WORKING PAPERS
The Centre for Research in Applied Economics (CRAE)

05052012, May 2012

Heterogeneity in ordered choice models: A review with applications to self-assessed health

William Greene
New York University
Mark N. Harris
Curtin University
Bruce Hollingsworth
Lancaster University
Timothy A. Weterings
Monash University

CRAE forms a collaborative applied economic research framework across Curtin University. By developing and enhancing new and existing research networks, it provides facilities for sharing research materials and data. CRAE also supports ongoing applied economic research activities relevant to the dynamic economic conditions of the local, regional, national and international concerns of our members through inter- and intra-disciplinary research.

Acknowledgments
This publication series is underwritten by Curtin University, the Curtin Business School and the School of Economics and Finance.

ISSN 1834-9536

Heterogeneity in Ordered Choice Models: A Review with Applications to Self-Assessed Health

William Greene
Department of Economics, Stern School of Business, New York University, New York

Mark N. Harris
Department of Econometrics and Quantitative Modelling, Curtin University, Perth

Bruce Hollingsworth
Division of Health Research, Lancaster University, UK

Timothy A. Weterings
Department of Econometrics and Business Statistics, Monash University, Clayton

Abstract Discrete variables that have an inherent sense of ordering across outcomes are commonly found in large datasets available to many economists, and are often the focus of research. However, assumptions underlying the standard Ordered Probit (which is usually used to analyse such variables) are not always justified by the data. This study provides a review of the ways in which the Ordered Probit might be extended to account for additional heterogeneity. Differing from other reviews in scope, application and relevance in economic settings, a series of issues pertaining to choices of variables, and the economic assumptions underlying each model are discussed in the context of measuring the underlying health of respondents. The models are applied to a wave of the Household, Income and Labour Dynamics in Australia survey, in order to check the appropriateness of such assumptions in an applied context.

Keywords: Ordered Choice Modelling, HOPIT model, Incorporating Heterogeneity, Maximum Likelihood, Self-Assessed Health

JEL classification: C01, C52

1 Introduction

Research involving analysis of discrete variables that have an inherent sense of ordering may summarise subjective aspects of individuals, such as self-evaluation of health and happiness or observable characteristics, such as employment status. Similarly, they may be used to censor consumption of tobacco, alcohol or illicit drugs. Although these variables may not be a perfect measure of the underlying variable that researchers are essentially attempting to analyse, they generally act as a good proxy. For example, variables resulting from subjective self-assessed health have been shown to be good indicators of underlying health under correct assumptions about the data generating process − on occasion even predicting mortality with more accuracy than objective measures (Idler and Benyamini, 1997; Benyamini, Leventhal, and Leventhal, 1999).

These variables may be used when the objective measurement of the underlying variable is considered too expensive, time-consuming or impractical. Indeed, most large datasets, such as the British Household Panel Survey, and the Household, Income and Labour Dynamics Survey of Australia (HILDA), do not administer medical examinations to determine levels of health. Thus, ordered variables appear crucial to the analysis of issues relevant to researchers in the areas of economics, health economics and other areas.

The economic frameworks usually utilised by researchers when analysing these variables are insufficient to account correctly for underlying data generating processes. For example, although ordinality is an assumption suitable for the use of ordered responses, researchers typically implicitly and restrictively assume cardinality, with the use of Ordinary Least Squares. Such models impose rigid structure on the data and potentially lead to incorrect conclusions being drawn. When measuring levels of health, for example, it is difficult to describe one person's underlying health to be twice as large as another's. Even when Ordered Probit models are used, thus alleviating the ordinality or cardinality issue, incorrect assumptions regarding the nature of responses given can inhibit correct conclusions being drawn from studies because of the common existence of reporting heterogeneity.

This paper is intended to help improve the quality of research on ordered choice variables by providing a review of extensions to the standard Ordered Probit model offered by the existing literature. While existing texts, including Greene and Hensher (2010), briefly discuss extensions while mathematically motivating the use of their contributions, this work differs by offering explanations of implied economic frameworks that would underlie resulting models. In particular, we focus on the context of self-assessed health as a specific example validating each of the extensions in theory, and discuss potential estimation problems that may arise. This research provides an intuitive explanation of assumptions required for each of the models to be valid.

