DISCUSSION PAPER SERIES

IZA DP No. 1175

Income and Happiness: New Results from Generalized Threshold and Sequential Models

Stefan Boes
Rainer Winkelmann

June 2004

Income and Happiness: New Results from Generalized Threshold and Sequential Models

Stefan Boes
University of Zurich

Rainer Winkelmann
University of Zurich and IZA Bonn

Discussion Paper No. 1175
June 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 1175
June 2004

ABSTRACT

Income and Happiness: New Results from Generalized Threshold and Sequential Models*

Empirical studies on the relationship between income and happiness commonly use standard ordered response models, the most well-known representatives being the ordered logit and the ordered probit. However, these models restrict the marginal probability effects by design, and therefore limit the analysis of distributional aspects of a change in income, that is, the study of whether the income effect depend on a person's happiness. In this paper we pinpoint the shortcomings of standard models and propose two alternatives, namely generalized threshold and sequential models. With data of two waves of the German Socio-Economic Panel, 1984 and 1997, we show that the more general models yield different marginal probability effects than standard models.

JEL Classification: C25, I31

Keywords: ordered response models, marginal effects, subjective well-being

Corresponding author:

Rainer Winkelmann
Socioeconomic Institute
University of Zurich
Zürichbergstr. 14
8032 Zürich
Switzerland
Email: winkelmann@sts.unizh.ch

* We thank Gerhard Tutz for valuable comments.

1 Introduction

The objective of this paper is to point out, and offer a solution to, a methodological shortcoming of the previous literature on the effect of income on happiness in cross sectional data. Under "happiness" we understand here the response to a survey question such as — in the German Socio-Economic Panel — "Taken all together, how satisfied are you with your life today?", on a scale from 0 to 10, where 0 means "completely dissatisfied" and 10 means "completely satisfied". There is by now a large literature on the effect of income on happiness, with data for different countries, different points in time, and using different specifications. This literature is surveyed, for example, in Frey and Stutzer (2002). Since the dependent variable "happiness" is measured on an ordinal scale, it is common to employ standard ordered response models, such as ordered probit and ordered logit, although some researchers also ignore this issue and use simply OLS (for a discussion see Ferrer-i-Carbonel and Frijters (2004)).

In either case, previous studies have focused almost entirely on mean effects, i.e. on estimating the expected happiness response (or expected latent happiness indicator in the case of ordered response models) to a change in income. Completely missing so far is any evidence whether the magnitude of the income effect depends on a person's happiness: is it possible that the effect of income on happiness is different in different parts of the outcome distribution? Could it be that "money cannot buy happiness, but buy-off unhappiness" as a proverb says? And if so, how can such distributional effects be quantified?

Note that the hypothesis of a – possibly negative – relationship between the effect of income and the level of happiness is different from the hypothesis of a decreasing marginal effect of income on happiness as income increases. The two hypotheses are not the same since income is but one factor affecting happiness, with a myriad of other factors having been identified as equally or more important in the recent literature (e.g. unemployment, health, age, etc.), in addition to chance, i.e. individual unobserved factors.

The standard ordered probit or logit models with their specific structure restrict the distributional effects a-priori without the possibility "to let the data speak". These models function as benchmark, against which results from more general models can be compared. Technically, we require a model that is sufficiently flexible such that the effect of income on the probabilitiy distribution of happiness is not fully determined by functional form. We consider two such more general models. The first is a model based on a latent threshold where the thresholds themselves are linear functions of the explanatory variables. As a second alternative, we discuss a sequential model which is based on a sequence of binary choice models for the conditional probability of choosing a higher response category, borrowing concepts from the literature on discrete duration models. Both models allow for category specific parameter vectors, possibly but not necessarily different in different parts of the underlying distribution.

The paper proceeds as follows. In the next section, we offer a partial review of the previous literature on income and happiness. One important lesson is that "getting it right", i.e. correctly measuring the effect of income on happiness, will be increasingly important, as proposals are being made to use happiness surveys in the valuation of public goods and intangibles. Such valuations can become standard in cost-benefit analyses one day, complementing other methods such as hedonic modeling and contingent valuation, and therefore influence whether projects are undertaken or not, and if so, how much compensation should be paid to inconvenienced parties. This should be reason enough to closely scrutinize the methodological underpinnings of the previously reported effects.

In Section 3, the methods are presented. We start with a review of the standard ordered logit and probit models and then contrast them with the generalized threshold and sequential model. In Section 4, we report on an analysis of the relationship between income and happiness using data from two waves of the German Socio-Economic Panel, 1984 and 1997. The results are somewhat mixed. Although we find that income has a stronger effect in the left part of the happiness distribution than in the utmost right part, the restriction of equal coefficients (under which the generalized threshold model reduces to the standard ordered model) can be rejected only for women and not for men. Also, when translated into marginal probability effects, the extra income sensitivity for low happiness responses (five or below on a scale from 0 to 10) is at most one percentage points above the prediction of the standard model. Also, the probability of responding with a high value (for example 8 or above) is in all cases an increasing function of income. In this sense, there is some evidence that "money can buy happiness".

2 Previous Literature

The relationship between income and happiness is perhaps the single most intensively studied question within the growing literature on the socio-economic determinants of individual happiness (i.e. reported subjective well-being from personal surveys). The early work in the field (Easterlin, 1973, 1974, 1995; Scitovsky 1975) answered the rethoric questions "does money buy happiness?" and "will raising the incomes of all increase the happiness of all?" with a resounding "no".

This research was motivated by the apparent infatuation of the economics profession with income maximization, and it reminded us that the ultimate goal of all economic activity must be human satisfaction and that, according to the empirical evidence, maximizing GDP would not be ministerial to that goal. The standard explanation why an increased income for everyone, as measured by GDP growth say, fails to deliver higher levels of happiness – either in international comparisons or over time – are by now relatively well understood. Although a ceteris paribus increase in a person's income leads to higher well being - as is commonly found in cross section studies (e.g. Frey and Stutzer, 2000, 2002; Gerdtham and Johannesson, 2001; Gardener and Oswald, 2001; or Shields and Wheatley Price, 2004) - such an effect is strongly moderated by status and comparison effects in the aggregate (Easterlin, 1995; Clark and Oswald, 1996; Blanchflower and Oswald, 2001; Frijters et al., 2004; among others).

