Why life satisfaction is U-shaped in age?*

Oleksandr Movshuk†

August 26, 2011

Abstract

While it is common to estimate the age effect on happiness with parametric terms (such as age and age-squared), this paper applies a semiparametric regression model that imposes much milder restrictions on the shape of age-happiness profile. Using data from the British Household Panel Study, the paper confirmed that the age-happiness profile remained U-shaped in the semiparametric specification as well. To trace sources of the U-shape, I applied the varying coefficient model of Hastie and Tibshirani (1993), and estimated differentiated age-profiles for major life circumstances, such as differences in marital status, the presence of children at home, and whether children were dependent or not. After introducing these differentiated age effects, the shape of the overall age-happiness profile changed in a notable way: it was no longer U-shaped, but flat. In contrast, the differentiated age profiles continued to have significant effects on happiness for most life circumstances. Evidently, the common finding that happiness is U-shaped in age may reflect a composite effect from differences in people's adjustment to major life circumstances over the life cycle.

1. Introduction

This paper examines a puzzling result from recent studies of life satisfaction by economists. Starting from the seminal paper by Blanchflower and Oswald (2004), a consensus has emerged that life satisfaction is U-shaped in age, with the lowest point at age in early 40s. The pattern was confirmed in a large number of countries by Blanchflower and Oswald (2008). The finding is unexpected, since a common view in other social sciences (such as psychology) is that age have little effect on happiness (Frijters and Beatton, 2008). Moreover, little progress has been made in explaining why the age effect on happiness is U-shaped (Stone *et al.*, 2010).

Initially, the U-shaped profile of life satisfaction was attributed to omitted cohort effect. For instance, Blanchflower and Oswald (2004, p. 1380) speculated that the estimate of increasing life satisfaction in the old age may be due to omitted cohort effects from generations that experienced the misery of the Second World War. During the post-war years, life experience greatly improved for subsequent birth cohorts, and increased life satisfaction in these birth cohorts. The possible role of cohort effects on happiness was examined by Clark (2007) and Blanchflower and Oswald (2008), but even after adding cohort effects, the U-shape changed little.

The addition of cohorts effects created a new problem that remains largely unsolved. Typically, cohort effects are used in specifications that already include age and period effects. This creates a simultaneity problem among these three effects, since they are linked by an exact identity ($current\ year - age = year\ of\ birth$). The simultaneity problem in age-period-cohort (APC) models has long been known in economics (Heckman and Robb, 1985), and it can be solved by imposing restrictions on parameter estimates of the APC effects. When Clark (2007) and Blanchflower and Oswald (2008) specified APC models for life satisfaction, they used a restriction that regression parameters for APC effects consist of different time blocks. For example, Clark (2007) represented age effects with 5-year age blocks, and left cohort and period effects unrestricted (as one-year dummy variables). A similar restriction was used by Blanchflower and Oswald (2008). However, the solution of the simultaneity problem was questioned by de Ree and Alessie (2011), who showed that slight modifications in the structure of time blocks may greatly change estimates of age effect. For example, the study demonstrated that when time blocks for age effect were modified from five to two years, life satisfaction was no longer U-shaped.

*Data from the British Household Panel Survey (BHPS) were made available through the Economic and Social Research Council (ESRC) Data Archive. I would like to thank Christoph Wunder for sharing STATA code for the BHPS data.

†Dept. of Economics, University of Toyama, 3190 Gofuku, Toyama, 930-8555, Japan. E-mail:movshuk@eco.u-toyama.ac.jp

In this paper, I apply a different solution to the simultaneity problem in APC models of life satisfaction, which specifies the age effect by a nonparametric term. The approach have been suggested by Wunder *et al.* (2009), who applied it to British and German data on life satisfaction. The present study uses a different semiparametric estimator of Marra and Wood (2011) with variable selection among nonparametric terms. In the limit, the estimator can reduce redundant terms to zero functions, effectively selecting them out of the model. Applying the estimator to data on life satisfaction from the British Panel of Household Survey (BHPS), I demonstrate that the age effect on life satisfaction preserves its U-shape in the semiparametric model, even after taking into account cohort effects.

