HETEROGENEITY IN REPORTED WELL-BEING: EVIDENCE FROM TWELVE EUROPEAN COUNTRIES

Andrew Clark* (CNRS and DELTA)

Fabrice Etilé (INRA-CORELA)

Fabien Postel-Vinay (INRA-LEA, CREST and CEPR)

Claudia Senik (DELTA and University Paris IV)

Karine Van der Straeten (CNRS and CECO, Ecole Polytechnique)

Approximate Word Count : 6821

January 2004

Abstract

This paper models the relationship between income and self-reported well-being using random-effect techniques applied to panel data from twelve European countries. We cannot distinguish empirically between heterogeneities in the utility function (translating income into utility) and the expression function (turning utility into self-reported well-being), but we strongly reject the hypothesis that individuals carry out these joint transformations in the same way. The "marginal well-being effect of income" is very different in the four classes we identify; we thus expect preferences for redistribution and behaviour to be different across these classes. Our results suggests that aggregating data across diverse populations, and countries, may be a dangerous practice.

JEL Codes: C14, C23, I30.

Keywords: Income, Utility, Well-being, Heterogeneity, Latent class.

* Corresponding author: DELTA, 48 Boulevard Jourdan, 75014 Paris, France. Tel: 33-1-43-13-63-29. E-mail: Andrew.Clark@ens.fr.

Heterogeneity in Reported Well-being: Evidence from Twelve European Countries

Andrew Clark, Fabrice Etilé, Fabien Postel-Vinay,

Claudia Senik, and Karine van der Straeten

1 Introduction

It is a moot point whether the most important developments drawing the social sciences together over the past thirty years have been the increasing availability of large-scale panel survey data, and the development of statistical techniques which control for unobserved fixed effects. These two together have allowed social scientists to respond systematically to the criticism that individuals do not answers questions in the same way. This development has been particularly useful in the realm of the analysis of subjective measures.

One class of variables that has attracted interest amongst economists are those providing proxy measures of individual utility. Some of these are global indices, such as happiness, life satisfaction, or psychological stress[1]; others are domain specific, such as job or income satisfaction.

These measures have a number of particular characteristics. First, they are ordinal. A life satisfaction score of 6, on a scale of 1 to 7, does not correspond to twice as satisfied as a score of 3. In this ordinal world, 6 only means more than 5 and less than 7. Second, proxy utility measures are bounded. In our example above, someone with a satisfaction score of 7 last year has no way of indicating that she is even happier this year. As such, ordered probit or ordered logit estimation is required in cross-sections, and panel estimation of well-being is something of an econometric minefield.

Even so, a careful researcher can make progress. Perhaps the greatest drawback to the analysis of the broad class of subjective measures, including well-being or happiness scores, has been hostile colleagues or seminar audiences. Various grades of incredulity have been registered, but the underlying consensus seemed to be that “these numbers don’t mean anything”, it typically being explained that individuals answered questions in very different ways.

Advances in econometric theory, and more pragmatically in the statistical packages that the majority of economists use for their applied work, have largely overcome one part of this objection, at least in a technical sense. It is now simple to control for an individual fixed effect in an ordinal regression. This takes care of the criticism that some people always look at life pessimistically or optimistically, even though there is "really" no difference in their level of happiness. Recent examples of fixed-effect estimation of well-being are Clark and Oswald (2002) and Ferrer-i-Carbonell and Frijters (2004).

Of course, controlling for fixed effects does not by itself make the information contained in the responses meaningful. Another strand of the literature has revealed the predictive power of proxy utility measures, for example linking life satisfaction to future marriage (Lucas *et al.*, 2003, and Stutzer and Frey, 2003) or job satisfaction to future quits (Clark, 2001, and Freeman, 1978).

Here we are interested in the relationship between utility, or more pragmatically proxy measures thereof, and income. This is one of the subjects which appeals to a number of different disciplines in social science. A number of recent papers have estimated this relationship in a fixed-effect framework. This approach covers heterogeneity in the intercept of the regression line between individuals. Perhaps of just as much interest is what could be thought of as second type of heterogeneity: differences in the <u>slopes</u> of the regression lines between individuals. This is the subject of this article.

We use an econometric technique that allows us to model slope heterogeneity in the relationship between a measure of well-being, income satisfaction, and income across twelve European countries. This therefore represents an attempt at modelling heterogeneity in marginal utility. The statistical technique endogenously divides the observations (in a probabilistic sense) into separate classes or groups; we then use predicted probabilities to characterise the relationship between income and satisfaction in each group. The data strongly reject the hypothesis that this relationship is identical across groups. We also describe the probability distribution of belonging to the different classes as a function of various demographic variables.

The rest of this paper is organised as follows. Section 2 describes the problem and the data that will be used. Section 3 presents the methods implemented in order to reveal heterogeneity, and section 4 the results. Section 5 concludes.

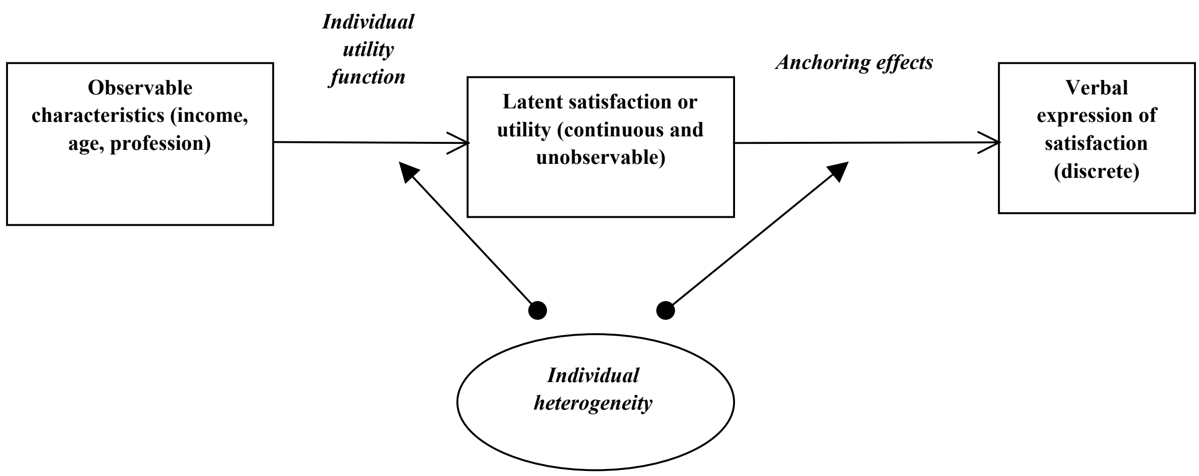
2 Income and Reported Well-Being

The relationship between income and utility or well-being is one of the great transversal questions in social science. The shape of the utility function is one of the keystones of microeconomics. The centrality of this relationship, however, has not been reflected in the quantity of empirical tests, although this has become increasingly less true in recent years. A recent survey is Senik (2004).

