Empirical Economics (2005) 30:749–761
DOI 10.1007/s00181-005-0255-7

EMPIRICAL ECONOMICS

Subjective well-being and the family: Results from an ordered probit model with multiple random effects

Rainer Winkelmann

Socioeconomic Institute, University of Zurich, Zürichbergstr. 14, 8032 Zürich, Switzerland
(e-mail: winkelmann@sts.unizh.ch)

First version received: May 2003/Final version received: August 2004

Abstract. The previous literature on the determinants of individual well-being has failed to fully account for the interdependencies in well-being at the family level. This paper develops an ordered probit model with multiple random effects that allows to identify the intra-family correlation in well-being. The parameters of the model can be estimated with panel data using Maximum Marginal Likelihood. The approach is illustrated in an application using data for the period 1984–1997 from the German Socio-Economic Panel in which both inter-generational and intra-marriage correlations in well-being are estimated.

Key words: Ordered probit model, error components, German Socio-Economic Panel

JEL classification: C23, C25, I31, J19

1. Introduction

Increasing well-being arguably should be the ultimate objective of economic and social policy. "Economic things matter only in so far as they make people happier" (Oswald 1997, p. 1815). Well-being is also one of the more elusive concepts of economic research. One only needs to mention the century old debate about the appropriate approaches to individual utility and welfare comparisons. And yet, a recent but rapidly increasing literature (surveyed in Frey and Stutzer 2001, among others) raises the hope that some real progress

I would like to thank Andrew Clark, Daniel S. Hamermesh, seminar participants at Tilburg University, and three anonymous referees for helpful comments.

can be made in understanding what people value and what makes them happy. The approach taken by this literature is empirical, taking at face value responses by individuals to questions of the type "On the whole, how satisfied are you with your life at present?", where individuals can respond on a scale from 0 (completely dissatisfied) to 10 (completely satisfied). Questions of this sort are routinely included in many current general purpose household surveys. I will refer to such responses henceforth as indicators of *subjective well-being* (SWB) in lieu of (but in principle interchangeable with) "happiness" or "general satisfaction". The economic interpretation of measurements of SWB is that a higher level of SWB corresponds to a higher utility level.

To understand the determinants of SWB, regression analysis has been the single most important tool. Yet, many factors that presumably are important determinants of SWB are inherently unobservable. Hence, latent structure models offer a promising avenue for future research, and this paper presents a step in this direction. In particular, it will be estimated how family background, as a latent factor, influences SWB of the individual.

This approach is in line with recent research by Ferrer-i-Carbonel and Frijters (2001) who provide persuasive evidence for the importance of time invariant factors in explaining SWB. They rightly conclude that the search for explanations for the sources of variation in "intrinsic" well-being should be at the forefront of further research in this area. Here, I explore the potential contribution that the family has to make in shaping the intrinsic well-being of an individual.

Family has been used before as explanatory variable in SWB regressions. Typical SWB equations include controls for marital status. For instance, divorce ceteris paribus has a very large negative contemporaneous effect on individual well-being (e.g., Clark and Oswald 2002). Winkelmann and Winkelmann (1995) went one step further by allowing interactions of SWB between family members. In particular, they found that male unemployment has a substantial negative effect, in terms of reduced SWB, on the female spouse, an instance of an externality that increases the social cost of unemployment. However, such an approach still falls short of modelling SWB within the family as a fully interdependent process. How does the SWB of one family member affect the SWB of others? A preferred way to model the interdependencies of preferences is by studying the joint distribution of SWB within the family.

The contribution of this paper is primarily methodological. It shows how family effects can be identified in an ordered probit model of SWB, and how the parameters of the model can be estimated by Maximum Marginal Likelihood. The application to data from the German Socio-Economic Panel for the years 1984–1997 demonstrates the methodology in a model of SWB for fathers, mothers, and children. The estimated correlations can be interpreted as the combined effects of a number of distinct factors, such as genes, nurture, and shared general economic conditions of the family. Disentangling these effects would require more differentiated samples, such as samples of siblings or twins living in different families, while the methodology would remain the same.