Another contribution of this paper is an explanation for the U-shape in life satisfaction[1]. Blanchflower and Oswald (2004) speculated "that this decline and then rise in well-being through the years may reflect a process of adaptation to circumstances; perhaps, by the middle of their lives, people relinquish some of their aspirations and thereby come to enjoy life more" (p. 1375). To verify this hypothesis, I applied a semiparametric varying coefficients model by Hastie and Tibshirani (1993) with the model-selection algorithm of Marra and Wood (2011) to estimate differentiated age effects for major life circumstances (such as marriage, divorce, widowhood, the presence of children at home, employment, unemployment, and so on). Once these differentiated age profiles for life circumstances were introduced, the overall age–happiness profile was no longer U-shaped, but flat (effectively, it was selected out of the model). On the other hand, most differentiated age effects displayed significant variability over the life cycle. However, when the sample was split into sub-samples of men and women, the overall age–happiness profile was not completely explained away by differentiated age effect. Still, the magnitude of U-shape became reduced by around one-half, especially for the sub-sample of men. Another noteworthy finding is significant asymmetry between men and women in estimates of differentiated age profiles for some life circumstances (especially for presence of dependent children and unemployment)[2].

2. Previous literature

Despite the rapidly growing interest among economists about determinants of happiness, most papers have examined only the general effect of age on happiness, and few studies discussed interactions between the age effect and other life circumstances (such as children at home, or unemployment). In a comprehensive review, MacKerron (in press) concluded that little is known how the impact of life events may interact with age, and pointed that the effect from children on subjective well-being (SWB) may be different with age (p. 26). Similarly, Stone *et al.* (2010) speculated that the effect from children on SWB may change over the live cycle, and may increase SWB of parents after children leave home, due to reduced levels of family conflict and smaller financial burden (p. 9986). Another interaction between age effect and widowhood was mentioned by de Ree and Alessie (2011), who pointed that the negative effect of widowhood may be especially traumatic in the young age (p. 182). Finally, some studies considered how the negative effect of unemployment may change over the life span. Clark and Oswald (1994, p. 657) reported estimates that the negative effect of unemployment on life satisfaction was relatively large at age between 30 and 49, while the the youngest individuals experienced relatively minor (but still negative) effects. Pichler (2006) reached a similar conclusion that negative effect from unemployment was less severe in the young age compared with the middle or old ages (p. 432).

On the whole, the previous studies were either just postulating possible interactions between the age effect and various life circumstances (like children at home or widowhood), or examined these interaction empirically, but for only a single life circumstance or event, with major interest focusing on age-related effects of unemployment. The interaction between the age effect on happiness and *several* life circumstances was examined by Yang (2008). The study estimated differentiated age profiles by three categorical variables (gender, race and education level), but considered only a parametric specification for the age effect (by age and age squared). Another study on the role of various life events for SWB was done by Clark *et al.* (2008). The study examined how marriage, divorce, widowhood, birth of child, unemployment and layoffs affected SWB before and after these life events. Though the present paper also deals with the effect from some of these life events (such as differences in marital and employment status), there is one important difference. While Clark *et al.* (2008) focused on the temporal variation in life satisfaction before and after the life events, the present paper focuses on inter-group differences between people that experience different life circumstances (such as being employed, unemployed, or out of labor force). With estimates

[1] I do not make distinction among 'life satisfaction', 'happiness', and 'subjective well being', which is a common approach among economists (MacKerron, in press).

[2] Due to the space limitation, the results by gender had to be omitted, but are available upon request.

of differentiated age effects, the present paper aims to evaluate whether alternative life circumstances have different effects on life satisfaction over the life cycle.

3. Model specification and estimation method

I apply the approach of Blanchflower and Oswald (2004), and postulate an experienced personal utility $U_{i,t}$ for individual i at time t that depends of a vector of personal and demographic characteristics $\boldsymbol{x}_{i,t}$, with $U_{i,t} = u(\boldsymbol{x}_{i,t})$. The utility $U_{i,t}$ is known only to the individual i, who reports it as reported happiness R, which is a function of $U_{i,t}$: $R_{i,t} = r(U_{i,t})$, or $R_{i,t} = r(u(\boldsymbol{x}_{i,t}))$. The reported happiness $R_{i,t}$ depends on $\boldsymbol{x}_{i,t}$ through parametric and nonparametric effects in a semiparametric regression model $R_{i,t} = r(u(\boldsymbol{x}_{i,t})) + \varepsilon_{i,t}$, where $\varepsilon_{i,t}$ is a conventional disturbance term.