Generally, this empirical literature postulates that the relationship between income and subjective well-being is identical across individuals, up to a constant. Our concern in this paper is to contribute to this research, but with a twist. We not only ask the question "Does money buy happiness", but also "for whom does it buy the most happiness"? We thus propose to take seriously the issue of unobserved individual heterogeneity which affects the relationship between income, on the one hand, and reported levels of satisfaction, on the other.

The general problem lies in the interpretation of reported satisfaction, as this latter is a representation of the unobserved variable (utility) which really interests us. Interpreting subjective responses requires (i) relating discrete verbal satisfaction judgements to levels of a latent continuous satisfaction variable, and (ii) associating these levels to observable characteristics. Figure 1 illustrates the process.

Figure 1. Heterogeneity problems with subjective variables



An explanatory variable, such as income, is correlated with (unobservable) utility. Individual heterogeneity likely plays a role at this point, in the sense that the link between observable variables (income for instance) and latent satisfaction is not the same for all individuals, *i.e.* the parameters of the individual utility function are not the same across individuals (Tinbergen, 1991, and Sen, 1992)[2]. The right-hand side of the figure shows the transformation of utility into reported satisfaction levels; again the relationship between verbal satisfaction labels and latent utility is unlikely to be the same for everybody. In either, or both, of these cases, any interpretation of subjective answers will be misleading (Winkelmann and Winkelmann, 1998). There is an obvious parallel between the two sides of the above diagram and the phenomena of the hedonic and satisfaction treadmills underscored in Danny Kahneman's work (see Kahneman, 2000). We will carry out a joint test for the presence of any heterogeneity in the above diagram.

Data

We use data from the European Community Household Panel (ECHP). The ECHP survey was conducted annually in EU Member States over the period 1994 - 2001. In the first wave, in 1994, a sample of some 60,500 nationally represented households - approximately 130,000 adults aged 16 years and over - were interviewed in the then 12 Member States. Austria joined the survey in 1995 and Finland in 1996. The ECHP is an extensive, sample-based panel survey in which the same households and individuals are interviewed annually. Data from the surveys are available on the main aspects of welfare, including income and employment, housing, education, social relationships, health. The data come from a standardised questionnaire and are designed to be cross-nationally comparable. Details of the ECHP are available on the Eurostat web site[3].

We take a 20% random sample of the data, with drawing probabilities weighted by country and satisfaction level. We have three waves of the ECHP data, 1994-96. This yields a balanced sample of 146853 observations. (48951 individuals over three waves and 12 countries).

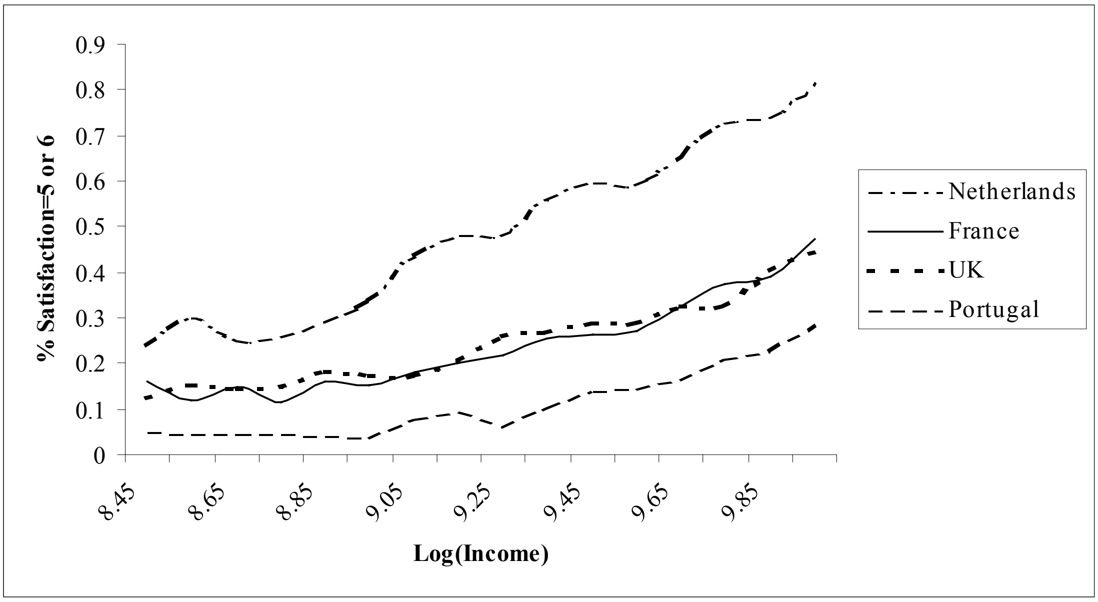
Our key variables are satisfaction, which is our proxy measure of utility, and income. The former is measured by satisfaction with financial situation, on a scale of one to six in increasing order of satisfaction. The latter is given by net household income in Euros, converted between countries using PPPs. This income is further converted into a household equivalent measure, using the modified OECD scale: weights of 1 for the first adult, 0.5 for subsequent adults (aged over 14), and 0.3 for children. The first and last percentile of the distribution of raw household income have been dropped, due to worries about the accuracy of the reported data. The distribution of all variables is presented in the first column of Appendix A.