In addition to providing a more applied description of the extensions, this paper also differs from other ordered choice literature reviews in its application and its scope of estimation. Estimation results are provided for most of the models discussed. Consequently, this review provides tools to assess trade-offs between complexity and computational efficiency in estimating these kinds of models. The estimation of models with unobserved heterogeneity is another important aspect of this paper, as other applications tend to limit unobserved heterogeneity to the latent regression function, in the form of individual effects. In contrast, this paper considers the lesser researched extension of unobserved threshold and variance heterogeneity. The estimation of a model combining these two areas of unobserved heterogeneity constitutes a further contribution. The discussion of results focuses on how the interpretation of model coefficients vary with choice of economic framework, the overall model fit, and the estimation and comparison of partial effects.

2 The Data and Self-Assessed Health Variable

Data for this analysis are obtained from the Household, Income and Labour Dynamics survey of Australia (HILDA) - a representative panel data set started in 2001. As will be discussed later, some of the variables that are of significant interest in our running application are not measured in every wave. Thus only a single wave (wave five) of HILDA is used for estimation. Depending on model specification, this leaves us with over 11,000 observations for estimation. [1]

The Data and Self-Assessed Health Variable

The key variable of interest in this study is self-assessed health (SAH) which, in HILDA, is the result of the question:

"In general, [what] *would you say your health is?"*

with respondents choosing one of the outcomes: *Poor*; *Fair*; *Good*; *Very Good*; or *Excellent*. Covariates assumed to influence the health of individuals are motivated with a specification by Hernandez-Quevedo, Jones, and Rice (2008), who suggest broad factors that influence levels of health include marital status, ethnicity, education, household dynamics, income, and age. Hernandez-Quevedo et al. (2008) provides an excellent motivation of these covariates, so an extensive review of covariates is not repeated here. These variables are consistent with those used in other studies of self-assessed health (Powdthavee, 2009; Carro and Traferri, 2010; Jones and Wildman, 2008). For parsimony, we keep the choice of variables limited to these socio-demographic variables.

Relevant variables retrieved from the HILDA dataset are described in Appendix 1. It should be noted that, similar to Hernandez-Quevedo et al. (2008), we consider a non-linear functional form for age (scaled by 100 years to aid the estimation process by avoiding near singularity in the Hessian matrix).

3 The Standard Ordered Probit Model

The ordered nature of the SAH variable provides strong economic motivation for the Ordered Probit (OP) model. As outlined by McKelvey and Zavoina (1975), the OP model assumes that the choice of outcome is driven by some latent utility, y_i^*, which is assumed to

[1]The data used in this paper was extracted using the Add-On package PanelWhiz for Stata. PanelWhiz (http://www.PanelWhiz.eu) was written by Dr. John P. Haisken-DeNew (john@PanelWhiz.eu). See Haisken-DeNew and Hahn (2006) for details. The PanelWhiz generated DO file to retrieve the data used here is available from Timothy Weterings upon request. Any data or computational errors in this paper are our own

be some linear (in parameters) function of covariates, x_i, in addition to a stochastic error term

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

For this review, we denote x_i as the row vector of covariates with a constant term, and x_i^0 as the same vector without a constant. Similar notation is used for other parameter vectors. Within the context of SAH, the latent utility would be assumed by the economist to be the level of underlying (true) health of individual i. Assuming a normally distributed error term, $u_i \sim N(0, \sigma^2 = 1)$ it follows that $y_i^* \sim N(x_i\beta, 1)$. Outcome probabilities are then determined via a comparison of this distribution to several threshold parameters, μ_j, $j = 0 \ldots J$, where $\mu_0 = -\infty$ and $\mu_J = \infty$. Other thresholds are estimated as part of the model. The probability of individual i selecting outcome j is thus

$$Pr(y_i = j|x_i) = \Phi(\mu_j - x_i\beta) - \Phi(\mu_{j-1} - x_i\beta), \;\; j = 1, \ldots, J$$

and

$$\mu_0 = -\infty, \mu_J = \infty.$$

The likelihood function is constructed as:

$$LogL = \sum_{i=1}^{n} \sum_{j=1}^{J} Ln \left(Pr(y_i = j|x_i) \right)^{I(y_i=j)}$$

and I() is and indicator function, taking the value 1 when the condition inside the brackets holds, and 0 otherwise. Threshold/constant identification issues are resolved by imposing $\mu_1 = 0$.

The model structure flows from the assumption that variation in underlying health solely drives responses to the SAH question, and that the effect of each covariate on underlying health is the same across individuals (i.e., the relative effects of similar individuals circumstances on underlying health are identical). However, factors other than underlying health may influence how individuals answer questions about their health. This merits the introduction of heterogeneity into the model that the standard OP model does not allow.