The vector of explanatory variables $\boldsymbol{x}_{i,t}$ includes income, age (specified as a smooth nonparametric term $s(age)$), time t and other demographic and personal characteristics. Cohort and period effects are estimated by sets of dummy variables D_c and D_t: $R_{i,t} = s(age_{i,t}) + \alpha_c' D_c + \alpha_t' D_t + \beta' \boldsymbol{x}_{i,t} + \varepsilon_{i,t}$.

To estimate age-happiness profiles that vary by life circumstances, I applied the varying-coefficient model of Hastie and Tibshirani (1993). Consider differences by work status, with the following three categories: 'employed', 'unemployed', and 'out of labor force'. Let d_1, d_2, and d_3 denote dummy variables for these three categories. In the varying-coefficient model, the shape of nonparametric age effect $s(age)$ depends on values of d_k: $R = s(age) + \sum_{k=1}^{3} s(age) d_k + \alpha_c' D_c + \alpha_t' D_t + \beta' \boldsymbol{x} + \varepsilon_{c,t}$.

For example, $s(age)d_1$ measures differentiated effects of age that depend on whether d_1 is either zero or one. Note that $s(age)$ in this model represents a nonparametric estimate of the *overall* age effect, while estimates of $s(age)d_k$ estimate differences (or differentials) from $s(age)$ as a result of belonging to kth life circumstance. Importantly, the model is not affected by the 'dummy-variable trap', since all nonparametric effects $s(age)d_k$ are nonlinear, so the model may include differentiated terms for all categories.

These semiparametric models were estimated by the modified generalized cross-validation (*mgcv*) library (Wood, 2011) in the *R* statistical package (R Development Core Team, 2011).

4. Data

I used data from 10 waves of the the British Household Panel Survey (BHPS). Life satisfaction was measured from responses to the following question: 'How dissatisfied or satisfied are you with your life overall', and seven possible answers range from 'not satisfied at all' to 'completely satisfied'[3].

5. Results

This section reports regression estimates only for the whole sample with 63,402 person-year observations. Table 1 reports estimates for six alternative regression models. They mainly differ in the composition of nonparametric components, which are shown in the lower part of Table 1. Model 1 is a baseline that mimics typical specifications in the previous studies. In particular, the model has only parametric terms, including age and age-squared for the age effect. The latter estimates indicate a U-shape in age, with the lowest happiness at age 41. In Model 2, age is modeled by a nonparametric term $s(age)$, and the model is estimated by the *mgcv* algorithm. Though Models 1 and 2 used different estimation methods, estimates in their parametric components were similar. Consider estimates for Model 2. Health status had the largest effect on happiness: compared with poor health, excellent health raised happiness by 0.987 (on the scale from 1 to 7). The next largest effects were for differences in work and marital status. Unemployed individuals had lower happiness by 0.329 points (compared with employed individuals). Similarly, marriage raised happiness by 0.303 points (compared with those who were never married), while divorce, widowhood or separation from partner reduced happiness by 0.212 points. The effect of children at home was negative, but the estimate had t-ratio of only 1.72.

The effect of income on happiness was examined by 10 income deciles. The 1st decile of the lowest income was the reference category. In 2nd and 3rd deciles, the level of happiness was similar to 1st decile, with t-ratios of only 0.69 and 0.91, respectively. The effect from income remained relatively minor even in the middle of income distribution. For example, for incomes in 5th and 6th deciles, happiness did not rise much compared with 1st decile, by only 0.063 and 0.116 points, respectively. Moreover, even at the highest level of income in 10th decile, the extra effect on happiness from 6th decile continued to be small, rising by just $0.180 - 0.116 = 0.064$ points. As for the nonparametric estimate of age effect $s(age)$, it produced a relatively complex nonlinear pattern, as shown by number of estimated degrees of freedom for the term

[3]More details about the survey are available at http://www.iser.essex.ac.uk/survey/bhps/

(7.03). This nonparametric estimate is plotted in top-left panel of Figure 1. A clear U-shaped pattern is evident, with the minimum level of well-being at age 46.