2. Related literature

The concept of a family used in this paper is based in comparisons of parents and their children while the children still live at home. In the empirical part, only children older than 16 are included, since only they fill out an individual questionnaire. Thus, the scope is somewhat different from a related prior literature in psychology, where the focus was on the specific contribution of heritability in explaining variation of SWB between individuals. This literature concentrates on comparisons of monozygotic and dizygotic twins. Because monozygotic twins share all of their genes, whereas dizygotic twins share only half of their genes on average, comparisons of different sets of twins (those who were reared together and others who grew up separately) potentially identify the effect of genes and early family environment on SWB. Diener and Lucas (1999) cite a study by Tellegen et al. (1988) according to which monozygotic twins are extremely similar in terms of SWB, regardless of whether they were reared together or apart. On the other hand, dizygotic twins were on average far less similar. Tellegen et al. (1988) estimated that genetics account for 48 percent of the variability in well-being. Diener and Schwarz summarize other twin studies and conclude (p. 217) that "there are consistent and stable individual differences in SWB that are, to some degree, inherited."

The reliance on twin studies in such context, despite its obvious advantages, has a number of drawbacks as well. First, sample sizes tend to be rather small, certainly much smaller than those found in economically motivated SWB studies based on general surveys. To give one example, Blanchflower and Oswald (2004) report regressions with as much as 50 thousand observations. Secondly, it is an open question, whether findings for twins are representative for the wider population. One particular issue, raised first by Solon (1989), is that of homogeneity of samples of twins (for instance with respect to age, education, and employment) relative to random samples of the population. Under the assumption of his model, such homogeneity will tend to understate the correlation between siblings. However, greater homogeneity is likely to go hand in hand with greater homogeneity of the environment. As a consequence, the relative contribution of heritability may well be overstated. As Diener and Schwarz (1999, p. 217) put it, "if the environment were completely uniform, intelligence would be 100% inheritable."

A good case can be made to extend the analysis of family correlation to more heterogeneous datasets, or even better, representative samples of the population. In some ways, the germ of such an extension is already contained in papers by Groot and Maassen van den Brink (2003) and Shields and Wheatley Price (2001). However, as we will see, their analysis suffers from a limitation to cross section data, where long-term effects cannot be identified. Moreover, in the case of Groot and Maassen van den Brink, the interest centers around the match-specific value of marriage rather than on the more general concepts of family or kinship effects.

3. Modelling intra-family correlation

SWB responses are observed on an ordinal scale. To link those responses to observed and unobserved factors, we formulate a latent model for an underlying cardinal SWB level. Let

$$y^*_{ijt} = x'_{ijt}\beta + \varepsilon_{ijt}, \tag{1}$$

where y^*_{ijt} denotes the latent SWB of the ith member in family j at time t. x_{ijt} is a vector of exogenous variables that includes a polynomial in age and a gender dummy while β is a conformable vector of coefficients. The family-index j is used in this section as a generic index for group membership of the individual. Depending on the application, this can be a family, household, siblings, or more generally any group.

The model assumes that the error term is of the following composite nature (the following presentation follows closely the models of intergenerational earnings mobility as developed in Björklund and Jäntti 1997, Björklund et al. 2002, or Solon et al. 1991):

$$\varepsilon_{ijt} = a_j + u_{ij} + v_{ijt}. \tag{2}$$

a_j is a family specific random effect that does not vary across persons within the family or over time. u_{ij} is an individual specific random effect that does not vary over time. v_{ijt} is a white noise error term. The first two error components capture long-term effects, the last short-term effets. The model has a hierarchical structure. Individuals are part of a particular family j. This is a natural assumption in applications where the family effect is interpreted as a proxy for family background shared by parents and children, or between siblings. In other contexts, it might be more natural to define the individual effect independently of the family, i.e., use u_i rather than u_{ij}. In this situation, the individual keeps the individual effect but changes the family effect, when family affiliation changes, such as is the case when children move out, marry and have their own children.