As the OP model is essentially a comparison of a probability distribution for latent utility and a set of thresholds, four obvious areas for heterogeneity are identified (as per Greene and Hensher 2010). Thresholds provide obvious scope for heterogeneity and other three components of the underlying utility distribution can be considered for additional heterogeneity: the distribution of the error term, the variance of the error term, and the parameters of the underlying utility function. Changes to the distribution of the error term are not considered here, due to the empirical tendency for minimal changes in parameter estimates to result from large increases in computational complexity (Greene and Hensher, 2010). For example, ordered logit or ordered gompertz models add complexity without insight and are not suitable for the useful extensions considered here.

4 Threshold Heterogeneity

The literature discussing approaches to accounting for heterogeneity in thresholds provides a useful way of thinking about the thresholds as their own propensities, that are functions of sets of covariates, z_i

$$\mu^*_{ij} = f(z_i).$$

Most of the significant advances in threshold heterogeneity can be summarised across three dimensions: the form of the threshold (linear or non-linear); the number of sources of threshold variation; and the types of threshold covariates included in the thresholds. Assumptions about the functional form of threshold propensities shape the focus of much of the threshold literature. One of the first assumptions about the functional form of the thresholds, made by Terza (1985), was linearity. Terza (1985) specifies

$$\mu_{ij} = \mu_j + z_i^0\gamma, \;\; j = 2 \dots J-1,$$

implying each threshold covariate influences the thresholds in the same way (i.e., the same direction and magnitude). If the first threshold varies in this way as well, this is a result referred to by Lindeboom and Doorslaer (2004) as *index shift*. This is characterised by the thresholds moving by the same amount and in the same direction as each other (such that their relative location does not change). In this specification, however, the spacing between the first threshold and the remaining thresholds changes. This is a constrained form of what Lindeboom and Doorslaer (2004) call *cut-point shift*, where the spacing between all thresholds is allowed to change.

Terza (1985) notes that if the threshold covariates are the same as in the latent regression ($x_i \equiv z_i$), the relevant probabilities are equivalent to the following (separating out the latent regression constant for clarity):

$$Pr(y_i = 1|x_i) = \Phi(-\beta_0 - x_i^0\beta)$$

$$Pr(y_i = 2|x_i) = \Phi(\mu_2 - \beta_0 - x_i^0\delta) - \Phi(-\beta_0 - x_i^0\beta)$$

$$Pr(y_i = j|x_i) = \Phi(\mu_j - \beta_0 - x_i^0\delta) - \Phi(\mu_{j-1} - \beta_0 - x_i^0\delta), \;\; j = 3 \dots J$$

where $\delta = \beta - \gamma$. It also follows from the linearity of the model that, when $x_i \equiv z_i$, an equivalent model can be estimated by allowing the covariates to influence only the first threshold, and forcing the other thresholds to be constant. Thus, the model by Terza (1985) can be considered to be equivalent to allowing heterogeneity through just one threshold.

The obvious alternative to having just one source of variation in the thresholds is to have the set of covariates influence each threshold separately, as exhibited by the *generalised ordered probit* model of Pudney and Shields (2000)

$$\mu_{ij} = \mu_j + z_i^0\gamma_j, \;\; j = 2 \dots J-1.$$

When $x_i \equiv z_i$ the effect of the above specification is essentially to give each probability its own set of parameters:

$$Pr(y_i \leq j|x_i) = \Phi(\mu_j - \beta_0 - x_i^0\delta_j)$$

where $\delta_j = \beta - \gamma_j$, and $\gamma_1 = 0$ (by assumption). See Boes and Winklemann (2006) for an application of this specification. The specification implies that the generalised ordered probit model can be estimated as a series of binary ordered probits. The hypothesis that the coefficients for each of the binary probits are the same can then be tested (Lindeboom and Doorslaer, 2004; Schneider, Pfarr, Schneider, and Ulrich, 2011). If the null hypothesis of this test were not rejected the conclusion would be that the covariates exhibit index shift. However, in order to interpret this result, thought must be given to the intended influence of the variables that are tested. In particular, if a variable is found to have an index shift effect, this does not necessarily mean that the variable belongs in the latent regression Hernandez-Quevedo et al. (2008) - the effect is still not identified. For example, if an index effect for a positive coefficient on income is found, this might imply that higher incomes are related to higher health only or related to lower reporting thresholds only or imply a combination of the two effects. If testing yields the presence of cut-point shift (changes in the relative positions of thresholds), however, one might reasonably conclude the presence of reporting heterogeneity related to those variables.