The figure below shows the results from non-parametric estimation of the probability of being satisfied (reported satisfaction of 5 or 6 on the six-point scale) on the log of per capita income. Results from four countries are shown: the Netherlands, the UK, France and Portugal.

The estimated relationship is mostly positive, but it is obvious that neither the intercepts nor the slopes are the same across countries (although the French regression line is remarkably close to the English).

Figure 2. Per Capita Income and Well-Being in Europe:

Some Non-Parametric Estimations



Our suspicion is that this graph shows heterogeneity in the relationship between income and reported well-being. However, bivariate correlations cannot prove anything, due to composition effects. The remainder of the paper presents and applies a multivariate model which provides a robust test of this hypothesis.

3 Modelling Heterogeneity in Well-being

Our statistical model applies the latent class approach, in which sub-groups of the population are identified endogenously, to an ordered dependent variable, here satisfaction with financial situation. While this type of model is widely-used in other disciplines, it is uncommon in Economics.

3.1 Econometric modelling

Consider an agent i who evaluates her well-being at time t using P different "naturally" ordered labels such as excellent, very good, good etc. Denote a_{it} her answer, which belongs to the ordered set of labels $L=\{l_1,l_2,...,l_P\}$. The most common way to model this choice assumes that there exists an underlying real-valued latent well-being index $WB^*_{i,t}$ and *P+1* ordered individual threshold parameters $s^{0*}_{i,t}=-\infty, s^{1*}_{i,t},..., s^{p*}_{i,t},..., s^{P*}_{i,t}=+\infty$ such that:

$$a_{it}=l_p \Leftrightarrow s^{p-1*}_{i,t} \le WB^*_{i,t} < s^{p*}_{i,t} \qquad \textbf{(1)}$$

Assume for the moment that only income affects well-being. We then consider the simple following relation between these variables:

$$WB^*_{i,t}=\alpha^*_{i,t}Y_{i,t}+\beta^*_{i,t} \qquad \textbf{(2)}$$

wherein $\alpha^*_{i,t}$ and $\beta^*_{i,t}$ are individual and potentially time-varying parameters. In this general model, heterogeneity is twofold, firstly because the "marginal utility" of income ($\alpha^*_{i,t}$) and the baseline (intercept) level of well-being ($\beta^*_{i,t}$) are individual-specific, and secondly because individuals may use different labels to express the same level of well-being. This is why the thresholds points, ($s^{1*}_{i,t},..,s^{P*}_{i,t}$), have an *i* subscript above. This second heterogeneity may reflect anthropological variations in attitudes towards pleasure and pain.

Henceforth, $Y_{i,t}$ includes income and all variables that may affect well-being. Adding a normal zero mean i.i.d. error term $\tilde{\varepsilon}^*_{i,t}$ to the right-hand side of equation **(2)** yields:

$$WB^*_{i,t}=\alpha^*_{i,t}Y_{i,t}+\beta^*_{i,t}+\tilde{\varepsilon}^*_{i,t} \qquad \textbf{(3)}$$

Some further assumptions are needed to identify this model from the data. First, we suppose that the parameter vector $\vec{v}^*_{i,t}$=($\alpha^*_{i,t},\beta^*_{i,t},s^{1*}_{i,t},...,s^{p*}_{i,t},...,s^{P-1*}_{i,t}$) is distributed over a finite number of points *C*. Hence the distribution of the heterogeneity parameters is discrete and the correlation between the distributions of any pair of parameters is set to 1. Second, the parameters are constant over time. As they are now randomly distributed we will note them $\tilde{\vec{v}}^*_i$ =($\tilde{\alpha}^*_i,\tilde{\beta}^*_i\ \ \tilde{s}^{1*}_i,...,\tilde{s}^{P-1*}_i$). Equation **(3)** thus becomes:

$$WB^*_{i,t}=\tilde{\alpha}^*_i Y_{i,t}+\tilde{\beta}^*_i+\tilde{\varepsilon}^*_{i,t} \qquad \textbf{(4)}$$

The vector $\tilde{\vec{v}}^*_i$ is distributed on C points of $\Re^{P+1}$, with associated conditional probabilities of $\omega_{i,c}$=Pr($\tilde{\vec{v}}^*_i=\vec{v}^*_c$|$Y_{i,1}$, …,$Y_{i,T}$): the value of the vector is $\vec{v}^*_c$=($\alpha^*_c,\beta^*_c,s^{1*}_c,...,s^{P-1*}_c$) with probability ω_c. The data provide us with the empirical probabilities Pr($a_{i,1}$, …,a_{iT}|$Y_{i,1}$,

$,\ldots,Y_{i,T})$. By integration over the support of the distribution of the random vector $\tilde{\vec{v}}_i^*$, we obtain the following decomposition:

$$\Pr(a_{i,1},\ldots,a_{,iT}\,|\,Y_{i,1},\ldots,Y_{i,T}) = \sum_{c=1}^{C}\Pr(a_{i,1},\ldots,a_{i,T}\,|\,Y_{i,1},\ldots,Y_{i,T},\tilde{\vec{v}}_i^* = \vec{v}_c^*)\Pr(\tilde{\vec{v}}_i^* = \vec{v}_c^*\,|\,Y_{i,1},\ldots,Y_{i,T}) \quad \textbf{(5)}$$

We now suppose that answers $a_{i,j}$ and $a_{i,k}$ at different periods *j* and *k* are independent given $\tilde{\vec{v}}_i^*$ and the $Y_{i,}$'s. Further, $a_{i,j}$ and $Y_{i,k}$ are independent given $\tilde{\vec{v}}_i^*$ and $Y_{i,j}$. Equation **(5)** becomes:

$$\Pr(a_{i,1},\ldots,a_{i,T}\,|\,Y_{i,1},\ldots,Y_{i,T}) = \sum_{c=1}^{C}\omega_{i,c}\left\{\prod_{t=1}^{T}\Pr(a_{i,t}\,|\,\tilde{\vec{v}}_i^* = \vec{v}_c^*, Y_{i,t}\right\} \quad \textbf{(6)}$$