Model 3 includes cohort effects on happiness. The model's estimate of $s(age)$ is plotted in the top-right panel of Figure 1. Evidently, even with cohorts effects, happiness was still U-shaped in age. The bottom of Figure 1 show estimates of cohort and period effects. Estimates for the cohort effect were volatile, and were in addition smoothed to identify a clear pattern. Across different birth cohorts, well-being was higher for those born in the 1930s and 1940s. Estimates of the period effect indicate a decreasing trend over 1996-2006, but the magnitude of the change was relatively small compared with the variation in the age effect.

Starting from Model 4, nonparametric effect include differentiated age effects for various life circumstances. The composition of these differentiated age effects can be read from the lower part of Table 1. For example, Model 4 contains two differentiated effects for children at home, denoted by *s(age):no child* and *s(age):child*. Model 6 has the largest number of differentiated age effects on happiness. First, it subdivides the category 'children at home' into dependent and independent children, and second, introduces three differentiated age effects for marital status ('never married', 'married', and a composite group of 'separated/divorced/widowed' individuals).

Figure 2 shows in panel A the overall age profile $s(age)$, along with three differentiated effects of children and marital status, shown in panels B.1-B.3 and C.1-C.3, respectively. It is noteworthy that after introducing the differentiated age effects in Model 6, the estimate for the overall age effect $s(age)$ was no longer U-shaped, but flat (essentially, the term was selected out from the model). Evidently, the U-shape in happiness was explained away by differentiated age effects for presence of children at home, and for three categories of marital status. Some of these differentiated estimates closely matched predictions from previous literature, which were mentioned in section 2.. For example, the effect of 'empty nest' after children left home is evident in relatively high happiness in old age (panel B.1), while the trauma from divorce or widowhood is relatively severe in young age (panel C.3), especially when compared with estimates for old age.

6. Conclusion

This paper examined variation in life satisfaction in the United Kingdom with a semiparametric regression model, and reached two conclusions. First, the U-shape in life satisfaction changed little even after controlling for cohort effects. Second, the source of U-shape in happiness appears to be variation in people's response to different life circumstances over the life cycle. Once this variation is taken into account, the U-shape in overall age effect was greatly reduced, and eventually became flat. On the other hand, differentiated age effects for most life circumstances remained significant over the life cycle.

References

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

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

Clark, A. E. (2007) Born to be mild? Cohort effects don't (fully) explain why well-being is u-shaped in age, IZA Discussion Papers 3170, Institute for the Study of Labor (IZA).

Clark, A. E., Diener, E., Georgellis, Y. and Lucas, R. E. (2008) Lags and leads in life satisfaction: a test of the baseline hypothesis, *Economic Journal*, **118**, F222–F243.

Clark, A. E. and Oswald, A. J. (1994) Unhappiness and unemployment, *Economic Journal*, **104**, 648–59.

de Ree, J. and Alessie, R. (2011) Life satisfaction and age: Dealing with underidentification in age-period-cohort models, *Social Science & Medicine*, **73**, 177 – 182.

Frijters, P. and Beatton, T. (2008) The mystery of the U-shaped relationship between happiness and age, NCER Working Paper Series 26, National Centre for Econometric Research.

Hastie, T. and Tibshirani, R. (1993) Varying-coefficient models, *Journal of the Royal Statistical Society. Series B (Methodological)*, **55**, 757–796.

Heckman, J. J. and Robb, R. (1985) Using longitudinal data to estimate age, period and cohort effects in earnings equations, in *Cohort Analysis in Social Research : Beyond the Identification Problem* (Eds.) W. M. Mason and S. E. Fienberg, New York: Springer-Verlag, pp. 137 – 150.

MacKerron, G. (in press) Happiness economics from 35000 feet, *Journal of Economic Surveys*, accepted for publication.

Marra, G. and Wood, S. N. (2011) Practical variable selection for generalized additive models, *Computational Statistics and Data Analysis*, **55**, 2372–2387.