While the full implications of this insight for economic policy are largely ignored, there exists another aspect of the cross-sectional relationship between income and happiness that may yet prove to be of great and lasting practical importance for policy analysis. This second aspect is related to the potential of income to act as a numeraire, and hence put a price on everything else that may affect happiness as well. The basic idea is one of compensation: in the case of a "bad", how much of an increase in income is required to offset the negative effect of the bad on happiness, and keep the person at the same level of happiness as in the absence of the bad? Similarly, in case of a good, one can implicitly determine the pecuniary value of the good by asking, how much income a person would be willing to give up in order to obtain the good, so that the overall happiness is unchanged. Examples for this line of research are Winkelmann and Winkelmann (1998) who estimate the money-equivalent value of the psychological cost of unemployment, and Schwarze (2003) who uses the principle to determine an income equivalence scale, i.e. the income compensation required to keep the same individual well-being level with one additional household member present. Frey, Luechinger and Stutzer (2004) estimate the value of public safety, or the absence of terrorism. Van Praag and Baarsma (2004) estimate the external cost of air traffic noise for people living near the Amsterdam Airport. This is a very active area of research, and new interesting applications are likely to be added in the future.

Importantly, all of this research makes quite restrictive assumptions regarding the way that income is assumed to effect happiness. Usually, income is entered in logarithms. In this case the marginal effect is inversely proportional to income: to achieve the same increase in (average) happiness, larger and larger absolute changes in income are necessary. As argued above, this may not be sufficiently flexible to measure the true impact of income on happiness, and the above results may therefore be unreliable.

3 Methods

3.1 Standard Model

Let y_i denote the observed ordinal dependent variable, where $i = 1, \ldots, n$ and n is the number of observations. There are J different outcomes coded without loss of generality in a rank preserving manner such that $y_i \in \{1, \ldots, J\}$. The values of y_i are determined by an unobserved continuous variable y_i^* and a partition of the real line:

$$y_i = j \text{ if and only if } \kappa_j \leq y_i^* < \kappa_{j+1} \qquad j = 1, \ldots, J$$

where $\kappa_1, \ldots, \kappa_{J+1}$ are threshold values and it is understood that $\kappa_1 = -\infty$ and $\kappa_{J+1} = \infty$. Finally, the model is completed by assuming that

$$y_i^* = x_i'\beta + u_i$$

where x_i is a $k \times 1$ vector of covariates (excluding a constant), β is a conformable vector of parameters and u is the error term.

In standard ordered models it is assumed that the threshold values κ_j are constants to be estimated together with the regression parameters β. Estimation of the $j - 1 + k$ parameters by maximum likelihood is straightforward once a distribution function $F(u)$ has been specified. The likelihood contributions are of the form

$$P(y_i = j|x_i) = F(\kappa_{j+1} - x_i'\beta) - F(\kappa_j - x_i'\beta) \tag{1}$$

Let $y_{ij} = 1$ if $y_i = j$ and $y_{ij} = 0$ else. For a sample of n independent observations (y_i, x_i) the log-likelihood function is given by

$$\ln L(\beta, \kappa_2, \ldots, \kappa_j; y, x) = \sum_{i=1}^{n} y_{ij} \ln P(y_i = j|x_i)$$

The ordered logit and probit models are obtained by substituting for F the cumulative density function of the logistic and the standard normal distribution, respectively. For more details see McKelvey and Zavoina (1975), McCullagh (1980).

3.2 Interpretation

There are a number of ways to interpret the parameters of this model. What does it mean for an element of β to be "large" or "small"? First, one might be tempted to interpret the coefficients in terms of the latent model for y_i^*, since this part of the model is a simple linear regression. However, the β's are identified only up to scale. Moreover, y_i^*, being an artificial construct, is not of interest. Potentially more interesting is a comparison based on "compensating variation". Let x_{il} denote the l-th element of the vector of covariates and β_l the corresponding parameter. Now consider changing two covariates x_{il} and x_{im} at the same time such that $\Delta y_i^* = 0$ (and therefore all probabilities are unchanged). This requires

$$\beta_l \Delta x_{il} = -\beta_m \Delta x_{im} \qquad \text{or} \qquad \frac{\Delta x_{il}}{\Delta x_{im}} = -\frac{\beta_m}{\beta_l}$$

If, for example, x_{il} is logarithmic income and x_{im} is unemployment, then the above fraction gives the relative increase in income required to compensate for the negative effect of unemployment (assuming that $\beta_l > 0$ and $\beta_m < 0$).

To move the interpretation closer to the observed outcomes y_i, the threshold mechanism needs to be taken into account. One way of doing this is to ask how much of a change in a covariate it takes to move over one response category. For this purpose, one can form the ratio of the interval length to the parameter $(\kappa_{j+1} - \kappa_j)/\beta_l$. The smaller this ratio (in absolute terms), the smaller the maximum change in x_{il} required to move the response from category j to category $j+1$.

Such measures stop short of the most natural way of interpreting the parameters in discrete probability models such as ordered response models, namely in terms of marginal probability effects. Indeed, the main concern of this paper is to determine how the change in a covariate, such as income, changes the distribution of the outcome variable, here subjective well-being. The marginal probability effects can be obtained directly from equation (1):

$$MPE_{jl}(x_i) = \frac{\partial P(y_i = j|x_i)}{\partial x_{il}} = [f(\kappa_j - x_i'\beta) - f(\kappa_{j+1} - x_i'\beta)]\beta_l \tag{2}$$

where f is the density function of u. In general, marginal probability effects are functions of x_i. Average marginal probability effects can be obtained by taking expectations:

$$AMPE_{jl} = \mathrm{E}_x \left[\frac{\partial P(y_i = j|x_i)}{\partial x_{il}}\right] \tag{3}$$

A consistent estimator of the $AMPE_{jl}$ is obtained by replacing β in equation (2) with the maximum likelihood estimator $\hat{\beta}$ and averaging over the sample:

$$A\widehat{MP}E_{jl} = \frac{1}{n}\sum_{i=1}^{n} \widehat{MP}E_{jl}(x_i)$$

It is interesting to note that despite their intuitive appeal, marginal probability effects are rarely reported in practice. Among all contributions on the determinants of subjective well-being listed in the references just three report them, where two of them consider the marginal effect on the highest outcome only.