For the sake of parsimony, we assume independence between $\tilde{\vec{v}}_i^*$ and $Y_i=(Y_{i,1},\ldots,Y_{i,T})$. The probabilities associated with each vector $\vec{v}_c^*$ are modelled as[4]:

$$\omega_{i,c} = \Pr(\tilde{\vec{v}}_i^* = \vec{v}_c^*\,|\,Y_i) = \frac{\exp(\gamma_c)}{\sum_{c'=1}^{C}\exp(\gamma_{c'})} \quad \textbf{(7)}$$

with $\gamma_1=0$. Last, assuming that $\tilde{\vec{v}}_i^*$ and $\tilde{\varepsilon}_{i,t}$ are independently distributed conditional on $Y_{i,t}$, we model $\Pr(a_{i,t}\,|\,\tilde{\vec{v}}_i^* = \vec{v}_c^*, Y_{i,t})$ according to the simple ordered probit model in equation **(4)**:

$$\Pr(a_{i,t}\,|\,\tilde{\vec{v}}_i^* = \vec{v}_c^*, Y_{i,t}) = \prod_{p=1}^{P}\left[\Phi(s_c^{p*} - \alpha_c^* Y_{i,t} - \beta_c^*) - \Phi(s_c^{p-1*} - \alpha_c^* Y_{i,t} - \beta_c^*)\right]^{1\{a_{i,t}=l_p\}} \quad \textbf{(8)}$$

It is easily seen from equation **(8)** that the β_c^* (the individual fixed effect) and the threshold parameters can not be separately identified. Hence, β_c^* is normalised to 0. The variance of the error-term $\tilde{\varepsilon}_{i,t}^*$ is also not identified. It is normalised to 1. We are able to identify the following quantities:

$$\begin{aligned} &\forall(p,c), s_c^p = \frac{s_c^{p*} - \beta_c^*}{\mathrm{var}(\tilde{\varepsilon}_{i,t}^*)} \\ &\forall c, \alpha_c = \frac{\alpha_c^*}{\mathrm{var}(\tilde{\varepsilon}_{i,t}^*)} \end{aligned} \quad \textbf{(9)}$$

The parameters α_c, s_c^p and γ_c are obtained through the maximisation of the following log-likelihood:

$$\sum_i \log\left\{\sum_{c=1}^{C}\left(\frac{\exp(\gamma_c)}{\sum_{c'=1}^{C}\exp(\gamma_{c'})}\prod_{t=1}^{T}\left[\prod_{p=1}^{P}\left[\Phi(s_c^p-\alpha_c Y_{i,t})-\Phi(s_c^{p-1}-\alpha_c Y_{i,t})\right]^{1\{a_{i,t}=1_p\}}\right]\right)\right\} \qquad \mathbf{(10)}$$

Our parametric approach for modelling heterogeneity is deeply rooted in the latent class analysis literature, and its applications in psychometry (see Uebersax, 1999, for instance). One straightforward interpretation of the model is in terms of mixtures of distinct subgroups or classes of the population. The observed sample is considered as a mixture of several classes of individuals, who differ in their latent ability to transform income into utility, and by the way they express their well-being. We posit that everyone has a conditional probability $\omega_{i,c}$ of belonging to class *c*. However, the class membership of individuals remains unobserved, and we have a standard problem of missing data. This is solved by using a variant of the standard iterative EM (Expected Maximisation) algorithm for missing data (Dempster *et al.*, 1977), the Simulated Annealing EM algorithm (Celeux *et al.*, 1995), which allows a better detection of a global maximum of the sample likelihood and avoidance of saddle points. An interesting by-product of any EM algorithm is a fuzzy classification of observations into the classes. For each individual, we compute the following ex-post conditional probability:

$$w_{i,c}=\Pr(\tilde{\vec{v}}_i=\vec{v}_c \mid a_{i,1},\ldots,a_{i,T},Y_{i,1},\ldots,Y_{i,T})=\frac{\omega_{i,c}\Pr(a_{i,1},\ldots,a_{i,T}\mid \tilde{\vec{v}}_i=\vec{v}_c,Y_{i,1},\ldots,Y_{i,T})}{\Pr(a_{i,1},\ldots,a_{i,T}\mid Y_{i,1},\ldots,Y_{i,T})} \qquad \mathbf{(11)}$$

where $\vec{v}_c=(\alpha_c,s_c^1,\ldots,s_c^{P-1})$ are the point estimates of the parameter vectors for class *c*. Hence, we take advantage of our distributional assumptions to build an ex-post fuzzy classification of the population into *C* different classes or groups.

Perhaps more important is the problem of theoretical identification: is it possible to find several set of parameters – thus several mixtures – that would fit the data just as well (*i.e.* produce the same likelihood)? Uebersax (1999) proposes an order condition: the number of parameters in the model should be less than the number *R* of empirical patterns of response. Here, we have *C* slope parameters, *C-1* unconditional probabilities ω_c (the *C* weights sum up to 1) and *(P-1)*C* threshold parameters, making a total of *C(P+1)-1*. There are P response modalities, and T waves. Moreover, conditional on class membership, the probabilities of response are time-independent. Hence, $R=\frac{(P+T-1)!}{T!(P-1)!}$.With *P=6* and *T=3*, we have *R=56*, so that the order condition inequality becomes 7C-1 < 56, and the maximum number of classes we can identify is 8[5].

To obtain the empirical optimal number of classes, we compare information criteria such as the entropy of the fuzzy classification, the BIC and the AIC for 1, 2 or more points. The BIC and AIC scores are commonly used in order to balance the gain in log-likelihood through an increase of *C*, and the loss of degrees of freedom from the greater number of parameters (see, for instance, Deb and Trivedi, 1997). We also use an entropy criterion proposed by Jedidi *et al.* (1997) to assess the accuracy of the fuzzy classification.