4.1 Specification Issues

The above specification raises two major issues. First, probabilities are not constrained between zero and one, as the ordering of thresholds is not enforced. Second, the specification may also lack parsimony if many independent variables are considered, resulting in inefficient coefficient estimates.

Interpretation problems also arise due to identification issues. Indeed, such variables might be incorporated into the thresholds to investigate heterogeneous reporting or to obtain better outcome selection probabilities. However, ordered choice models are commonly used because of the sound economic reasoning behind the assumption of one latent index.

In general, any ordered choice specification that involves a linear utility function, additive error, linear thresholds and an overlap in threshold and latent regression covariates will have poor interpretation properties due to the lack of identification between the thresholds and latent regression. This is owing to the standardisation of ordered choice models by the first (or other) threshold. When this standardisation is required, it takes the form of a downward shift in both the latent index, and each of the thresholds, by $x_i\gamma_1$ (the variation in the first threshold). Although this does not affect the estimation probabilities, because of

$$x_i\gamma_j - x_i\beta = x_i(\tilde{\gamma}_j - \gamma_1) - x_i(\tilde{\beta} - \gamma_1) = x_i\tilde{\gamma}_j - x_i\tilde{\beta},$$

where $\tilde{\beta}$ is the true effect of x_i on underlying health and $\tilde{\gamma}_j$ is the true effect of x_i on threshold j. This implies that the resulting estimate of β is actually a relative measurement of $\beta - \gamma_1$ (and likewise for the threshold coefficients). For the SAH context, this means that the directional effect of covariates on underlying health cannot be identified. Similarly, the directional effects of covariates on the thresholds cannot be identified.

In order to obtain directional effects of covariates on these model "components" (the latent regression and thresholds), an ordered choice model specification requires suitable assumptions about the effect of such variables on the underlying propensity or thresholds (Jones, 2007). Selection of threshold covariates should therefore be based on the criterion that these variables do not affect the latent regression and vice versa. An alternative assumption that could be made is that the effect of the shared covariates on the first threshold, γ_1, is equal to zero. If other thresholds are truly affected by a covariate, it seems unlikely that the first threshold would be unaffected.

Use of distinct threshold covariates allows the parameterisation of μ_1 with the same covariates as with the other thresholds (assuming no multi-collinearity issues). However, by default, standard econometric software estimation procedures tend not to include a parameterisation of the first threshold. A way to bypass this issue is to include threshold covariates in the latent regression. If this approach is used, then the parameters can be related to the fully parameterised threshold specification as follows (denoting the parameters with double dots as the newly estimated parameters):

$$y_i^* = \beta_0 + x_i^0\beta + z_i^0\ddot{\gamma}_1$$

$$\ddot{\mu}_{ij} = \ddot{\mu}_j + z_i^0\ddot{\gamma}_j, \quad \ddot{\mu}_{i1} = 0 \;\forall\, i, \tag{1}$$

where $\ddot{\gamma}_1 = -\gamma_1, \ddot{\gamma}_j = \gamma_j - \gamma_1$, and $\ddot{\mu}_j = \mu_j$. Thus, the true direction of partial effects on each of the thresholds would need to be worked out ex-post. A statistically significant estimate of $\ddot{\gamma}_j, j \neq 1$, would indicate cut-point shift.

4.2 Selection of threshold covariates

This section looks at which covariates might be used for the thresholds in the SAH context if we treat the threshold covariates as being completely distinct from the latent regression covariates. Before such variable selection is possible, it is important to identify what the thresholds represent. Depending on the application this might vary significantly. For example, in Cameron and Heckman (1998), which investigates the completion of schooling, the thresholds separating education categories are motivated as the marginal costs of attending different schooling levels. Pudney and Shields (2000) likewise motivate their application on the promotion of nurses, with thresholds indicating the relative waiting times required for promotion.

In the preceding examples, the choice of threshold covariates is quite clear. However, there is not always an obvious distinction between variables which should be included in the thresholds, and those which should be included in the latent index. Selection of suitable threshold covariates is made more difficult in the context of SAH, due to the large range of variables that could foreseeably be related to underlying health. We discuss the use of three different types of variables for use in the thresholds. In the HILDA dataset, options may include participant understanding of the English language, interviewer perceptions of participants, and personality scores. In practice, determination of these covariates will depend on the dataset used and the range of questions asked of the respondents.