Once we focus on the full distribution of outcomes, and the marginal probability effects, it becomes immediately apparent that standard ordered models are quite restrictive, and perhaps unnecessarily so. A first way to pinpoint the restrictive nature of the marginal effects is the observation that their relative magnitude is not allowed to vary over the outcomes. We obtain

$$\frac{MPE_{jl}(x_i)}{MPE_{jm}(x_i)} = \frac{\beta_l}{\beta_m}$$

The relative marginal effects do not depend on j (nor do they depend on x_i); in other words, they are the same in every part of the distribution. It is not possible, for instance, that income is more important (relative to unemployment, say) in the left part of the outcome distribution than in the right part. This property holds regardless of the choice of F.

A second restrictive property is that the sign of the marginal effects for increasing j is entirely determined by the distribution function F. For example, with F being either standard normal or logistic, f is bell-shaped with a maximum at 0. Also note that in order to be well defined, we must have $\kappa_j \leq \kappa_{j+1}$. It follows from equation (2) that

$$\text{sgn}[MPE_{jl}(x_i)] = -\text{sgn}[\beta_l] \;\text{ if }\; \kappa_j < x_i'\beta \text{ and } \kappa_{j+1} \leq x_i'\beta$$

$$\text{sgn}[MPE_{jl}(x_i)] = \text{sgn}[\beta_l] \;\text{ if }\; \kappa_j \geq x_i'\beta \text{ and } \kappa_{j+1} > x_i'\beta$$

The sign is indeterminate for $\kappa_j < x_i'\beta$ and $\kappa_{j+1} > x_i'\beta$. The model has a "single crossing" property. As one moves from the probability of the smallest outcome to the probability of the largest outcome, marginal probability effects are either first negative and then positive, or they are first negative and then positive. A reversion is precluded by the assumptions of the model.

More specific results can be obtained once we consider a specific distribution function F. The best known result is the proportional log-odds assumption of the ordered logit model. From equation (1)

$$P(y_i < j|x_i) = \Lambda(\kappa_j - x_i'\beta) = \frac{\exp(\kappa_j - x_i'\beta)}{1 + \exp(\kappa_j - x_i'\beta)}$$

and therefore

$$\frac{P(y_i \geq j|x_i)}{P(y_i < j|x_i)} = \frac{1 - \Lambda(\kappa_j - x_i'\beta)}{\Lambda(\kappa_j - x_i'\beta)} = \exp(-\kappa_j + x_i'\beta)$$

Hence, the logarithmic odds of an outcome greater or equal than j relative to an outcome less than j are a linear function of x_i. The slope does not depend on the category j.

The heart of the matter is the single index assumption. In order to obtain more flexible response patterns for probabilities or odds, one will need to look for a richer parametric model where index functions are allowed to vary across response categories. In the next two sections, such models will be discussed. The presented models allow any covariate to have different effects in different parts of the distribution, relative to what the base model with constant thresholds would predict.

3.3 Generalized Threshold Models

When searching for more flexible parametric models for ordered dependent variables, the multinomial logit model stands at one extreme in terms high flexibility. However, it does not make any use of the ordering information and therefore is inefficient. A very flexible model that uses the ordering information can be obtained by making the threshold parameters linear functions of the covariates (e.g., Maddala, 1983, Terza, 1985). Let $\kappa_{ij} = \tilde{\kappa}_j + x_i'\gamma_j$, $j = 1, \ldots, J$ where x_i is a $k \times 1$ vector of covariates excluding the constant as before.

After substitution κ_{ij} for κ_j into equation (1), we obtain

$$\begin{aligned} P(y_i = j|x_i) &= F(\tilde{\kappa}_{j+1} + x_i'\gamma_{j+1} - x_i'\beta) - F(\tilde{\kappa}_j + x_i'\gamma_j - x_i'\beta) \\ &= F(\tilde{\kappa}_{j+1} - x_i'\beta_{j+1}) - F(\tilde{\kappa}_j - x_i'\beta_j) \end{aligned}$$

where $\beta_j = \beta - \gamma_j$ and it is understood that $\tilde{\kappa}_1 = -\infty$ and $\tilde{\kappa}_{J+1} = \infty$ as before. Thus, $F(\tilde{\kappa}_1 - x_i'\beta_1) = 0$ and $F(\tilde{\kappa}_{J+1} - x_i'\beta_{J+1}) = 1$.

The model now contains $J-1$ parameter vectors $\beta_2, \ldots, \beta_J$, plus $J-1$ constants $\tilde{\kappa}_2, \ldots, \tilde{\kappa}_J$ that can be estimated jointly by maximum likelihood. The generalized model nests the standard ordered model under the restriction

$$\beta_2 = \ldots = \beta_J$$

Hence, the restricted model has $(J-2) \times k$ additional degrees of freedom. Clearly, the proliferation of parameters, in particular when J is large, is a potential disadvantage. However, a test can be easily conducted, and one can economize on degrees of freedom by imposing partial restrictions in subsets of outcomes, such as $\beta_2 = \beta_3$, while allowing parameters to differ in other parts of the distribution.

The model has substantially more flexible marginal probability effects, since

$$MPE_{jl}(x_i) = f(\tilde{\kappa}_j - x_i'\beta_j)\beta_{jl} - f(\tilde{\kappa}_{j+1} - x_i'\beta_{j+1})\beta_{j+1,l} \tag{4}$$

All the statements in the previous subsection on constant relative effects and single crossing no longer need to hold. Rather, these effects can be determined empirically.