3.2 Testable assumptions

Our ambition is to challenge somewhat the existing literature, by showing that heterogeneity is present in the process of well-being production, but is also created by the instrument used to measure of well-being (the standard question about subjective well-being). The two main assumptions made in the literature to date are that (i) there is no heterogeneity in the slope parameters $\alpha^*_{i,t}$ (i.e. the marginal effect of $Y_{i,t}$ on the latent well-being index is the same), and (ii) the threshold parameters $(s^{1*}_{i,t},..,s^{P-1*}_{i,t})$ should not vary across the population, whatever the heterogeneity of the well-being latent index $WB^*_{i,t}$. In the estimates, we find that condition (i) is false for *our* model. However, we are fully aware that this claim is contingent to a strong assumption, namely the independence of the distribution of the parameter and the distribution of the error-term.

Indeed, consider a model wherein the distribution of the error-term is conditioned on class membership, and class-conditional error-terms are independent. Instead of the following class conditional well-being relation for each class: $\forall c = 1,...C,\quad WB^*_{i,t|i\in c} = \alpha^*_c Y_{i,t} + \beta^*_c + \tilde{\varepsilon}_{i,t}$, we have $WB^*_{i,t|i\in c} = \alpha^*_c Y_{i,t} + \beta^*_c + \tilde{\varepsilon}^c_{i,t}$. The latter model[6] can not be distinguished from our model, since its likelihood is the same (in both cases the variance(s) of the error-term(s) are not identified). Even if we add other regressors, the normalisation of the variance of the residuals $\tilde{\varepsilon}^c_{i,t}$ in this model would prevent us from proving the heterogeneity of the slope parameters: we may have $\alpha_c \neq \alpha_{c'}$ for some c and c' but $\alpha^*_c = \alpha^*_{c'}$. Hence, we can not prove that the slope parameters are heterogeneous, only that there is an indeterminacy concerning proposition (i).

The latent index $WB^*_{i,t}$ is identified up to an affine transformation[7] (see equation **9**). Hence, for all *c*, c' (c'≠c), $s_c^{p*} = s_{c'}^{p*}$ for all *p* if and only if there exists non-zero λ_1 and λ_2 such that $s_c^p = \lambda_1 s_{c'}^p + \lambda_2$. Therefore, the way individuals express their well-being is heterogeneous if thresholds for class *c* are not an affine transformation of thresholds for class *c'*. This condition holds if and only if the column vectors of thresholds for each class $\vec{s}_c$, $\vec{s}_{c'}$ and the (P-1)x1 vector of 1, $\vec{1}$, are not multicollinear. We use the test of rank proposed by Robin and Smith (2000) to test this condition for each couple (c, c'). We are thus able to show that even if proposition (i) is true, proposition (ii) is false.

Ultimately, it is not possible to disentangle the effects of a set of control variables ($Y_{i,t}$ plus other regressors) on the latent well-being index on one hand, and the expression of this well-being on the other hand. For such a task, we would need restrictions of identification, *i.e.* a variable that affects the response thresholds but not "utility" (Pudney and Shields, 2000).

4 Results

Information criteria and entropy measures suggest that there are at least four classes. We consider keep four classes, as we do not wish to overload the model with a very large number of competing outcomes.

Table 1 presents the results relating satisfaction to income and standard demographic variables, both in the whole sample and in each of the four groups.

Table 1. Regression results, ordered probit models

	Ordered probit	**Latent class ordered probit model – 4 classes**			
		Class 1	**Class 2**	**Class 3**	**Class 4**
Latent index parameters: α_c					
Ln(income)	0.629** (0.006)	0.762** (0.002)	0,706** (0,001)	0,597** (0,004)	0,517** (0,006)
Age/10	-0.171** (0.011)	-0.214** (0.007)	-0,186** (0,007)	-0,188** (0,012)	-0,207** (0,025)
Age-squared/100	0.027** (0.001)	0.033** (0.001)	0,030** (0,001)	0,034** (0,001)	0,030** (0,002)
Number of children at home under age 16.	0.023** (0.006)	-0.021** (0.004)	0,040** (0,005)	0,046** (0,009)	-0,035 (0,018)
Number of children-squared	-0.004* (0.002)	0.009** (0.001)	-0,020** (0,002)	-0,008** (0,003)	0,014** (0,004)
Male	-0.062** (0.006)	-0.058** (0.003)	-0,093** (0,003)	-0,120** (0,007)	-0,060** (0,012)

Marital status(ref: single & never married)					
Married.	0.120** (0.010)	0.248** (0.006)	0.213** (0.007)	0.079** (0.012)	0.173** (0.025)
Living in couple	-0.071** (0.015)	0.027** (0.007)	0.051** (0.008)	-0.170** (0.016)	-0.029 (0.037)
Widowed, separated, divorced.	-0.213** (0.013)	-0.190** (0.008)	-0.231** (0.008)	-0.280** (0.018)	-0.150** (0.033)
Education (ref: less than secondary education)					
Higher Education	0.167** (0.009)	0.332** (0.005)	0.341** (0.005)	0.265** (0.012)	0.264** (0.018)
Secondary Education	0.112** (0.007)	0.193** (0.004)	0.164** (0.004)	0.225** (0.008)	0.129** (0.014)
Labour force status (ref: inactive)					
Works over 15hrs per week.	0.106** (0.007)	0.239** (0.004)	0.263** (0.005)	0.178** (0.009)	0.122** (0.016)
Works under 15hrs per week.	-0.051** (0.017)	-0.026** (0.009)	-0.080** (0.011)	-0.024 (0.018)	-0.147** (0.050)
Unemployed	-0.809** (0.014)	-0.868** (0.017)	-0.905** (0.010)	-0.669** (0.017)	-0.814** (0.033)
Country effect (ref: France)					
Germany	0.139** (0.013)	1.392** (0.014)	-1.695** (0.014)	2.340** (0.042)	-1.128** (0.041)
Belgium– Luxembourg	0.313** (0.014)	1.180** (0.014)	-2.362** (0.013)	2.010** (0.039)	1.013** (0.038)
Netherlands	0.706** (0.013)	1.131** (0.014)	1.125** (0.006)	0.668** (0.038)	1.154** (0.040)
Denmark	0.848** (0.016)	2.391** (0.014)	-0.889** (0.008)	3.726** (0.037)	-0.598** (0.045)
United Kingdom	0.125** (0.011)	0.023 (0.017)	-0.172** (0.006)	2.316** (0.039)	-1.283** (0.045)
Ireland	0.127** (0.015)	1.526** (0.014)	-1.732** (0.016)	0.040 (0.038)	0.628** (0.041)
Italy	-0.161** (0.011)	0.761** (0.014)	-2.123** (0.026)	1.925** (0.037)	-1.364** (0.044)
Portugal	-0.013 (0.013)	0.719** (0.022)	-1.703** (0.031)	0.189** (0.059)	0.482** (0.043)
Spain	-0.083** (0.012)	-0.990** (0.086)	-0.982** (0.007)	2.033** (0.038)	-0.389** (0.042)
Greece	-0.165** (0.013)	0.434** (0.018)	-2.072** (0.036)	1.765** (0.037)	-1.153** (0.052)
1994	0.076** (0.007)	0.109** (0.004)	0.010** (0.004)	0.187** (0.008)	0.042 (0.022)
1995	0.040** (0.007)	0.034** (0.005)	0.011** (0.005)	0.103** (0.009)	0.048* (0.022)
Threshold parameters: $s_c^1,...,s_c^{P-1}$					
Cut 1	4.334** (0.056)	5.295** (0.004)	3.178** (0.004)	5.425** (0.005)	3.223** (0.007)
Cut 2	5.007** (0.056)	6.198** (0.007)	4.323** (0.008)	6.181** (0.008)	3.847** (0.014)
Cut 3	5.801** (0.056)	7.585** (0.011)	5.453** (0.012)	6.858** (0.011)	4.376** (0.019)
Cut 4	6.630** (0.057)	9.045** (0.012)	6.751** (0.014)	7.504** (0.015)	4.912** (0.021)