Pichler, F. (2006) Subjective quality of life of young Europeans. Feeling happy but who knows why?, *Social Indicators Research*, **75**, 419–444.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Stone, A. A., Schwartz, J. E., Broderick, J. E. and Deaton, A. (2010) A snapshot of the age distribution of psychological well-being in the United States, *Proceedings of the National Academy of Sciences*, **107**, 9985–9990.

Wood, S. N. (2011) *Package 'mgcv' (version 1.7-6).*, http://cran.r-project.org/web/packages/mgcv/mgcv.pdf.

Wunder, C., Wiencierz, A., Schwarze, J., Kchenhoff, H., Kleyer, S. and Bleninger, P. (2009) Well-being over the life span: semiparametric evidence from British and German longitudinal data, IZA Discussion Papers 4155, Institute for the Study of Labor (IZA).

Yang, Y. (2008) Social Inequalities in Happiness in the United States, 1972 to 2004: An Age-Period-Cohort Analysis, *American Sociological Review*, **73**, 204–226.

Table 1. Regression estimates for semiparametric models

	Model 1		*Model 2*		*Model 3*		*Model 4*		*Model 5*		*Model 6*	
	Estimate	t-value	Estimate	t-value	Estimate	t-value	Estimate	t-value	Estimate	t-value	Estimate	t-value
Intercept	5.816	94.70	4.252	120.35	4.454	32.90	4.444	32.78	4.491	32.88	4.484	32.88
Age	-0.084	30.28										
Age^2	0.001	30.72										
Married/couple	0.314	21.51	0.303	23.11	0.298	19.77	0.290	18.81	0.284	15.72	0.284	15.26
Widowed, separated	-0.201	9.75	-0.212	11.12	-0.211	10.09	-0.215	10.27	-0.252	8.52	-0.250	8.45
Have child	-0.016	1.53	-0.018	1.72	-0.014	1.30	-0.013	1.19	-0.017	1.55		
Have dependent child											-0.039	2.79
Have independent child											-0.007	0.50
Female	0.047	5.01	0.048	5.17	0.050	5.30	0.048	5.10	0.048	5.15	0.047	4.99
Unemployed	-0.331	12.67	-0.329	12.60	-0.329	12.59	-0.327	12.53	-0.325	12.43	-0.324	12.39
Not in labor force	-0.107	8.68	-0.117	9.46	-0.120	9.63	-0.121	9.67	-0.122	9.77	-0.122	9.76
Education: medium	-0.070	5.36	-0.073	5.69	-0.065	4.93	-0.066	5.01	-0.066	5.01	-0.066	5.00
Education: high	-0.082	6.39	-0.082	6.45	-0.075	5.80	-0.076	5.91	-0.076	-5.88	-0.074	5.75
Health: good	0.638	58.15	0.636	57.94	0.634	57.86	0.633	57.77	0.633	57.69	0.633	57.71
Health: excellent	0.990	76.50	0.987	76.22	0.983	75.93	0.982	75.81	0.981	75.72	0.982	75.72
House: owned outright	0.198	12.72	0.196	13.08	0.185	11.89	0.186	11.94	0.185	11.82	0.187	11.86
House: owned with mortgage	0.109	8.56	0.104	8.15	0.099	7.70	0.101	7.87	0.098	7.60	0.101	7.81
Income: decile 2	-0.015	0.73	-0.014	0.69	-0.018	0.86	-0.014	0.69	-0.016	0.77	-0.014	-0.67
Income: decile 3	0.019	0.93	0.019	0.91	0.016	0.78	0.020	0.97	0.018	0.88	0.020	0.95
Income: decile 4	0.047	2.22	0.049	2.32	0.046	2.18	0.050	2.39	0.049	2.30	0.049	2.32
Income: decile 5	0.059	2.79	0.063	2.97	0.061	2.84	0.066	3.11	0.065	3.02	0.065	3.01
Income: decile 6	0.111	5.19	0.116	5.46	0.112	5.25	0.119	5.53	0.117	5.44	0.116	5.39
Income: decile 7	0.116	5.44	0.124	5.80	0.121	5.68	0.129	6.01	0.127	5.93	0.127	5.84
Income: decile 8	0.126	5.86	0.133	6.22	0.128	5.99	0.137	6.34	0.135	6.22	0.134	6.11
Income: decile 9	0.139	6.41	0.146	6.73	0.141	6.51	0.151	6.92	0.149	6.82	0.148	6.66
Income: decile 10	0.175	7.95	0.180	8.18	0.176	7.96	0.188	8.39	0.185	8.24	0.183	8.08
Cohort effect	no		no		yes		yes		yes		yes	
Estimated degrees of freedom for nonparametric effects												
s(age)			7.03**		6.75**		6.79**		0.01		0.01	
s(age):no child							2.28**		2.59*		2.95**	
s(age):have child							0.04		0.12			
s(age):have dependent child											0.77	
s(age):have independent child											0.86*	
s(age):not married									6.12**		6.12**	
s(age):couple									4.66**		4.37**	
s(age):{divorced,widowed, separated}									6.04**		6.05**	
R-sq /Deviance explained	0.150		0.151		0.156		0.157		0.157		0.157	
AIC score	218,415		218,353		218,076		218,067		218,040		218,033	
Sample size	43,402		43,402		43,402		43,402		43,402		43,402	