In the logit case, we obtain

$$P(y_i < j|x_i) = \Lambda(\tilde{\kappa}_j - x_i'\beta_j)$$

and therefore

$$\begin{aligned} \frac{P(y_i \geq j|x_i)}{P(y_i < j|x_i)} &= \frac{1 - \Lambda(\tilde{\kappa}_j - x_i'\beta_j)}{\Lambda(\tilde{\kappa}_j - x_i'\beta_j)} \\ &= \exp(x_i'\beta_j - \tilde{\kappa}_j) \end{aligned}$$

The effects of covariates on the log-odds are now category specific.

The greater flexibility in modeling ordered responses with generalized thresholds does not come without costs. First, the constraint of ascending constants in the standard model to obtain a well-defined likelihood now extends to

$$\kappa_j - x_i'\beta_j \leq \kappa_{j+1} - x_i'\beta_{j+1} \qquad \forall i, j \tag{5}$$

In a model with generalized thresholds, it is necessary that the multiple indices satisfy the order restrictions for all observations i. As a practical consequence, the large number of parameters in conjunction with the order restriction (5) may increase computation time considerably, as attempts of unproductive likelihood steps are routinely made.

3.4 Sequential Model

For an alternative approach for modeling an ordinal response variable, we now consider the class of sequential models. This kind of model has been discussed in the statistics literature (see, for example, Tutz, 1991). However, previous applications in econometrics are, to the best of our knowledge, non-existent. As before, we assume that the ordered response variable is coded as $j = 1, \ldots, J$, where $j = 1$ designates the smallest outcome and $j = J$ the largest. The basic idea is to cast the model in terms of conditional transition probabilities $P(y_i = j | y_i \geq j)$, $j = 1, , \ldots, J$. These conditionals fully characterize the probability function of y. For example,

$$P(y_i = 1) = P(y_i = 1 | y_i \geq 1) P(y_i \geq 1) = P(y_i = 1 | y_i \geq 1)$$

$$P(y_i = 2) = P(y_i = 2 | y_i \geq 2) P(y_i \geq 2) = P(y_i = 2 | y_i \geq 2)(1 - P(y_i = 1 | y_i \geq 1))$$

$$\begin{aligned} P(y_i = 3) &= P(y_i = 3 | y_i \geq 3) P(y_i \geq 3) \\ &= P(y_i = 3 | y_i \geq 3)[1 - P(y_i = 1) - P(y_i = 2)] \\ &= P(y_i = 3 | y_i \geq 3)(1 - P(y_i = 1 | y_i \geq 1))(1 - P(y_i = 2 | y_i \geq 2)) \end{aligned}$$

and in general

$$P(y_i = j) = P(y_i = j | y_i \geq j) \prod_{r=0}^{j-1} (1 - P(y_i = r | y_i \geq r)) \qquad j = 1, \ldots, J$$

where it is understood that $P(y_i = 0|y_i \geq 0) = 0$ and $P(y_i = J|y_i \geq J) = 1$.

This approach is in close analogy to discrete time hazard rate models in duration analysis. Let t_j, $j = 1, \ldots, J$ denote the possible exit points, ordered by time. Then $P(T = t_j|T \geq t_j)$ gives the probability of exit at time t_j, conditional on survival until t_j.

The sequential model naturally accounts for the ordering of the responses without imposing any arbitrary cardinality assumption. We don't claim that the model is a literal representation of the cognitive processes that are at work when a respondent answers this type of question. It even appears rather unlikely that individuals actually think that way, starting with the lowest category, here one, and sequentially choose between one or at least two, two or at least three, and so on. We rather see the sequential model as a flexible tool for obtaining a model for ordered responses with unrestricted marginal probability effects.

In order to model the effects of explanatory variables, we can parameterize the model as follows:

$$P(y_i = j|y_i \geq j, x_i) = F(\alpha_j + x_i'\beta_j)$$

where α_j is a category specific constant, β_j is a vector of category specific slopes, and F is any function mapping real numbers onto the unit interval. The corresponding probability function is

$$P(y_i = j|x_i) = F(\alpha_j + x_i'\beta_j) \prod_{r=0}^{j-1} (1 - F(\alpha_r + x_i'\beta_r)) \qquad j = 1, \ldots, J \tag{6}$$

where it is understood that $\alpha_0 = -\infty$ and $\alpha_J = \infty$ such that $F(\alpha_0 + x_i'\beta_0) = 0$ and $F(\alpha_J + x_i'\beta_J) = 1$. A model with category specific constants only is obtained as a special case.

An important advantage of this model over the generalized threshold model is that no restrictions on the parameter space are required to ensure the existence of a proper probability function. This simplifies estimation considerably. On the other hand, due to the increasing number of terms in (6), the calculation of marginal probability effects becomes somewhat intractable, in particular for large J, since repeated application of product and chain rules is required.

But as before, more specific results can be obtained by assuming that F is the cumulative distribution function of the logistic distribution. In this case, the marginal probability effects can be derived as

$$MPE_{jl}(x_i) = \frac{\partial P(y_i = j|x)}{\partial x_{il}} = P(y_i = j|x_i) \left[[1 - \Lambda(\alpha_j + x_i'\beta_j)]\beta_{jl} - \sum_{r=0}^{j-1} \Lambda(\alpha_r + x_i'\beta_r)\beta_{rl} \right]$$

Like in the generalized ordered logit model the effects of a change in one explanatory variable are now local, i.e. they vary by category j. To estimate the parameters of the model one needs to perform $J - 1$ consecutive logit regressions. The dependent variable y_{ij} is equal to one if $y_i = j$ and equal to zero if $y_i > j$. In each step, only observations "at risk" are included, i.e. those for which it is the case that $y_i \geq j$.

4 Data and Results

In this section we report on the estimation results of the relationship between income and happiness. We use data from two waves — 1984 and 1997 — of the A-sample of the German Socio-Economic Panel (GSOEP), i.e. we consider the sample of West Germans (SOEP Group 2001). The dependent variable *happiness* is the response to the survey question "Taken all together, how satisfied are you with your life today?", on a scale from 0 to 10, where 0 means “completely dissatisfied” and 10 means “completely satisfied”. To avoid cells with low frequency we combined the categories 0, 1 and 2, as well as the categories 3 and 4, leaving $J = 8$ categories for estimation. Explanatory variables include a quadratic in age (age between 25 and 64), dummy variables for health status, unemployment, logarithmic household income and logarithmic household size. By entering income and size separately, we allow for flexible scale effects in utility from consumption at the household level. The alternative, using per-capita income (or its logarithm) would impose constant returns to scale (see also Schwarze, 2003). All analyses are performed separately by sex.