Cut 5	7.649** (0.057)	10.686** (0.005)	9.300** (0.005)	8.677** (0.009)	5.663** (0.015)
% of the sample	100 %	37.43%	33.96%	15.52%	13.09%
BIC	-225493	-211037			

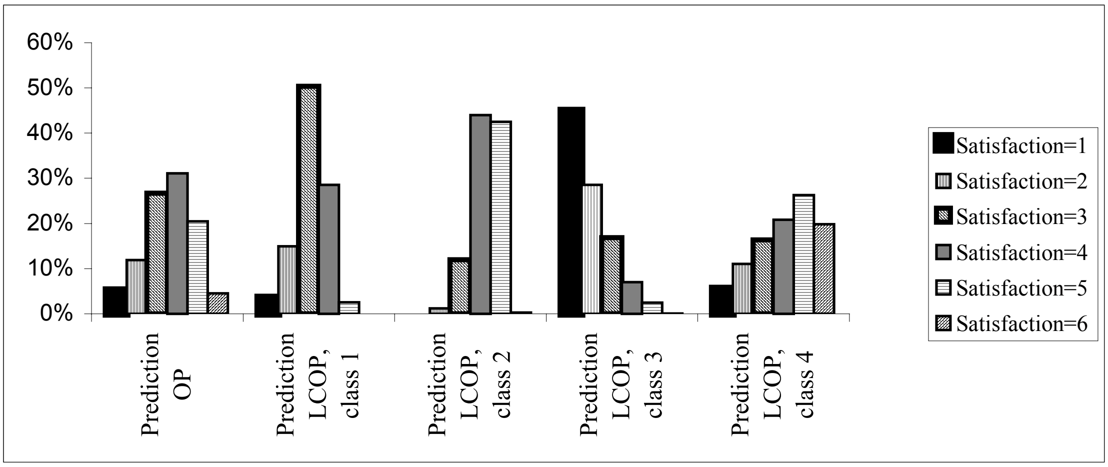
Note: standard errors in parentheses; *= significant at the 5% level; ** at the 1% level. Robust standard errors clustered on households.

Table 1 shows that satisfaction is U-shaped in age, lower for men and the unemployed, but higher for the married, the better-educated, and for full-time workers. Most of these partial correlations are well-known in the literature.

Rank tests reject the hypothesis that the cut points for one group are an affine transformation of the thresholds for another group[8]. Hence, the way individuals qualitatively assess their well-being varies greatly in the sample or/and the well-being functions are heterogeneous. Table 1 suggests that, in this simple set up, there are very sharp differences in the effect of income on declared satisfaction.

We are interested in the well-being/income relationships within each class, conditional on the other control variables. Figure 3 shows the predicted probabilities of reporting a given satisfaction level, for the "average" agent (having the average sample demographic characteristics), conditional on her membership of class *c* (c=1,…,4; LCOP means Latent Class Ordered Probit Model).

Figure 3: Predicted satisfaction probabilities (computed at $\bar{X}$, $\bar{I}$)



Compared to the predicted distribution from an ordered probit without heterogeneity, shown on the left, those in classes 2 and (to an extent) 4 are more likely satisfied, while those in class 3 are more likely dissatisfied.

To interpret the effect of income on reported well-being, we need to take into account both the estimated coefficient on income, and the cut-points. We also hold all other characteristics constant (at the sample mean). Last, higher income will not induce a large movement of probability mass out of low satisfaction levels if there were very few people there to start with. Equally, money can't buy (reported) happiness if most people are already very happy. A normalisation is called for which controls for the different initial distribution of reported satisfaction between classes.

