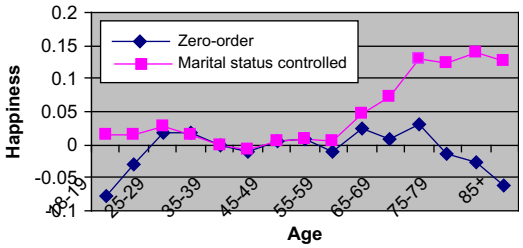

Fig. 3. Regression (unstandardized) of happiness score on age dummy variables, females, combined data from 1972 to 2006 American General Social Surveys.

relationship of marital status to age is substantially weaker for young adult males than for older ones and is relatively weak for females at both of the extreme adult age levels, an interaction that would make the inclusion of marital status as an intervening variable inappropriate in a simple additive model even if there were not other problems with doing so. When happiness is regressed on a dichotomous measure of marital status (married vs. unmarried) at the different age levels, the unstandardized regression coefficient is only 0.184 for males ages 18–19 but rises monotonically with age to 0.379 at age 70 and older, which of course is a statistically significant trend ($p < 0.001$ on a two-tailed test). For females, the coefficients are 0.245 and 0.250 for the two extreme levels but 0.329 for ages 30–49 and 0.359 for ages 50–69, and the difference between each pair of extreme and intermediate categories is significant at the 0.001 level.

Blanchflower and Oswald's inclusion of the other control variable that can only be an intervening variable in their model—log-income—is also problematic, but less so than in the case of marital status. There is little reason to think that happiness substantially affects income, but its relationship with happiness may be spurious to an important extent, being, for instance, partially the result of common influences from intelligence, physical health, mental health, and/or childhood family structure, among other variables. Furthermore, the magnitude of the relationship varies with age, being weaker among the oldest adults than among the younger ones (data not shown). Therefore, log-income should also be dropped from the analyses.

It seems to me that exclusion of these intervening variables would be advantageous even if they met the requirements for inclusion. Their exclusion turns the regression results into estimates of total age effects, "total effects" being defined as the sum of direct effects and of any indirect effects through other variables included in the model. In other words, the regression results become estimates of what really has happened to people as they have grown older, and I argue that these are of greater theoretical and practical importance than estimates of direct effects. The latter alone are not very useful because they are too dependent on which variables through which the total effects occur are included in the analyses. For instance, Blanchflower and Oswald's estimates would change, probably substantially, if measures of physical health were controlled. If all mechanisms through which the effects occur could be identified, represented by variables, and held constant, the estimates of direct effects would be zero.

Most if not all of the several journalists who wrote stories about the Blanchflower and Oswald research interpreted the regression results to represent what really has happened to people as they have aged, not what would have happened if income and marital status had not changed as people grew older. For instance, an article in the prestigious *University of California at Berkeley Wellness Letter* (April, 2008) says that "midlife is a hassle" and that "happiness reaches a low point" in midlife—statements that I show below to be incorrect, at least in the United States. The journalists wanted descriptively accurate data, not the counterfactual abstractions in the Blanchflower and Oswald article, and they interpreted the latter as though they were the data they wanted.

The only variables that need to be controlled to estimate total effects are those that are causally prior to both the independent and the dependent variable and thus that could create a spurious relationship between those two variables or mask a causal relationship between them. Most of the control variables that are left after marital status and income are dropped from Blanchflower and Oswald's analysis do not qualify, strictly speaking, because such variables as education, race, and whether or not the person was living with both parents at age 16 cannot affect age. These variables differ by birth cohort, and one could argue that they could usefully be used as surrogates for that variable if it were not included in the analyses. However, it *is* included, and there are reasons to think that the effects of such variables as years of school completed and race vary by cohort. Therefore, all of these cohort-related variables should be dropped from the analyses.

What variables are causally prior to age and thus should be included as control variables in the analyses? Only two, birth cohort (time of birth) and period (time of measurement of the dependent variable), of which age is a perfect linear function. However, the linear dependence among age, period, and cohort means that when age is the independent variable, and when all three variables are measured comparably and precisely, one cannot simply include both period and cohort as control variables; when one is controlled, the other has zero variance. This is the much discussed age-period-cohort conundrum, which is usually addressed when cohort analyses (of which the critiqued study is an example) are reported but which Blanchflower and Oswald do not acknowledge and address (for discussion of the conundrum, see Converse, 1976; Firebaugh, 1997; Glenn, 2003, 2005; Mason, Mason, Winsborough, & Poo1e, 1973). Rather, they simply include all three of the interrelated variables in the analyses, which they are able to do only because of inconsistency in measurement, with 5-year intervals being used for age, 1-year intervals for period, and 10-year intervals for cohort. This inconsistency allows enough "slippage" for the regression program to run, but there is still a severe collinearity problem and thus the resulting separation of age, period, and cohort effects is not likely to be very accurate. Either period or cohort should be dropped from the analyses, and at least in the case of the United States data on happiness, the dropped variable should be period; the data vary little over the total time covered—evidence (though not proof) of minimal period effects. There logically could be offsetting period and cohort effects that produce no trend, but that is unlikely. The only remaining viable control variable is birth cohort.