— Insert Table 1 about here —

Table 1 summarizes the sample means of selected variables. The descriptive statistics show that the share of responses in the highest “happiness” categories was markedly lower in 1997 compared to the year 1984. Overall, people are mostly satisfied with their life: more than two-thirds of women and men state a happiness level of seven or higher. From a descriptive point of view, happiness and income are positively related over most of the range, with the exception of the highest happiness category. Respondents in the highest category have on average less income than those reporting a high happiness, but not the highest. A common result known from the previous literature and also supported by the present data is that among unhappy people the unemployment rate is relatively high, and that reported health status and happiness are positively correlated.

In the following, we report results from three different models: the standard ordered response model, the generalized threshold model, and the sequential model, all of them with a logistic specification of the error terms. In the discussion, we will emphasize two important aspects: model selection and testing; and interpretation of the models in terms of marginal probability effects.

4.1 Model Selection and Testing

Before we start with different procedures of model selection we give estimation results for one subsample, women 1984, and then illustrate selection criteria by means of this example. Table 2 reports estimation results for the standard ordered logit model, Table 3 and Table 4 for the generalized threshold and the sequential model, respectively.

— Insert Table 2, 3 and 4 about here —

In the standard ordered logit model the parameter vector β is estimated together with $J-1=7$ thresholds. All significant parameters have the expected sign. For example household income and health status have a positive effect on happiness, whereas unemployment has a negative effect. The relationship between age and happiness is insignificant. We caution that such an interpretation falls short of the most natural way of interpretation in terms of marginal probability effects, and can be misleading in a sense that the standard model might be too restrictive with equal parameter vectors in each part of the outcome distribution.

In the generalized threshold model the estimated parameters differ in each category. By visual inspection, the coefficients on logarithmic household income are decreasing from the lowest to the highest category. To test the hypothesis of no change we can conduct a likelihood ratio (LR) test by computing minus two times the difference of the restricted and the unrestricted log-likelihood values. The LR statistic is asymptotically χ^2-distributed with degrees of freedom equal to the number of restrictions, $J-2$ in the case of a single regressor such as income. For women, 1984 and 1997, we can clearly reject the null hypothesis of equal income coefficients, whereas for men in both years we cannot reject the hypothesis at conventional levels of significance.

The null hypothesis of no category specific parameters can be tested in the same manner, i.e. by computing minus two times the difference of the log-likelihoods of the standard and the generalized ordered model. For women in 1984 the LR statistic is 104.64, the degrees of freedom are 36, and thus the null hypothesis of equal coefficients can be highly rejected. The same result applies for women 1997 and men in both waves. Therefore in either case the generalized model is preferred to the standard model.

In the sequential logit model the hypothesis of equal coefficients is not a relevant one. Indeed, equal coefficients, and therefore equal *conditional* transition probabilities while moving from $j = 1$ to $j = J$, would mean that the *marginal* probabilities $P(y_i = j)$ decrease monotonically following a geometric decay function. This clearly is not an interesting benchmark. The sequential model does not nest the standard ordered logit model. To make the results comparable across models, we need to consult the marginal probability effects rather than the parameter values *per se*. To compare the goodness-of-fit of the two models, we can use the *Akaike Information Criterion*. Since the number of parameters is the same in the sequential and generalized threshold model, this amounts to comparing the log-likelihoods and choosing the model with the higher value. In the sequential model we can simply sum the log-likelihoods of each binary logit regression to obtain the overall log-likelihood. In three cases the generalized ordered logit model is preferred to the sequential model (e.g. women

1984: $-5474.75 > -5479.82$), only for men 1997 the log-likelihood value of the sequential logit is slightly higher.

4.2 Marginal Probability Effects of Income

The use of probability models for ordered responses lends itself to the interpretation of parameters in terms of marginal probability effects. The question of interest is: How does the probability of observing a certain outcome $j = 1, \ldots, J$ change if one of the explanatory variables changes? In this paper, we are mainly interested in the effect of a change in income on the probability of being more or less happy. For female observations in 1984 we reported the estimated average marginal probability effects ($\widehat{AMPE}$) in Tables 2,3 and 4. The results for women 1997 and men in both years are summarized in Table 5.

— Insert Table 5 about here —

How can we interpret the marginal probability effects precisely? Consider, for example, the results for men, 1997, and take the ceteris paribus effect of increasing logarithmic household income by a small amount on the probability of choosing the sixth category ($j = 6$). For the standard ordered logit model Table 5 shows a value of 0.0465 (the rows for the standard model are denoted by "OLogit"). That is, the probability of $j = 6$ increases by 0.0465 percentage points if we increase $\ln(hhincome)$ by 0.01 (this is approximately a 1% change in level household income). Now take a doubling of household income, i.e. a change in logarithmic household income of 0.693, keeping household size and everything else constant. The associated change in the probability of $j = 6$ is about $0.0465 \times 0.693 \times 100$, a 3.2 percentage point increase. In the generalized ordered logit as well as in the sequential logit (rows “GOLogit” and “SLogit”, respectively) the estimated marginal probability effects of income are considerably larger. They exceed the standard ordered logit effect by more than 50 percent. A doubling in income is predicted to increase the probability of a 6 response by approximately 5.1 percentage points.

If we compare the average marginal probability effects among the three different models and over all possible outcomes, we get somewhat mixed results. On one hand, all models suggest that more income reduces the probability of being unhappy, and increases that of being happy. On the other hand, there is no systematic under- or over-estimation in the “restrictive” ordered logit model compared to the generalized models. However, our data point out in a nice way the shortcomings of the standard model. Consider for example the estimated marginal probability effects for men in the year 1997. In the generalized ordered logit and the sequential logit these effects are at first negative, then positive, and then again negative in the highest category. By design, such a pattern is impossible in the standard ordered logit model. A similar result appears for women 1984, where the signs switch three times if marginal probability effets are estimated by the generalized threshold model.