Table 2 presents the results from one such normalisation, using:

$$\overline{\Delta}(\text{sat}) = \left\{ \Pr(\hat{WB}_{\bar{i}} \leq \text{sat} \mid \bar{i} \in c, \overline{X}, \ln(1.01 * \bar{I})) - \Pr(\hat{WB}_{\bar{i}} \leq \text{sat} \mid \bar{i} \in c, \overline{X}, \ln(\bar{I})) \right\} \Big/ \Pr(\hat{WB}_{\bar{i}} \leq \text{sat} \mid \bar{i} \in c, \overline{X}, \ln(\bar{I}))$$

Table 2: Normalised marginal income effects (computed at $\overline{X}, \bar{I}$***)***

Latent class ordered probit	$\overline{\Delta}$ **(1)**	$\overline{\Delta}$ **(2)**	$\overline{\Delta}$ **(3)**	$\overline{\Delta}$ **(4)**	$\overline{\Delta}$ **(5)**
LCOP – Class 1	-14.9%	-10.1%	-3.8%	-0.5%	-0.01%
LCOP – Class 2	-21.9%	-16.2%	-10.5%	-4.6%	-0.1%
LCOP – Class 3	-5.0%	-2.6%	-1.1%	-0.4%	-0.02%
LCOP – Class 4	-9.6%	-7.2%	-5.4%	-3.6%	-1.7%

This table should be interpreted as follows. The percentage figures show the change in the probability of having satisfaction lower or equal than the number in parentheses in the column head; these can be thought of as exit rates from low satisfaction. A one per cent rise in income decreases the probability that someone in class 2 (the "happy" class, from Figure 2) has satisfaction of three or lower (on the one to six scale) by 10.5 percentage points. On the contrary, it has little effect on the same probability for someone in class 3 (the "unhappy").

The results here are unambiguous. The effect of income on subjective well-being depends on unobserved heterogeneities relating either to the underlying utility functions or to the way people label their utility. Further one group (class 2) is both highly satisfied and has large marginal well-being effects of income, while another (class 3) is the least satisfied and has the lowest marginal well-being effect of income. Groups one and four occupy intermediate positions.

Appendix A shows the distribution of observable characteristics across the four well-being classes in Table 1. Taking the two classes of most interest, we see that those in class 2 (satisfied, high marginal well-being effect of income) are conformist, in the sense that they are mostly close to sample mean characteristics. Those in class 3 (dissatisfied, low marginal well-being effect of income) are less likely to be married and in full-time work; they are also somewhat older, and are the richest of the four classes.

Table 3 shows predicted class membership by country. There is some initial *prima facie* evidence for groupings of countries. We concentrate on classes 2 (the satisfied with greater sensitivity to income) and 3 (the dissatisfied who are less sensitive to income).

Table 3: Comparison of within-class proportions

	Class 1	Class 2	Class 3	Class 4
Belgium	37.7%	28.4%	18.2%	15.6%
Luxembourg	40.8%	26.3%	18.9%	14.0%
Netherlands	37.6%	33.3%	14.8%	14.2%
Denmark	37.0%	29.2%	19.2%	14.6%
Germany	36.5%	34.5%	16.0%	12.9%
UK	41.3%	36.0%	11.7%	10.9%
Ireland	29.9%	31.6%	18.0%	20.5%
France	38.1%	34.4%	16.0%	11.5%
Italy	36.3%	35.1%	15.6%	13.0%
Portugal	43.3%	37.0%	11.8%	8.0%
Spain	33.9%	33.2%	17.6%	15.3%
Greece	36.2%	34.5%	16.1%	13.1%
Whole sample	*37.4%*	*34.0%*	*15.5%*	*13.1%*

It is noticeable in particular that individuals from Belgium, Luxembourg, Denmark, and Ireland are under-represented in class 2 but over-represented in class 3; the opposite holds for individuals from Portugal and the UK. These different groups of countries being associated with different rates of transformation of income into reported well-being, we expect their preferences to differ in terms of redistribution and political economics. Four of the six founding members of the European Union, Germany, Italy, the Netherlands and France, are closest to the European average; Ireland, Luxembourg and Portugal are the least representative.

More generally, there is a great deal of heterogeneity within countries as well, and predicted class membership at the individual level will likely correlate with various behaviours. We believe that future applied work in microeconomics will increasingly take slope heterogeneity into account in order to better model individual behaviour.

5 Conclusion

This paper modelled the relationship between income and self-reported well-being using random-effect techniques applied to panel data from twelve European countries. We show that people are different, and in more complicated ways than just having different intercepts. We are not able to distinguish between heterogeneity in the utility function (translating income into utility) and heterogeneity in the expression function (turning utility into reported well-being). We can, however, strongly reject the hypothesis that individuals carry out these joint transformations in the same way.

We identify four classes of individuals, and show that the "marginal well-being effect of income" is very different across these classes. In particular one class is satisfied and has a high marginal well-being effect, while another is dissatisfied and has a low marginal well-being effect. Descriptive statistics reveal demographic and country patterns between classes. This has at least two important implications. First, in a political economy sense, as the effect of income differs sharply across classes (and classes are not independently distributed between countries), we would expect average opinion regarding economic policies to differ across countries. To the extent that we have identified country groups in Table 3, we *a priori* expect these groups to vote similarly with respect to European-level reforms, and to behave differently. This is a subject for ongoing research.

Perhaps more importantly, our results suggests that aggregating data across diverse populations may be a dangerous practice. Individuals, who seem to fall naturally into a number of different classes, differ in ways that are far more complicated than those picked up by a simple fixed effect. The trend towards comparative research in social science, whereby data from different countries are compared, is laudable. Nonetheless, our results suggest that the blind aggregation of diverse populations risks producing empirical results that are false for everybody.

Footnotes

---

[1] The GHQ-12 score, used by Clark and Oswald (1994), is an example of the latter.

[2] Lelkes (2002) uses Hungarian data to show that the marginal utility of income is lower for the religious than for the non-religious.

[3] http://forum.europa.eu.int/irc/dsis/echpanel/info/data/information.html

[4] Alternatively, we can specify $\Pr(\tilde{\vec{v}}_i^* = \vec{v}_c^* | Y_i)$ as multinomial logit probabilities, with $Y_i=(Y_{i,1},\ldots,Y_{i,T})$ as regressors to take into account any possible correlation between the distributions of heterogeneity and the control variables. Such a distributional assumption is as arbitrary as is our independence assumption.

[5] While this condition is obviously necessary, Uebersax (1999) does not prove its sufficiency. Intuitively, identification also requires the presence of a continuous right-hand side regressor (for instance income). Otherwise, it is always possible to classify individuals perfectly according to their response patterns and any set of discrete characteristics.

[6] This model is a true latent class model, whereas our model is strictly speaking a random-coefficient model, with coefficients distributed according to a finite discrete distribution.