The three error components are assumed mutually independent and distributed with mean zero and constant variances σ^2_a, σ^2_u, and σ^2_v (later normalized to 1), respectively. It follows that the variance of ε_{ijt} is given by

$$\sigma^2_\varepsilon = \sigma^2_a + \sigma^2_u + \sigma^2_v.$$

Moreover, the covariance between observations at two different points in time, t and s, is given by

$$\mathrm{Cov}(y^*_{ijt}, y^*_{ijs} | x_{ijt}, x_{ijs}) = \mathrm{Cov}(\varepsilon_{ijt}, \varepsilon_{ijs}) = \sigma^2_a + \sigma^2_u \qquad t \neq s$$

if the same individual is considered, and by

$$\mathrm{Cov}(y^*_{kjt}, y^*_{ijs} | x_{ijt}, x_{ijs}) = \mathrm{Cov}(\varepsilon_{kjt}, \varepsilon_{ijs}) = \sigma^2_a \qquad t \neq s, k \neq i$$

if two individuals within the same family are considered. The *long-term* within-family correlation (for any $t \neq s$) in SWB is

$$\rho = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_u}. \tag{3}$$

It is long-term, because it is by assumption purged of the transitory error component and thus exists only due to a family specific time invariant component. By contrast, the overall correlations in latent SWB,

$$\mathrm{Corr}(\varepsilon_{ijt}, \varepsilon_{ijs}) = \frac{\sigma^2_a + \sigma^2_u}{\sigma^2_a + \sigma^2_u + \sigma^2_v} \tag{4}$$

and

$$\mathrm{Corr}(\varepsilon_{ijt}, \varepsilon_{kjs}) = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_u + \sigma^2_v} \tag{5}$$

depend on the variance of the transitory error. The larger the transitory fluctuations, the smaller the correlation. This is the well known "regression to the mean" phenomenon. Attempts of estimating long-term correlations in SWB or economic status from cross-section data result in downward bias.

The goal of the present analysis is a decomposition of the long-term correlation into a part that is shared between members of the same family (the family effect) and a part that is specific to the individual (the individual effect) as given by (3). The family effect measures the correlation in long-run well-being between individuals of the same family. The correlation can be identified if repeated measurements (panel data) are available for different members of the same family.

The error components approach is closely related to an alternative regression approach, where y_{ij}, the outcome of individual i in family j, is regressed on y_{kj}, the outcome of individual k in the same family j. Assume that a measure in long-run well-being is available. Then

$$y_{ij} = \alpha y_{kj} + \xi_{ij},$$

where $y_{kj} = a_j + u_{kj}$. The population least squares coefficient α measure the correlation in long-term well-being and is equal to ρ in Eq. (3). Now assume that y_{ij} is unobserved. Rather, we observe $y_{ijt} = y_{ij} + v_{ijt}$. This is a typical measurement error problem. The least squares coefficient in a regression of y_{ijt} on y_{kjt} has probability limit

$$\begin{aligned}\tilde{\alpha} &= \frac{\mathrm{Cov}(y_{ijt}, y_{kjt})}{\mathrm{Var}(y_{ijt})}\\ &= \frac{\mathrm{Cov}(\alpha y_{kjt} - \alpha v_{kjt} + \varepsilon_{ij} + v_{ijt}, y_{kjt})}{\mathrm{Var}(y_{ijt})}\\ &= \frac{\alpha\sigma_\varepsilon^2 - \alpha\sigma_v^2}{\sigma_\varepsilon^2}\\ &= \alpha\left(\frac{\sigma_a^2 + \sigma_u^2}{\sigma_a^2 + \sigma_u^2 + \sigma_v^2}\right).\end{aligned}$$

Thus, as before (see Eq. (5)), using cross-section data alone leads to a downward bias in the estimated family correlation. Otherwise, if measurements of SWB purged by period effects were available, one could obtain the family correlation in one of two ways, either directly through a regression of the sort presented here, or through the appropriate ratio formed by the error variances as before.