Apart from point estimation of marginal probability effects, sampling variability is also important. In principle, standard errors can be obtained by the delta method and asymptotic normality follows from maximum likelihood estimation. However, due to the increasing number of terms in the marginal probability expressions of the sequential logit model, multiple chain and product rules are required and bootstrapped standard errors offer an “easy-to-implement” alternative.

5 Conclusions

This paper had two main objectives. First, we stressed that the previous literature has mostly neglected distributional aspects of the effect of income on happiness, and therefore ruled out different effects on low and high happiness responses, respectively. As mentioned above, this can become important in future research when happiness surveys are used for valuation of public goods or intangibles. Secondly, we pinpointed the shortcomings of the standard ordered response models in analyzing these issues, and proposed several alternatives (the generalized threshold and the sequential model) that are flexible enough to answer questions like "Is it true that money cannot buy happiness, but buy-off unhappiness?"

We illustrated these models with data from the 1984 and 1997 waves of the GSOEP. The main focus was put on model selection and average marginal probability effects. The latter was discussed in detail since this is the most natural way of interpretating regression coefficients in conditional probability models. The restrictiveness of standard methods compared to the more general alternatives was highlighted, e.g. with signs of marginal effects that are impossible in the standard model. In this sense we let the data speak, and determined empirically that there is some evidence for "Money can buy happiness".

Our results are related to three other recent papers who also relax the strict assumptions of standard ordered models. In a first paper, Clark et al. (2004) take explicitly into account that unobserved individual heterogeneity may cause different transformations of observable characteristics into the verbal expression of satisfaction. They apply a latent class approach to the ordered probit model and show with panel data from twelve European countries that marginal income effects differ among classes. Secondly, Headey and Wooden (2004) allow for asymmetries in the well-being response to income by using two measures for the lefthand-side variable from the same survey, one being well-being (or happiness) and the other being ill-being (or psychological distress). Their results support the hypothesis that income has the same impact on both measures. Thirdly, Kerkhofs and Lindeboom (1995) consider a model for state-dependent reporting errors in subjective health measures. They generalize the threshold values as function of explanatory variables, but focus on predicted probabilities and not marginal probability effects. A main result is that predictions of reported health status are sensitive to the labor market status.

However, with very few exceptions (Clark et al., 2004; Shields and Wheatley Price, 2004; Frey and Stutzer, 2002: ch.4) marginal probability effects have been neglected in the literature on happiness and income. This paper argued that this does not have to be so necessarily, and that one needs to go beyond the standard ordered response models if one wants to estimate these effects properly. Therefore, we conclude with a strong recommendation that future work in this area seriously considers the use of more general models.

References

Blanchflower, D., and A. Oswald (2001): "Well-Being Over Time in Britain and the USA," *revised NBER working paper no. 7487.*

Clark, A., F. Etil, F. Postel-Vinay, C. Senik, and K. V. der Straeten (2004): "Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries," *unpublished working paper*, DELTA.

Clark, A., and A. Oswald (1996): "Satisfaction and Comparison Income," *Journal of Public Economics*, 61, pp. 359–381.

Easterlin, R. (1973): "Does Money Buy Happiness?," *Public Interest*, 30, 3–10.

Easterlin, R. (1974): "Does Economic Growth Improve the Human Lot? Some Empirical Evidence," in *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*, ed. by P. David, and M. Reder, pp. 89–125. Academic Press, New York.

Easterlin, R. (1995): "Will Raising the Incomes of All Increase the Happiness of All?," *Journal of Economic Behaviour and Organization*, 27, pp. 35–48.

Easterlin, R. (2001): "Income and Happiness: Towards a Unified Theory," *The Economic Journal*, 111, pp. 465–484.

Ferrer-i-Carbonel, A., and P. Frijters (2004): "How important is methodology for the estimates of the determinants of happiness?," *Economic Journal* forthcoming.

Frey, B.S., S. Luechinger, and A. Stutzer (2004) Valuing Public Goods: The Life Satisfaction Approach, CESifo Working Paper No. 1158.

Frey, B.S., and A. Stutzer (2000): "Happiness, Economy and Institutions," *The Economic Journal*, 110, pp. 918–938.

Frey, B.S., and A. Stutzer (2002): *Happiness & Economics.* Princeton University Press, Princeton and Oxford.

Frijters, P., J. Haisken-DeNew, and M. Shields (2004): "Money does matter! Evidence from increasing real incomes and life satisfaction in East Germany following reunification," *American Economic Review*, forthcoming.

Gardener, J., and A. Oswald (2001): "Does Money Buy Happiness? A Longitudinal Study Using Data on Windfalls," *unpublished working paper*, Warwick University.

Gerdtham, U.-G., and M. Johannesson (2001): "The Relationship Between Happiness, Health and Socio-Economic Factors: Results based on Swedish Micro Data," *Journal of Socio-Economics*, 30, pp. 553–557.

Headey, B., and M. Wooden (2004): "The Effects of Wealth and Income on Subjective Well-Being and Ill-Being," *IZA Discussion Paper No. 1032.*

Kerkhofs, M., and M. Lindeboom (1995): "Subjective Health Measures and State Dependent Reporting Errors," *Health Economics*, 4, pp. 221–235.

Maddala, G. (1983): *Limited-dependent and qualitative variables in econometrics.* Cambridge University Press, Cambridge.

McCullagh, P. (1980): "Regression Models for Ordinal Data," *Journal of the Royal Statistical Society. Series B (Methodological)*, 42, pp. 109–142.

McKelvey, R., and W. Zavoina (1975): "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," *Journal of Mathematical Sociology*, 4, pp. 103–120.

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

Scitovsky, T. (1975): "Income and Happiness," *Acta Oeconomica*, 15, pp. 45–53.

Shields, M., and S. Wheatley Price (2004): "Exploring the Economic and Social Determinants of Psychological Well-Being and Perceived Social Support in England," *forthcoming in Journal of The Royal Statistical Society.*

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

Terza, J. (1985): "Ordinal Probit: A Generalization," *Communications in Statistics – Theory and Methods*, 14, pp. 1–11.