Some new estimates of age effects on happiness

Estimates of age effects on happiness computed with the United States General Social Survey data and with only birth cohort controlled (using the same 10-year cohort dummy variables used

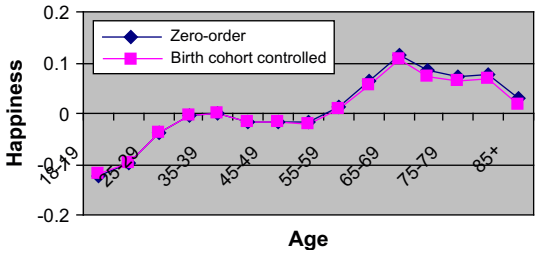

Fig. 4. Regression (unstandardized) of happiness score on age dummy variables, males, combined data from 1972 to 2006 American General Social Surveys.

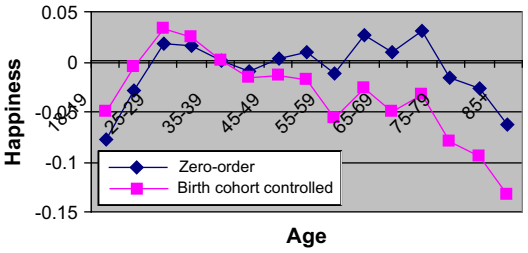

Fig. 5. Regression (unstandardized) of happiness score on age dummy variables, females, combined data from 1972 to 2006 American General Social Surveys.

by Blanchflower and Oswald) are shown for males and females, respectively, in Figs. 4 and 5. For males, controlling birth cohort changes the data very little from the bivariate relationship but does show a slight dip in happiness in the forties and early fifties. However, the data do not come close to a U-shaped pattern, and the lowest happiness by far is indicated for the youngest adults. For females, controlling cohort makes considerably more difference, but again, the pattern of the data is not U-shaped, with the lowest indicated happiness being by far in advanced old age and with that for the youngest adults being lower than that for persons in their forties. The pattern is closer to an inverse U-shape than to a U-shape.

The estimates with cohort controlled should not be considered precise, because period is not controlled and thus any period effects are confounded with age and cohort effects in the estimates. As I point out above, there apparently were no appreciable net period effects, but there probably were small ones given that the age-adjusted happiness of males increased slightly from 1972 through 2006 while that of females decreased by about the same amount. (Regressing happiness on year of survey with the age dummy variables in the equation yields unstandardized coefficients of 0.001 and −0.001, for males and females, respectively, both coefficients being significant at the 0.01 level on a two-tailed test.) These changes probably resulted primarily from cohort succession, but if more than a negligible portion resulted from period influences on adults, the two lines in Figs. 4 and 5 should be slightly closer together than they are because cohort is "over-controlled." However, the estimates in the figures should be reasonably accurate.

Confidence in the rough accuracy of the estimates is reinforced by the intra-cohort trend data shown in Figs. 6 and 7. I compiled these data by taking two 10-year birth cohorts (1945–1954 and 1944–1964) and computing the mean level of happiness within each for each survey year starting when the cohort was ages 18–27 up until the ages when the yearly samples sizes became too small for reliable estimates. I then computed 5-year running means to smooth out some of the fluctuation due to sampling variability and identified the age ranges on the X axis of the figure using the middle year of each 5-year range used for the running means.

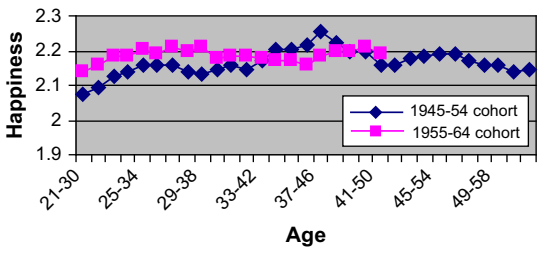

Fig. 6. Mean happiness score for males, 1945–1954 and 1955–1964 birth cohorts, by age, combined data from 1972 to 2006 American General Social Surveys.

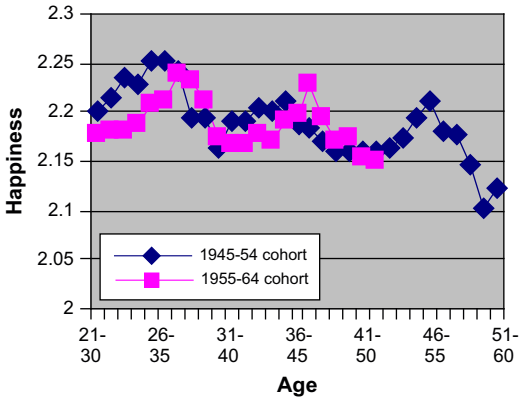

Fig. 7. Mean happiness score for females, 1945–1954 and 1955–1964 birth cohorts, by age, combined data from 1972 to 2006 American General Social Surveys.