Participant understanding of English may influence how a respondent perceives the SAH question and which words or components of the question a respondent focused on when answers (see, for example, Vaillant and Wolff (2010), who adjusted for the cross-cultural bias stemming from country of origin and language proficiency in a model for the SAH of older migrants throughout Europe). Further evidence for the validity of such variables is given by the array of literature discussing the wording of subjective questions with ordinal responses even for English-speaking respondents (see Hernandez-Quevedo et al., 2008, or Baron-Epel and Kaplan, 2001). There is a concern, however, that such variables may be correlated to factors such as ethnicity which, if not properly controlled for in the latent regression, could result in endogenous effects of these variables through the thresholds. Language-based variables are not preferred if alternatives exist.

Interview or interviewer related questions provide information such as whether or not other adults were present at the time of the survey, how much a participant co-operated throughout the survey, and whether or not a participant seemed to be suspicious of the survey. Each of these factors likely influences responses. For example, a lack of trust in an interviewer might result in an individual understating his or her health, especially if he or she is unemployed and feels a need to justify this (referred to as *justification bias* by Jones, Rice, and Roberts, 2010). Some of these variables, such as whether the respondent seemed suspicious, may not be particularly appropriate here, as they are based on third party subjective assessments, which may be measuring latent factors such as interviewer mood or bias. A variable that does seem appropriate, however, indicates the presence of other adults at the time of the survey. This is objective in nature, and not likely to be correlated to the underlying health of the individual, making it a good candidate for inclusion in the thresholds.

Personality variables are also considered as potentially good choices for threshold covariates. These variables, based on a series of indicator questions, are used to construct personality scales for each of the *big five* personality traits *openness to experience*, *conscientiousness*, *extraversion*, *agreeableness* and *neuroticism* (Losonscz, 2004). In particular, we include neuroticism and conscientiousness in the thresholds. Studies in psychology, including Jorm, Christensen, Hendersen, Korten, Mackinnon, and Scott (1993), Korotkov and Hannah (2004) and Michel (2006), provide evidence of personality traits (especially *neuroticism*, but also *conscientiousness*) influencing the perception of health, rather than underlying health. However, in the majority of papers, this is shown via the comparison of models for objective measures of health with models for subjective measures of health. For example, Korotkov and Hannah (2004) investigates the significance of the big five personality traits (and interactions) on objective behavioural measures, such as restricting activities, the number of days in bed, and the frequency of physician use. This analysis was compared to subjective measures of health, such as SAH, self-reporting of physical symptoms and positive and negative affect. The result of the Korotkov and Hannah (2004) study was that, while very few cases of significant personality traits were found for objective measures of health, personality variables (neuroticism especially) were more frequently found to be significant in explaining subjective measures of health. As differences in perceptions would reasonably be expected to influence subjective measures more than objective measures of health, measures for neuroticism and conscientiousness are used in the thresholds for our application (in addition to an indicator variable for the presence of other people). The derived variable available in the HILDA dataset measures emotional stability, which we use as thecomplement to neuroticism (Losonscz, 2004).

4.3 Non-linear Thresholds: The Hierarchical Ordered Probit

Another issue that the linearly heterogeneous threshold specifications face is lack of restrictions enforcing the ordering of the thresholds. The thresholds need to remain ordered in order for the probabilities to remain positive. However, it is clear that with linear thresholds, such an ordering is not imposed, and thresholds might cross if particular values of the threshold covariates are taken. This motivates what Greene (2007) calls the *Hierarchical Ordered Probit* (HOPIT) model. The model has two alternatives, as shown below.

$$\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma)} \qquad [\textit{Case 1}] \tag{2}$$

$$\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma_j)} \qquad [\textit{Case 2}] \tag{3}$$

with λ_j indicating constants for each threshold, j, and the normalising restriction that $\mu_1 = 0$.

It is important to acknowledge that covariates (and estimated threshold parameters) influence thresholds differently from previous specification of thresholds. While linear threshold specifications have focused on directly estimating the level of each of the thresholds, the covariates in the HOPIT specification instead estimate the spacing between each of the thresholds, for example, in *case 2*, this spacing is:

$$\mu_{ij} - \mu_{i,j-1} = e^{(\lambda_j + z_i^0 \gamma_j)}.$$

In both forms of the model, the exponential function constrains each of these differences to be non-negative.