[7] Of course, this is true conditional on our choice of a linear functional form.

[8] The critical values for the statistics proposed by Robin and Smith (2000) were obtained after simulation of their distribution functions. Multicollinearity was overwhelmingly rejected for all pairs of classes. Given the precision of the estimates, it is not surprising that critical values for the test statistics are all very close to 0. The lowest value we obtained is 128.1 for classes 2 and 3. Other statistics are available from the authors.

Appendix A: Descriptive statistics

The table is to be read as follows. The percentage figures show the probability of having the demographic characteristic in question conditional on belonging to the different classes. For example, 67% of respondents in class 1 are married, compared to 61% of respondents in class 4.

Variable	Whole Sample	Class 1	Class 2	Class 3	Class 4
Ln(income)	9.126	9.135	9.110	9.158	9.106
Age	46.3	45.8	46.5	47.2	46.3
Number children	0.650	0.673	0.648	0.623	0.620
Male	47.4%	48.4%	47.2%	46.2%	46.3%
Married	65.1%	66.7%	65.9%	63.2%	60.8%
Living together	4.3%	4.5%	4.3%	3.6%	4.5%
Widow,sep.,div.	10.8%	10.2%	10.5%	12.5%	11.4%
Higher Education	15.3%	15.9%	15.4%	14.0%	14.7%
Sec. Education	29.2%	29.6%	28.2%	30.1%	29.6%
Over 15hrs/week	50.3%	54.0%	51.1%	44.8%	44.2%
Under 15hrs/week	2.8%	2.8%	2.8%	2.7%	3.0%
Unemployed	5.4%	4.7%	5.2%	5.8%	7.2%
Belgium	5.0%	5.0%	4.2%	5.9%	6.0%
Luxembourg	1.5%	1.7%	1.2%	1.9%	1.6%
Netherlands	7.5%	7.6%	7.4%	7.2%	8.2%
Denmark	4.2%	4.2%	3.6%	5.2%	4.7%
Germany	7.7%	7.5%	7.8%	7.9%	7.6%
UK	12.7%	14.0%	13.4%	9.6%	10.6%
Ireland	5.1%	4.1%	4.7%	5.9%	8.0%
France	10.6%	10.7%	10.7%	10.9%	9.3%
Italy	14.4%	14.0%	14.9%	14.5%	14.4%
Portugal	9.4%	10.9%	10.2%	7.1%	5.7%
Spain	12.3%	11.1%	12.0%	14.0%	14.4%
Greece	9.6%	9.3%	9.7%	9.9%	9.6%
1994	33.3%	33.3%	33.3%	33.3%	33.3%
1995	33.3%	33.3%	33.3%	33.3%	33.3%
Satisfaction = 1	10.3%	6.0%	7.8%	13.9%	24.6%
Satisfaction = 2	14.0%	10.1%	18.5%	12.5%	15.4%
Satisfaction = 3	24.3%	30.7%	25.8%	14.3%	13.7%
Satisfaction = 4	25.8%	34.7%	25.5%	15.5%	12.9%
Satisfaction = 5	19.3%	16.3%	21.2%	26.2%	14.8%
Satisfaction = 6	6.4%	2.2%	1.2%	17.6%	18.6%

Note: There are 146853 observations, representing 48951 individuals over three waves.

References

Celeux, G., Chauveau, D. & Diebolt, J. (1995). "On Stochastic Versions of the EM algorithm" , Working Paper n°2514, Paris: Institut National de Recherche en Informatique et en Automatique.

Clark, A.E. (2001). "What Really Matters in a Job? Hedonic Measurement Using Quit Data". *Labour Economics,* **8**, 223-242.

Clark, A.E. & Oswald, A.J. (1994). "Unhappiness and Unemployment". *Economic Journal,* **104**, 648-659.

Clark, A.E. & Oswald, A.J. (2002). "Well-Being in Panels". DELTA, mimeo.

Dempster, A.P., Laird, N.M. & Rubin, P.B. (1977), "Maximum-Likelihood from incomplete data via the E-M algorithm", *Journal of the Royal Statistical Society B*, 39, 1-38.

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

Freeman, R.B. (1978). "Job Satisfaction as an Economic Variable". *American Economic Review,* **68**, 135-141.

Jedidi, K., Jagpal, H.S. & DeSarbo, W.S. (1997). "Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity". *Marketing Science*, **16**, 39-59.

Kahneman, D. (2000). "Experienced Utility and Objective Happiness: A Moment-Based Approach". In D. Kahneman & A. Tversky (Eds.), *Choices, Values and Frames*. New York: Cambridge University Press and the Russell Sage Foundation.

Lelkes, O. (2002). "Tasting freedom: Happiness, religion and economic transition". London School of Economics, CASE Paper 59.

Lucas, R., Clark, A.E., Georgellis, Y., & Diener, E. (2003). "Re-Examining Adaptation and the Setpoint Model of Happiness: Reaction to Changes in Marital Status". *Journal of Personality and Social Psychology,* **84**, 527-539.

Pudney, S., & Shields, M. (2000), "Gender, Race, Pay and Promotion in the British Nursing Profession: Estimation of a Generalised Ordered Probit Model", *Journal of Applied Econometrics*, **15**, 367-399.

Robin, J-M., & Smith, R.J. (2000), "Tests of rank", *Econometric Theory*, **16**, 151- 175.

Sen, A. (1992). *Inequality Reexamined*. Oxford: Oxford University Press.

Senik, C. (2004). "Income and Well-Being: What Can we Learn from Subjective Data?", *Journal of Economic Surveys*, forthcoming.

Stutzer, A., & Frey, B.S. (2003). "Does Marriage Make People Happy, Or Do Happy People Get Married?". University of Zurich, mimeo.

Tinbergen, J. (1991). "On the Measurement of Welfare". *Journal of Econometrics,* **50**, 7-15.

Uebersax, J.S. (1999). "Probit latent class analysis with dichotomous or ordered category measures: conditional independence/dependence models", *Applied Psychological Measurement*, **23**, 283-297.

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

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.