The regression approach faces a number of practical problems, though. First, it is not easily extended to the latent variable framework where observed variables are ordinal. Second, long-term realisations are by definition unobserved. One approach would take averages over time. Such averaging is incompatible with the ordinal character of the data. And finally, the regression approach is restricted to pairwise comparisons (such as between fathers and sons or between brothers) whereas in the present application there are multiple observations per family. These disadvantages are absent from the error components model. In the next section, I will describe an ordered probit model with multiple variance components, and its estimation.

4. An ordered probit model with error components

The link between the observed subjective well-being responses (measured in our data on an eleven-point ordered scale) and the latent well-being index is assumed to be of the ordered probit type (McKelvey and Zavoina 1975). In particular, let v_{ijt} be normally distributed with mean 0 and variance 1, and let

$$y_{ijt} = 0 \text{ if } y^*_{ijt} \leq \kappa_1$$

$$y_{ijt} = 1 \text{ if } \kappa_1$$

$$\vdots$$

$$y_{ijt} = 10 \text{ if } y^*_{ijt} > \kappa_{10}.$$

Conditional on the two random effects, a_j and u_{ij}, and x_{ijt}, the probability function for a single observation of the dependent variable can be written as

$$f(y_{ijt}|x_{ijt}, a_j, u_{ij}) = \Phi(\kappa_{y+1} - x'_{ijt}\beta - a_j - u_{ij}) - \Phi(\kappa_y - x'_{ijt}\beta - a_j - u_{ij}),$$

where Φ denotes the cumulative density function of the standard normal distribution, $\kappa_0 = -\infty$ and $\kappa_{11} = \infty$. The remaining ten threshold parameters $\kappa_1, \ldots, \kappa_{10}$ are freely estimated together with β. In this parameterization, identification requires that the x-vector includes no constant.

The model has a hierarchical structure. There are J different families with N_j persons in each family. Clearly, observations are independent across families. Within the family, observations are not independent, since a_j is common to all persons living in the family (and time periods) and u_{ij} is common to all time periods for a given individual. Let $y_j = (y_{1j1}, \ldots, y_{1jT_i}, \ldots, y_{N_jj1}, \ldots, y_{N_jjT_{N_j}})$ denote the vector of responses of all persons living in family j. Define the vector x_j analogously. The joint probability function of $y_j|x_j$ is then given by

$$f(y_j|x_j) = \int_{-\infty}^{\infty} \prod_{i=1}^{N_j} \left[\int_{-\infty}^{\infty} \prod_{t=1}^{T_i} f(y_{ijt}|x_{ijt}, a_j, u_{ij}) h(u_{ij}) \mathrm{d}u_{ij} \right] g(a_j) \mathrm{d}a_j. \tag{6}$$

The intuition behind this expression is as follows. Conditionally on a_j and u_{ij}, observations are independent. The term in brackets marginalizes with respect to the individual specific effect u_{ij} (with density $h(u_{ij})$). What is left is an expression that depends on a_j only. The outer integral then marginalizes in a second step with respect to a_j with density $g(a_j)$. Sequential integration is possible, since a_j and u_{ij} are independent by assumption.

To complete the model, the distributions $h(u_{ij})$ and $g(a_j)$ need to be specified. We assume that u_{ij} and a_j are independently normal distributed with mean zero and variances σ_u^2 and σ_a^2, respectively. Thus, changing variables from u_{ij} to $z_{ij} = u_{ij}/\gamma_2\sigma_u$ and from a_j to $\tilde{z}_j = a_j/\gamma_2\sigma_a$, we obtain

$$h(\sigma_u z_{ij}\gamma_2\sigma_u) = \frac{1}{\sqrt{\pi}} e^{-z_{ij}^2}, \qquad g(\sigma_a \tilde{z}_j \gamma_2 \sigma_a) = \frac{1}{\sqrt{\pi}} e^{-\tilde{z}_j^2}.$$

Thus, the integrals are in a form amenable to Gauss-Hermite integration where one has first to form the weighted sum over the inner integrands, evaluated at the quadrature points, and then repeat the approximation for the outer integral. With q quadrature points at each step, there are a total of q^2 function calls required for computation of points and weights (see Press et al. 1999). In practice, one can obtain starting values from the ordered probit model without random effects, then use a small value for q (e.g., $q = 10$) to obtain a first round of estimates, and then increase q to larger values, such as 40, in order to check whether the results are stable.