Tutz, G. (1991): "Sequential Models in Ordinal Regression," *Computational Statistics and Data Analysis*, 11, pp. 275–295.

van Praag, B.M.S. and B. E. Baarsma (2004): "Using Happiness Surveys to Value Intangibles: The Case of Airport Noise," IZA Discussion Paper No. 1096.

Winkelmann, L. and R. Winkelmann (1998): "Why Are the Unemployed So Unhappy? Evidence from Panel Data," *Economica*, 65, pp. 1–15.

Table 1: Sample means for selected variables by sex and year

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
women, 1984, $n=2981$								
rel. frequency	0.023	0.037	0.126	0.083	0.148	0.248	0.131	0.203
hhincome	28830	35691	34752	37477	40836	40613	44776	41872
health	0.429	0.649	0.806	0.863	0.912	0.931	0.959	0.950
unemp	0.086	0.108	0.029	0.024	0.016	0.023	0.018	0.013
age	45.96	44.23	44.80	43.552	43.562	43.312	40.980	45.337
women, 1997, $n=2208$								
rel. frequency	0.018	0.065	0.110	0.121	0.233	0.296	0.118	0.039
hhincome	42290	51322	52823	56139	63727	64748	67597	60978
health	0.462	0.657	0.844	0.910	0.942	0.969	0.977	0.977
unemp	0.179	0.098	0.078	0.060	0.033	0.035	0.019	0.011
age	44.59	42.43	43.86	43.17	41.62	42.00	40.09	45.31
men, 1984, $n=2844$								
rel. frequency	0.029	0.039	0.103	0.083	0.170	0.277	0.127	0.172
hhincome	33472	35782	36548	41045	42186	44395	45985	41902
health	0.470	0.645	0.788	0.886	0.936	0.953	0.958	0.949
unemp	0.253	0.091	0.075	0.051	0.021	0.011	0.006	0.010
age	41.96	42.98	43.71	42.13	41.10	43.20	42.33	44.60
men, 1997, $n=2169$								
rel. frequency	0.019	0.076	0.100	0.117	0.253	0.303	0.097	0.036
hhincome	52383	56184	56437	60865	65826	70312	70009	63046
health	0.690	0.652	0.796	0.921	0.954	0.977	0.976	0.961
unemp	0.261	0.189	0.106	0.102	0.051	0.035	0.024	0.052
age	40.83	43.95	43.88	42.33	41.20	42.27	40.87	42.43

Table 2: Estimating *happiness* by standard ordered logit, sample: women, 1984

		coeff.	std. err.
$\ln(hhincome)$		0.3725***	(0.0570)
$\ln(hhsize)$		-0.1479*	(0.0836)
$unemp$		-0.7394***	(0.2242)
$health$		1.5472***	(0.1149)
age		0.0043	(0.0255)
$agesq/100$		0.0005	(0.0290)
Mean$(-x'\hat{\beta})$		-5.295093	
κ_j	$j=1$	$-\infty$	
	$j=2$	1.3864	(0.7397)
	$j=3$	2.4210	(0.7351)
	$j=4$	3.7767	(0.7356)
	$j=5$	4.2871	(0.7366)
	$j=6$	4.9917	(0.7379)
	$j=7$	6.0579	(0.7401)
	$j=8$	6.7543	(0.7415)
$\widehat{AMPE}_{j,\ln(income)}$	$j=1$	-0.0085	(0.0018)
	$j=2$	-0.0124	(0.0023)
	$j=3$	-0.0327	(0.0056)
	$j=4$	-0.0152	(0.0026)
	$j=5$	-0.0159	(0.0027)
	$j=6$	0.0056	(0.0018)
	$j=7$	0.0207	(0.0034)
	$j=8$	0.0585	(0.0099)
Number of obs.		2981	
$\ln L(\hat{\beta})$		-5527.067	

significance levels : * 10% ** 5% *** 1%

Table 3: Estimating *happiness* by generalized ordered logit, sample: women, 1984

	Estimated parameters $\hat{\beta}_j$ (standard errors in parentheses)							
	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$\ln(hhincome)$	—	0.6796***	0.6140***	0.6143***	0.4000***	0.3232***	0.2948***	0.2196***
		(0.1666)	(0.1162)	(0.0857)	(0.0703)	(0.0654)	(0.0689)	(0.0719)
$\ln(hhsize)$	—	-0.2385	-0.3987**	-0.3449***	-0.1846*	-0.0759	-0.1247	-0.1081
		(0.2848)	(0.1880)	(0.1238)	(0.1100)	(0.0986)	(0.1020)	(0.1184)
$unemp$	—	-1.1903***	-1.5020***	-0.8372***	-0.7299***	-0.4254*	-0.4984*	-0.5377
		(0.4586)	(0.2927)	(0.2551)	(0.2465)	(0.2462)	(0.2975)	(0.3781)
$health$	—	2.4983***	2.0840***	1.5642***	1.4868***	1.2943***	1.1882***	1.0625***
		(0.2771)	(0.1715)	(0.1301)	(0.1263)	(0.1312)	(0.1671)	(0.1989)
age	—	-0.0477	-0.0483	0.0366	0.0319	-0.0147	0.0015	0.0392
		(0.0954)	(0.0602)	(0.0371)	(0.0322)	(0.0297)	(0.0315)	(0.0376)
$agesq/100$	—	0.0650	0.0602	-0.0429	-0.0351	0.0189	0.0018	-0.0226
		(0.1082)	(0.0683)	(0.0419)	(0.0366)	(0.0337)	(0.0357)	(0.0421)
$\tilde{\kappa}_j$	$-\infty$	-3.9650*	-3.8478**	-6.5689***	-4.9264***	-3.8511***	-4.8255***	-5.7848***
		(2.3184)	(1.6294)	(1.0903)	(0.9151)	(0.8440)	(0.8982)	(1.0277)
Mean($\hat{\tilde{\kappa}}_j - x'\hat{\beta}_j$)	$-\infty$	-4.3207	-3.0941	-1.5861	-1.0442	-0.3304	0.7264	1.4180
Mean($\hat{P}(y = j)$)	0.0238	0.0374	0.1265	0.0828	0.1481	0.2477	0.1312	0.2026
$\widehat{AMPE}_{j,\ln(income)}$	-0.0146	-0.0171	-0.0529	0.0116	-0.0018	0.0110	0.0290	0.0348
(bootstr. std. err.)	(0.0048)	(0.0060)	(0.0111)	(0.0108)	(0.0104)	(0.0130)	(0.0111)	(0.0124)
Number of obs.	2981							
$\ln L(\hat{\beta})$	-5474.748							