As Greene and Hensher (2010) note, the first case can be re-written as

$$\mu_{ij} = \sum_{h=1}^{j} \left[e^{(\lambda_h)} \right] e^{(z_i^0 \gamma)},$$

meaning the covariates can be interpreted as scaling factors for the spaces between the thresholds. The directional effect of a covariate on the threshold is also clear from the direction of the coefficient.

In contrast, *case 2*, allows much more flexibility in the effect of covariates. With this flexibility comes more difficulty in interpretation of threshold coefficients. As threshold j can be re-written as:

$$\mu_{ij} = \sum_{h=1}^{j} e^{(\lambda_h + z_i^0 \gamma_h)}$$

the direction of the partial effect of a covariate on the level of threshold j depends on the derivative of the summation of several exponential components.

$$\frac{d\mu_{ij}}{dz_i} = \sum_{h=1}^{j} \gamma_h e^{(\lambda_h + z_i^0 \gamma_h)}.$$

That is, a weighted sum of the coefficients, where the weights are given by the threshold differences which are, in turn, non-linear functions of the covariates.

As per the Pudney and Shields (2000) specification, and likewise in King, Murray, Salomon, and Tandon (2004) full flexibility in this model can be achieved through the parameterisation of the first threshold. However, it is more straightforward to include these variables in the underlying regression, and then calculate the direction of the effects of these variables on the first threshold as the opposite of the estimated coefficients on the latent regression. One advantage of the functional form of the HOPIT specification is that, unlike the linear threshold models, no transformations on the threshold covariates are required if the first threshold is allowed to vary by threshold-specific covariates, as these covariates influence the differences between thresholds, rather than the level of the thresholds.

5 Variance Heterogeneity

Another area for heterogeneity is the variance of the error term, also referred to as utility function scaling. In our example, heterogeneity in this form may account for individuals having different extremes in their health responses. For example, it seems reasonable to have an expectation that individuals who are less emotionally stable (more neurotic) are more likely to choose extreme responses than individuals who are more stable. Likewise, it seems appropriate to expect that levels of health might be more extreme in individuals who are at particular life stages. That is, most young individuals may consider themselves quite healthy, while older individuals exhibit wide variation in health due to long term effects of lifestyle factors as well as genetic predispositions.

We can allow covariates to influence the variance of the error term through an exponential transformation (again ensuring positivity). Thus we specify the variance in the following form:

$$\sigma_i^2 = [e^{w_i^0\theta}]^2,$$

where a constant is omitted from w_i for identification. That is, the constant is omitted for the reason that σ was assumed to equal one at the outset.

Estimation of the probabilities, and thus the likelihood function, consequently requires standardisation of the latent utility and thresholds by the standard deviation,

$$Pr(y_i = j|x_i, w_i) = \Phi\left(\frac{\mu_{ij} - x_i\beta}{e^{(w_i^0\theta)}}\right) - \Phi\left(\frac{\mu_{i,j-1} - x_i\beta}{e^{(w_i^0\theta)}}\right).$$

As mentioned in Greene and Hensher (2010), not accounting for heteroskedastic variance in ordered choice models has more significant implications than for a linear regression due to the maximum likelihood methodology often used to estimate these models. In contrast to the potential loss of accurate inference in OLS, heteroskedastic misspecification in OP models results in inconsistent parameter estimates.

The literature varies with respect to choice of variance heterogeneity covariates, with some authors, such as Lemp, Kockelman, and Unnikrishnan (2011), Ritter and Vance (2011) and Litchfield, Reilly, and Veneziani (2010) placing many of the latent regression factors in the variance component, and observing *ex-post* which variables are significant. The reason often cited for this is that the form of the variance is unknown. While this may be an appropriate strategy for many contexts (especially if some variables are then removed using an iterative process to improve efficiency), for this application variables are motivated economically. As motivated earlier, *age* and *age*^{2} variables as well as the threshold covariates, *personality* and *other adults* are used as indicators of the extremeness in the underlying health of individuals.

6 Parameter Heterogeneity

If the effect of latent regression covariates on underlying health is likely to differ across individuals, even after other variables are controlled for, a Random Parameters Ordered Probit (RPOP) specification might be appropriate. This may be relevant for quite a few of the variables considered for self-assessed health. For example, the effect of education on underlying health might be different across individuals depending on what discipline a degree is in or different levels of intelligence.