The model has some similarity to the polychoric correlation model used in psychometrics (Olsson 1979), where the multivariate normal distribution of y_j^* is taken as point of departure. However, the particular structure of the covariance matrix implied by the hierarchical error components model makes it much more amenable to estimation by maximum marginal likelihood, as suggested here. The numerical integration is embedded in a procedure that maximizes the likelihood function with respect to $\theta = (\beta, \kappa_1, \ldots, \kappa_J, \sigma_u^2, \sigma_a^2)$. The likelihood function is simply

$$L(\theta|y,x) = \prod_{j=1}^{J} f(y_j|x_j, \theta), \tag{7}$$

where $f(y_j|x_j)$ is defined as in Eq. (6). The likelihood function can be maximized using numerical first and second derivatives and the BFGS algorithm as implemented for instance in the GAUSS Maxlik routine. Note that the likelihood can be simplified by taking logs only after the integration step has been performed.

The maximum likelihood estimator $\hat{\theta}$ has the usual properties. If the model is correctly specified the estimator is consistent, efficient and approximately normally distributed in finite samples with covariance matrix equal to the inverse of the expected Hessian.

5. Application

In this section I report estimation results for ordered probit models of well-being with and without random effects. The data come from the first 14 years (1984–1997) of the A-sample (the West German sample, excluding the immigrant "guest worker" sample) of the *German Socio-Economic Panel* (SOEP Group 2001). The most important selection criterion is dictated by the focus on family relations: only observations are retained where, for a given year, both spouses (or partners) plus at least one child lived in the same household, filled each out the personal questionnaire, and provided valid information on the variables involved in the analysis. The requirement for a valid child observation is quite restrictive, as children start filling out the personal questionnaire only in the year during which they turn 16.

In cases where individuals moved out of one household (and family) unit and formed a new one, only the original, pre-move, observations were retained. Otherwise, such observations would violate the hierarchical structure of the proposed model (the same individual would be observed in two families). For this to be a valid approach, we essentially require the family composition to be exogenous. Issues arising from endogenous formation and dissolution of families are certainly of great interest but beyond the scope of the present paper.

The thus obtained sample has 25,168 person-year observations and 7,485 family-year observations (for 1,309 different families). The size distribution of family-year observations is as follows: there are 5,197 three-person family years, 1,941 four-person family years, 285 five-person family years, and 62 family-years for families between 6 and 8 members. Of course, this distribution is not representative for family size in West Germany more generally, as it overrepresents families who participated longer in the survey and, more importantly, includes only children aged 16 or above who still live in the parental household. Among the 25,168 person-year observations, 14,970 are for parents and 10,198 for children. Note also that in the GSOEP, it is not possible to distinguish between biological and non-biological children, a distinction that has recently been emphasised in the context of inter-generational earnings mobility by Björklund and Chadwick (2003).

The dependent variable is obtained from the response to the question "How satisfied are you with your life on the whole at present" (SWB). The relative frequencies are displayed in Fig. 1, separately for parents and children. The two distributions look quite similar. The distributions are skewed to the right. The mode is the response "8". Children report more frequently a high subjective well-being of "9" or "10" than parents.

A first indication of the dependence between subjective well-being at the family level can be obtained through simple cross-tabulations. For example, among spouses the responses coincide exactly in 35% of the cases, as compared to 1/11 = 9% of perfect agreement to be expected under complete randomness. The chi-squared test statistics (100 d.f.) has the value of 3,874, and the independence hypothesis is clearly rejected. To test the hypothesis of independence between parent and child well-being in bivariate cross-tabulations, one can for example compare the SWB response of fathers and mothers each with those of a randomly selected child. The chi-squared statistics are somewhat lower than between spouses (966 for father and child, and 1018 for mother and child) but still strongly significant.