Notes: t-ratios are in parentheses; statistically significant estimates at the level of 10, 5, and 1 percent are shown with *, **, ***, respectively; all models also included year and region models; Models 2-6 also include cohort dummies; reference categories are (1) never married; (2) no children at home; (3) male (4) employed, (5) education: low (6) health: poor (7) house: renting, (8) income: decile 1.

Figure 1. Estimates of age, cohort and period effects on life satisfaction

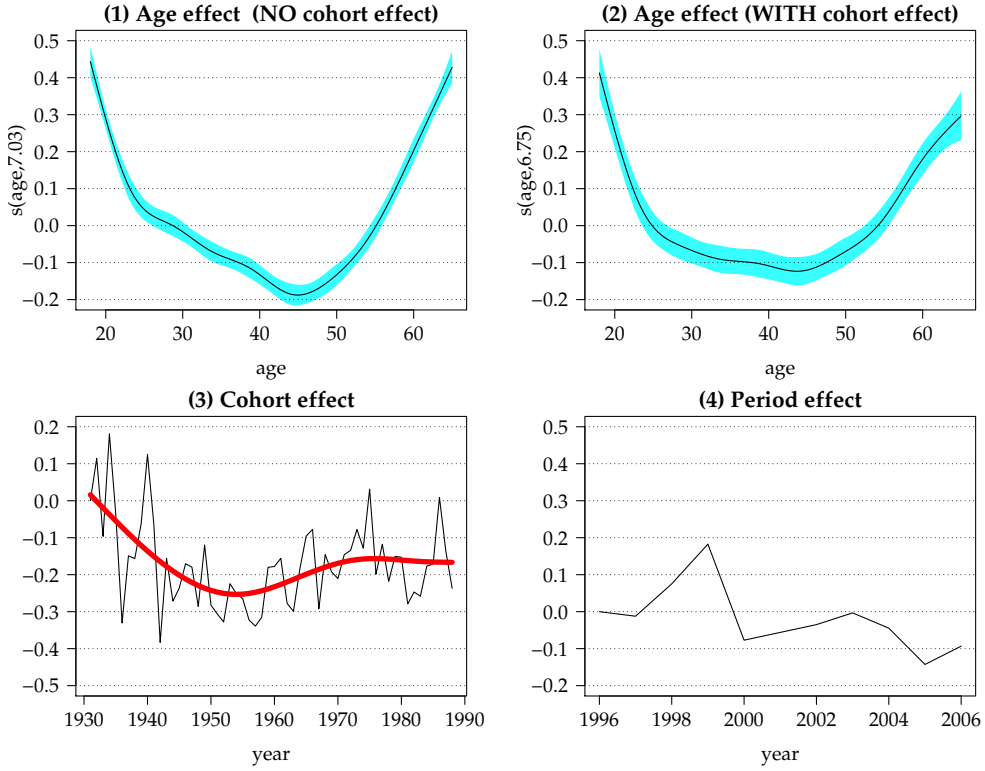

Figure 2. Estimates of the overall and differentiated age effects for Model 6.