Greene and Hensher (2010) suggests a maximum simulated likelihood approach to account for variation in the parameters across individuals. This specification assumes

$$\beta_i = \beta + e_i$$

where β is a vector of population means, and e_i is a multivariate stochastic disturbance term. The variance-covariance matrix of random components, Σ, is decomposed into a lower triangular matrix, L, with ones on the diagonal, as well as a diagonal matrix, D, with strictly positive elements (Greene and Hensher, 2010). Thus, while L determines covariances between the parameters, D acts as the scaling matrix for each of the random components,

$$\Sigma = LD^2L'.$$

The overall form of the parameters can then be written as follows:

$$\beta_i = \beta + LD\epsilon_i$$

where ϵ_i has a mean of zero, a variance independent of the random parameters and follows some multivariate distribution of the dimension of the number of random parameters.

Maximisation of the likelihood function requires the unobserved components to be integrated out, such that

$$LnL_i = Ln\sum_{j=1}^{J}\int_{\epsilon_i}\left[\Phi(\mu_j - x_i\beta - x_i(LD\epsilon_i)) - \Phi(\mu_{j-1} - x_i\beta - x_i(LD\epsilon_i))\right]^{I(y=j)} f(\epsilon_i).d\epsilon_i.$$

Estimation involves maximising the simulated log-likelihood function, LnL_s. R random draws are used to integrate out the random components, and ϵ_{ir} is the r^{th} draw for individual i from the multivariate distribution of the parameters.

$$LnL_s = \sum_{n=1}^{N} Ln\sum_{j=1}^{J}\frac{1}{R}\sum_{r=1}^{R}\left[\Phi(\mu_j - x_i\beta - x_i(LD\epsilon_{ir})) - \Phi(\mu_{j-1} - x_i\beta - x_i(LD\epsilon_{ir}))\right]^{I(y=j)}$$

Halton draws or other intelligent draw methodologies can be used to speed up the simulation (Train, 1999). Due to the computational complexity arising in estimation of these models, empirical applications usually only allow a small subset of parameters to be random.

7 Stochastic Components Models

While random parameters techniques are applicable to many different types of models, randomness can also be induced in areas more specific to ordered choice models. In particular we refer to randomness (unobserved heterogeneity) in the thresholds and in the variance. Although the literature does not seem to have reached agreement on the naming of such models, we believe *stochastic thresholds* and *stochastic variance* are appropriate for such specifications. Academic discussion concerning this area is quite rare compared to the deterministic extensions, but includes papers such as Greene (2010), and Eluru, Bhat, and Hensher (2008), as well as Cunha and Navarro (2007). A brief treatment is given in Greene and Hensher (2010). It is noted that these models are likely to be better identified with the use of panel data, although for our expositional purposes cross section data are used.

7.1 Stochastic Thresholds

Stochastic threshold models can be a particularly useful tool to allow for threshold heterogeneity in health applications because of the relatively small number of threshold covariates that seem appropriate to the application. That is, we know that there may be factors that are likely to influence health status reporting; however, the covariates that influence thresholds may be difficult to observe or, indeed, unobservable. For example, personality variables in the HILDA dataset are only observed intermittently and are likely to be measured with a degree of error because of the construction of variables taking place outside the time of the survey. If this is the case, the practitioner may be wise to treat the threshold variables as unobservable, and integrate them out using stochastic threshold models (Weterings, Harris, and Hollingsworth, 2012).

The inclusion of stochastic thresholds requires careful consideration about the form of the thresholds. For example, caution must be used if an additive stochastic component is used, as the stochastic components of the thresholds may cause thresholds to cross over. A HOPIT specification is used with the stochastic components of the thresholds included in the exponential function to ensure ordering in the thresholds. Using the *case 2* form of the HOPIT model, the thresholds in these applications are estimated as follows:

$$\mu^*_{ij} = \mu^*_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j + \eta_j\omega_{ij})}, \;\; j = 2 \ldots J-1,$$

where ω_{ij} is assumed to be normally distributed (although other distributional forms for the random components could be chosen, normality has desirable properties in estimation). The parameter, η_j, allows for scaling of the random components of the thresholds.