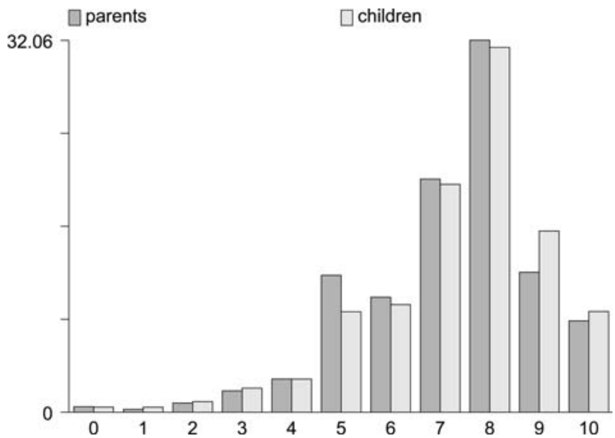

Fig. 1. Frequency distribution of SWB responses on the 0-10 scale

In order to shed more light on the nature of these dependencies (long-run or short-run, conditional or marginal), we now move from these non-parametric results to the proposed parametric ordered probit model with multiple random effects. The regression component includes a number of explanatory variables that have been found important determinants of individual subjective well-being in previous research. These include a second order polynomial in age (expected to be convex, or *u*-shaped), and an indicator for gender (male). Furthermore included are logarithmic family income (post-government transfer) and logarithmic household size. This flexible specification is preferred over the alternative of imposing an arbitraty "income equivalence scale" to account for size effects. According to previous research, bad health and unemployment are among the strongest predictors of low well-being, and both variables are included. Finally, the regression component includes a simple linear trend.

Table 1. Ordered probit models for long-term intra-family correlation in subjective well-being (N = 25168)

	Ordered probit	Ordered probit with random effects
Age/10	−0.1600**	−0.2615**
	(0.0241)	(0.0405)
(Age squared)/100	0.0194**	0.0303**
	(0.0030)	(0.0052)
Male	−0.0121	−0.0260
	(0.0130)	(0.0234)
Unemployed	−0.5807**	−0.6634**
	(0.0380)	(0.0442)
Good health	1.1205**	0.9000**
	(0.0243)	(0.0294)
Log of family income	0.1537**	0.1317**
	(0.0141)	(0.0250)
Log of number of persons	−0.1063**	−0.0556
	(0.0300)	(0.0604)
Trend	−0.0142**	−0.0352**
	(0.0017)	(0.0029)
σ_a		0.5931**
		(0.0134)
σ_u		0.6686**
		(0.0198)
κ_1	−0.7025	−2.1046
κ_2	−0.4900	−1.8598
κ_3	−0.1682	−1.4824
κ_4	0.2190	−1.0188
κ_5	0.5422	−0.6283
κ_6	1.1717	0.1584
κ_7	1.5324	0.6242
κ_8	2.1025	1.3724
κ_9	3.0040	2.5663
κ_{10}	3.6254	3.3966
Log-likelihood	−46096.0	−42736.5

Notes: Data from German Socio-Economic Panel, 1984–1997
Estimated standard errors in parentheses.
**, * Indicate statistical significance at the 5 and 10 % level, respectively.

The regression results for the full sample of 25,168 person-year observations are shown in Table 1. The first column shows the pooled ordered probit results without random effects. The second column shows the results for the ordered probit with three error components, estimated using the maxlik routine in GAUSS. The model was estimated using 10-point Gauss-Hermite integration. The two models are nested, and a comparison between them can be based on a likelihood ratio test or a Wald test. In either case, the zero-variance restrictions are strongly rejected, and the model with error components is the preferred specification.