significance levels : * 10% ** 5% *** 1%

Table 4: Estimating *happiness* by sequential logit, sample: women, 1984

	Estimated parameters $\hat{\beta}_j$ (standard errors in parentheses)							
	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
ln(*hhincome*)	-0.5667***	-0.4028***	-0.4139***	-0.1256	-0.1318	-0.1523*	0.1684	—
	(0.1451)	(0.1270)	(0.0804)	(0.1078)	(0.0846)	(0.0810)	(0.1102)	
ln(*hhsize*)	0.1119	0.3441	0.1799	-0.0146	-0.0499	0.1337	-0.1314	—
	(0.2969)	(0.2544)	(0.1425)	(0.1734)	(0.1398)	(0.1276)	(0.1762)	
unemp	1.0573**	1.6170***	0.3058	0.1607	-0.2059	0.3654	0.1816	—
	(0.4791)	(0.3555)	(0.3510)	(0.4508)	(0.4243)	(0.3607)	(0.5338)	
health	-2.5276***	-1.7540***	-1.1078***	-0.8761***	-0.5048**	-0.4236**	-0.0171	—
	(0.2684)	(0.2247)	(0.1578)	(0.2104)	(0.2010)	(0.2122)	(0.3268)	
age	0.0353	0.0503	-0.0649	-0.0211	0.0719*	-0.0162	-0.0430	—
	(0.0975)	(0.0787)	(0.0432)	(0.0527)	(0.0428)	(0.0389)	(0.0532)	
agesq/100	-0.0507	-0.0634	0.0775	0.0191	-0.0840*	0.0155	0.0084	—
	(0.1090)	(0.0891)	(0.0487)	(0.0598)	(0.0486)	(0.0441)	(0.0606)	
α_j	3.2047	1.0118	4.4684***	0.4745	-0.8987	1.9440*	-0.3743	—
	(2.4179)	(1.9816)	(1.1575)	(1.4534)	(1.1663)	(1.0890)	(1.5217)	
Mean($\hat{\alpha}_j + x'\hat{\beta}_j$)	-4.3111	-3.4624	-1.9009	-2.1582	-1.3514	-0.2693	-0.4701	—
Mean($\hat{P}(y=j)$)	0.0235	0.0370	0.1265	0.0832	0.1480	0.2477	0.1310	0.2030
$\widehat{AMPE}_{j,\ln(income)}$	-0.0121	-0.0125	-0.0387	-0.0017	-0.0015	0.0077	0.0357	0.0230
(bootstr. std. err.)	(0.0029)	(0.0043)	(0.0082)	(0.0054)	(0.0108)	(0.0109)	(0.0118)	(0.0112)
Number of obs.	2981	2911	2800	2423	2175	1734	996	—
$\ln L(\hat{\beta})$	-276.493	-429.916	-1065.487	-791.097	-1091.061	-1177.630	-648.137	—
$\sum_j \ln L(\hat{\beta})$				-5479.820				

significance levels : * 10% ** 5% *** 1%

Table 5: $\widehat{AMPE}_{j,\ln(hhincome)}$ (bootstrapped standard errors in parentheses)

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
				women, 1997				
OLogit	-0.0066	-0.0208	-0.0271	-0.0209	-0.0139	0.0393	0.0354	0.0145
	(0.0022)	(0.0062)	(0.0071)	(0.0053)	(0.0038)	(0.0104)	(0.0095)	(0.0042)
GOLogit	-0.0149	-0.0252	-0.0290	-0.0363	0.0397	0.0304	0.0335	0.0018
	(0.0058)	(0.0104)	(0.0145)	(0.0116)	(0.0149)	(0.0193)	(0.0095)	(0.0063)
SLogit	-0.0079	-0.0211	-0.0239	-0.0255	0.0276	0.0224	0.0277	0.0008
	(0.0039)	(0.0081)	(0.0105)	(0.0099)	(0.0134)	(0.0206)	(0.0091)	(0.0060)
				men, 1984				
OLogit	-0.0058	-0.0068	-0.0148	-0.0092	-0.0112	0.0053	0.0130	0.0295
	(0.0014)	(0.0016)	(0.0034)	(0.0020)	(0.0025)	(0.0015)	(0.0029)	(0.0065)
GOLogit	-0.0058	-0.0069	-0.0246	-0.0059	-0.0035	0.0117	0.0215	0.0135
	(0.0060)	(0.0078)	(0.0082)	(0.0080)	(0.0094)	(0.0161)	(0.0094)	(0.0094)
SLogit	-0.0080	-0.0064	-0.0219	-0.0061	-0.0031	0.0142	0.0226	0.0087
	(0.0036)	(0.0047)	(0.0063)	(0.0078)	(0.0107)	(0.0174)	(0.0126)	(0.0107)
				men, 1997				
OLogit	-0.0072	-0.0235	-0.0247	-0.0212	-0.0145	0.0465	0.0311	0.0136
	(0.0019)	(0.0049)	(0.0052)	(0.0045)	(0.0028)	(0.0094)	(0.0063)	(0.0028)
GOLogit	-0.0117	-0.0104	-0.0328	-0.0351	-0.0106	0.0731	0.0334	-0.0059
	(0.0068)	(0.0169)	(0.0157)	(0.0169)	(0.0182)	(0.0203)	(0.0123)	(0.0120)
SLogit	-0.0095	-0.0143	-0.0289	-0.0314	-0.0100	0.0732	0.0273	-0.0064
	(0.0048)	(0.0109)	(0.0083)	(0.0120)	(0.0171)	(0.0224)	(0.0118)	(0.0096)