These data should be viewed with caution, because they reflect period as well as age effects (though the former should be small), and because the number of respondents in some cohorts for some survey years is not much more than 100. Nevertheless, it is important that the male data in Fig. 6 show the same early sharp increase in happiness as the cross-sectional data and that the female data show a general decline after the late twenties or early thirties and show the lowest level of happiness at the oldest ages represented. Similar intra-cohort happiness trend data reported in Glenn (2005) show essentially the same patterns of change.

A final caution is in order about interpreting both the data I report and those Blanchflower and Oswald report. At the oldest ages, the persons interviewed represent only a small proportion of the original members of their birth cohorts, and these survivors may be systematically different from all of the original cohort members. Specifically, they may on average have had a greater life-long tendency to be happy, and if so, the trend lines in Figs. 4 and 5 from about age 60 up are higher than they would be if they accurately represented average life course change. This brief comment is not the proper place for an in-depth treatment of this issue, but such differential mortality is predicted by both theory and considerable evidence that negative emotions tend to have negative health effects (for a summary of much of this evidence, and also of some contrary evidence, see Zautra, 2003). This means that the real pattern of happiness through the life course in the United States has probably departed even more from the U-shaped pattern than the data in Figs. 4 and 5 indicate.

Discussion

It is very important that the analyses reported by Blanchflower and Oswald be amended so that they provide estimates of what has really typically happened to well-being through the life course in the countries studied. That is what journalists and lay persons want, and that, in my opinion, is what has theoretical and practical utility. If these authors do not want to do that, someone else should do it. If Blanchflower and Oswald think that some kind of *ceteris paribus* approach is useful, they should drop the most questionable control variables from their analyses, explain why they think the resulting data are useful, and take special pains to assure that journalists do not interpret their findings as being descriptively accurate of what has happened or predictive of what is likely to happen in the future.

The analyses that I recommend could be immensely useful. Consider, for instance, the value of knowing how common the huge differences between males and females shown in Figs. 4 and 5 are cross-nationally and the extent to which they exist in regard to well-being variables other than happiness. Such findings would be especially relevant to the work of scholars and researchers in gender studies and almost certainly would stimulate a great deal of follow-up research.

My critique has relevance for kinds of research other than that dealing with age and well-being. A reviewer of an earlier version of this comment noted that the control variables used by Blanchflower and Oswald are commonly used in well-being research. That is true, and it is also true that they are commonly used in a routine fashion in other kinds of social scientific research. However, common practice is not necessarily good practice. It seems to me that there should be an explicit purpose for each control variable used in research, that there should be a clear specification of how each control variable fits in a causal model, and that it is not a good idea to use a bundle of control variables just because others have used them.

Social scientists increasingly realize that a valid explanation of an association between two variables solely or largely by one-way causation is not as common as we once thought. This realization has led to increasingly complex causal models, as it should, but there are occasionally happy instances, such as the one addressed by this critique, in which the best way to deal with the complexity of reality is to change to a simpler analysis.

References

Blackman, L., Clayton, O., Glenn, N., Malone-Colon, L., & Roberts, A. (2005). *The consequences of marriage for African Americans: A comprehensive literature review*. New York: Institute for American Values.

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

Converse, P. E. (1976). *The dynamics of party support: Cohort analyzing party identification*. Beverly Hills, CA: Sage Publications.

Firebaugh, G. (1997). *Analyzing repeated surveys*. Thousand Oaks, CA: Sage Publications.

Glenn, N. D. (2001). Is the current concern about American marriage warranted? *Virginia Journal of Social Policy and the Law, 9*, 5–47.

Glenn, N. D. (2003). Distinguishing age, period, and cohort effects. In J. T. Mortimer, & M. J. Shanahan (Eds.), *Handbook of the life course*. New York: Kluwer Academic/Plenum Publishers.

Glenn, N. D. (2005). *Cohort analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Horwitz, A. V., & White, H. R. (1991). Becoming married, depression, and alcohol problems among young adults. *Journal of Health and Social Behavior, 32*, 221–233.

Kessler, R. C., & Essex, M. (1982). Marital status and depression: the importance of coping resources. *Social Forces, 61*, 484–498.

Marks, N. F., & Lambert, J. D. (1998). Marital status continuity and change among young adults: longitudinal effects on psychological well-being. *Journal of Family Issues, 19*, 652–676.

Mason, K. O., Mason, W. M., Winsborough, H. H., & Poole, K. W. (1973). Some methodological issues in the cohort analysis of archival data. *American Sociological Review*, 38, 242–258.

Nock, S. L. (1998). *Marriage in men's lives*. New York: Oxford University Press.

Stack, S., & Eshleman, J. R. (1998). Marital status and happiness: a 17-nation study. *Journal of Marriage and the Family, 60*, 527–536.

Stutzer, A., & Frey, B. S. (2006). Does marriage make people happy or do happy people get married? *Journal of Socio-Economics, 35*, 326–347.

Waite, L. J. (1995). Does marriage matter? *Demography, 32*, 483–497.

Waite, L. J., & Gallagher, M. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially*. New York: Doubleday.

Zautra, A. J. (2003). *Emotions, stress, and health*. New York: Oxford University Press.