The main effects confirm previous findings from the literature on subjective well-being. SWB is *u*-shaped in age. Unemployment has a large negative effect on SWB whereas SWB stongly increases with good health. To obtain a feeling for the magnitude of the unemployment and health effects, one can compute the required compensating variation in post-government income to keep the latent SWB index constant (i.e., to stay on the same "indifference curve"). Since a doubling in income is predicted to increase the SWB index by about 0.13 points, whereas the negative effect of unemployment amounts to a 0.66 point reduction in the index, one sees that the implicit cost of unemployment, measured in terms of reduced SWB for the individual experiencing unemployment, are enormous. In the lin-log specification for income and household size adopted here, one can also notice that SWB increases if income and household size are increased proportionally, i.e., if per-capita income remains the same. One interpretation is that income equivalence scales should give a weight lower than one for additional household members, as is usually done (see also Schwarze 2003). Finally, there is a significant negative trend in SWB.

The point estimates for the variances are $\sigma_a^2 = 0.352$ and $\sigma_u^2 = 0.447$, respectively. Since the variance of the temporary error component is normalized to 1, we can conclude that $(0.352 + 0.447)/\,(0.352 + 0.447 + 1) = 44.4$ percent of the total variance is long-term as opposed to transitory. The three-effects model allows now to decompose the long-term variance into a part attributable to the family (i.e., shared by all family members) and another part specific to the individual. The results show that the family effect is important indeed. 44.0 % in variation of long-term well-being is due to family effects, and 56.0 % is due to individual effects. In other words, the correlation in long-term SWB between members of the same family is substantial, namely 0.44. Earlier estimates, not reported here, indicated an even higher correlation in a model without the household level variables income and size. The present result shows that the correlation persists once we control, albeit admittedly in a crude fashion, for material well-being at the household level.

To shed some further light on the nature of the family effect, the analysis was repeated on different subsamples. In a first sample, only children were considered. After dropping family-year observations with a single child only, 5001 observations remained. A second sample considered the correlation in SWB among spouses only, resulting in a sample of 14,970 observations. This sample should provide an interesting benchmark, since spouses are not biologically related, but they are together by choice. The "family effect" can be interpreted in this context as an indicator of the quality of the match.

The results are given in Table 2. Again, the simple ordered probit model is rejected against the ordered probit model with multiple random effects, as the variances of the error components are highly significant. The regression

Table 2. Ordered probit models for long-term correlation among siblings and spouses in subjective well-being

	Siblings	Spouses
Age/10	−0.4167	−0.3176*
	(0.2858)	(0.1631)
(Age squared)/100	0.0386	0.0266*
	(0.0557)	(0.0152)
Male	0.0059	0.0061
	(0.0502)	(0.0241)
Unemployed	−0.7223**	−0.4403**
	(0.0887)	(0.0632)
Good health	0.9549**	0.9189**
	(0.0840)	(0.0338)
Log of family income	0.1287**	0.2026**
	(0.0624)	(0.0319)
Log of number of persons	0.0769	−0.2539**
	(0.1716)	(0.0801)
Trend	−0.0099	−0.0386**
	(0.0069)	(0.0040)
σ_a	0.5117**	0.4320**
	(0.0252)	(0.0183)
σ_u	0.5399**	0.8680**
	(0.0278)	(0.0242)
κ_1	−1.8561	−2.1585
κ_2	−1.5747	−1.9526
κ_3	−1.1615	−1.5489
κ_4	−0.7001	−1.0620
κ_5	−0.3386	−0.6401
κ_6	0.2658	0.2621
κ_7	0.6801	0.7461
κ_8	1.3997	1.5255
κ_9	2.5014	2.8036
κ_{10}	3.3214	3.6075
Log-likelihood	−8757.0	−24749.0
Number of Observations	5001	14970

Notes: See Table 1.

coefficients for the control variables are qualitatively the same as before. However, the correlation results reveal some interesting patterns. Among siblings, the correlation in long-term SWB is even higher than for all family members taken together, namely 0.47. For spouses, on the other hand, it is markedly lower, as the intra-spouse correlation in long-term SWB implied by the estimates amounts only to 0.2. A tentative explanation would be that the long-term correlation indeed is related to biological factors, and not just to living together.

6. Conclusion

The paper had two main objectives. First, it argued that interdependencies at the family level are an issue that has been neglected by the previous economic literature on subjective well-being. Secondly, it demonstrated how to model and test for such interdependencies using the framework of an ordered probit model with multiple random effects.

The application using data on families in the *German Socio-Economic Panel* suggested that the underlying long-term well-being of individuals within the same family is highly correlated, with an estimated correlation coefficient of 0.44. The correlation in long-term well-being is even higher among siblings (0.47), but markedly lower among spouses (0.20).

The paper therefore provides evidence that a purely individualistic view of explaining subjective well-being misses part of the story. There clearly are important interdependencies in reported well-being among members of the same family, some of which may have biological origins. These need to be reckoned with if one wants to understand the determinants of subjective well-being.

References

Björklund A, Jäntti M (1997) Intergenerational mobility in Sweden compared to the United States. American Economic Review 87: 1009–1018

Björklund A, Chadwick L (2003) Integenerational income mobility in permanent and separated families. Economics Letters 80: 239–246

Björklund A, Eriksson T, Jäntti M, Raaum O, Österbacka E (2002) Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. Journal of Population Economics 15: 757–772

Blanchflower DG (2004) Well-being over time in Britain and the USA. Journal of Public Economics 88: 1359–1386

Clark AE, Oswald AJ (2002) A simple statistical method for measuring how life events affect happiness. International Journal of Epidemiology 31: 1139–1144

Diener E, Lucas RE (1999) Personality and subjective well-being. In: Kahneman D, Diener E, Schwarz N (eds.) Well-being: the foundations of hedonic psychology. Russell Sage, New York

Ferrer-i-Carbonel A, Frijters P (2001). The effect of methodology on the determinants of happiness. Mimeo, Free University, Amsterdam

Frey BS, Stutzer A (2001) Happiness and economics: how the economy and institutions affect human well-being. Princeton University Press, Princeton, NJ

Groot W, Maassen van den Brink H (2003) Match specific gains to marriages: a random effects ordered response model. Quality and Quantity 37: 317–325

McKelvey R, Zavoina W (1975) A statistical model for the analysis of ordinal level dependent variables. Journal of Mathematical Sociology 4: 103–120

Olsson U (1979) Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika 44: 443–460

Oswald A (1997) Happiness and economic performance. Economic Journal 107: 1815–1831.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1999) Numerical recipes in FORTRAN 77. Cambridge University Press, New York

Shields MA, Wheatley Price S (2001) Exploring the economic and social determinants of psychological and psychosocial health. IZA Discussion Paper No. 396

SOEP Group (2001) The German Socio-Economic Panel (GSOEP) after more than 15 years - Overview. In: Holst E, Lillard DR, DiPrete TA (eds.) Proceedings of the 2000 Fourth International Conference of German Socio-Economic Panel Study Users (GSOEP2000) Vierteljahrshefte zur Wirtschaftsforschung 70: 7–14

Schwarze J (2003) Using panel data on income satisfaction to estimate the equivalence scale elasticity. Review of Income and Wealth 49: 359–372

Solon G (1989) Biases in the estimation of intergenerational earnings correlations. Review of Economics and Statistics 71: 172–174

Solon G (1999) Intergenerational mobility in the labor market. In: Ashenrelter O, Card O (eds.) Handbook of labor economics, Vol. 3 Elsevier North Holland

Solon G, Corcoran M, Gordon R, Laren D (1991) A longitudinal analysis of sibling correlations in economic status. Journal of Human Resources 26: 509–534
Tellegen A, Lykken DT, Bouchard TJ, Wilcox KJ, Segal N L, Rich S (1988) Personality similarity in twins raised apart and together. Journal of Personality and Social Psychology 54: 1031–1039
Winkelmann L, Winkelmann R (1995) Happiness and unemployment: a panel data analysis for Germany. Konjunkturpolitik 41: 293–307