In cases where z_i contains components not in the latent regression, such a specification fails to allow for the full amount of explanatory power in the thresholds in the absence of a linear first threshold. Thus, if the covariates in z_i are distinct from x_i, it might be desirable to specify the first threshold as

$$\mu_{i1}^* = \lambda_1 + z_i^0\gamma_1 + \eta_1\omega_{i1},$$

where $\lambda_1 = 0$ if the latent regression includes a constant.

However, the lack of identification of this linear threshold from the latent regression means that any stochastic component could also account for random variation in the level of underlying health through a stochastic constant in the latent regression in the cross-section case. This is inconsequential for this study, where the main parameters of interest are β (and potentially γ or θ). However, it does mean that the scale parameter on the first threshold cannot be interpreted directly (i.e., in terms of either the threshold or the underlying regression).

As with estimation of the random parameters model, the stochastic components of the thresholds must be integrated out, to get the following individual log-likelihood.

$$LnL_i = Ln\sum_{j=1}^{J}\int_{\omega_1,\ldots,\omega_{J-1}}\left[\Phi(\mu_{ij}^* - x_i\beta) - \Phi(\mu_{i,j-1}^* - x_i\beta)\right]^{I(y=j)} f(\omega_1,\ldots\omega_{J-1})d\omega_1\ldots d\omega_{j-1}$$

where the μ_{ik}^*'s are functions of ω_{ih} and potentially z_i, as defined earlier. In order to estimate the model, the simulated likelihood function is maximised,

$$LnL_s = \sum_{n=1}^{N} Ln\frac{1}{R}\sum_{r=1}^{R}\sum_{j=1}^{J}\left[\Phi(\mu_{ijr}^* - x_i\beta) - \Phi(\mu_{i,j-1,r}^* - x_i\beta)\right]^{I(y=j)}$$

where μ_{ijr}^* is the r^{tj} simulated value of μ_{ij}^*. Interpretation of the stochastic component scale parameters depends on the HOPIT form of the thresholds. That is, even if the stochastic component scale parameter for a particular threshold is not significant, that threshold may still be considered to be stochastic according to the influence of the previous thresholds level on the level of that threshold.

7.2 Stochastic Variance

Under reasoning similar to that of stochastic threshold models, there may be unobserved factors that influence the variance of an individual's SAH response. Introduction of a stochastic component for the variance of an ordered choice model may resolve this issue. As with other stochastic components models, maximum simulated likelihood can be used to introduce this component. Along the lines of Greene and Hensher (2010), the variance is specified as

$$\sigma_i^2 = [e^{(w_i^0\theta + \eta_v v_i)}]^2$$

where η_v acts as a scale parameter for the stochastic component of the variance. In order to estimate the model, we once again need to integrate out the stochastic component

$$LnL_i = Ln\sum_{j=1}^{J}\int_{v_i}\left[\Phi\left(\frac{\mu_j - x_i\beta}{e^{(w_i^0\theta+\eta_v v_i)}}\right) - \Phi\left(\frac{\mu_{j-1} - x_i\beta}{e^{(w_i^0\theta+\eta_v v_i)}}\right)\right]^{I(y=j)} f(v_i)dv_i$$

The model is again estimated using the simulated log-likelihood function as follows:

$$LnL_s = \sum_{n=1}^{N} Ln \sum_{j=1}^{J} \frac{1}{R} \sum_{r=1}^{R} \left[\Phi(\frac{\mu_j - x_i\beta}{e^{(w_i^0\theta + \eta_v v_{ir})}}) - \Phi(\frac{\mu_{j-1} - x_i\beta}{e^{(w_i^0\theta + \eta_v v_{ir})}}) \right]^{I(y=j)}.$$

As per the stochastic components of the thresholds, the stochastic component of the variance is assumed to be normally distributed in this application.

7.3 Stochastic Thresholds and Stochastic Variance

It may be the case that unobservable factors affect both the variance and the thresholds. If so, both unobserved components can be incorporated by maximising the simulated log-likelihood function

$$LnL_s = \sum_{n=1}^{N} Ln \sum_{j=1}^{J} \frac{1}{R} \sum_{r=1}^{R} \left[\Phi(\frac{\mu_{ijr}^* - x_i\beta}{e^{(w_i^0\theta + \eta_v v_{ir})}}) - \Phi(\frac{\mu_{i,j-1,r}^* - x_i\beta}{e^{(w_i^0\theta + \eta_v v_{ir})}}) \right]^{I(y=j)},$$

where

$$\mu_{i1}^* = \lambda_1 + z_i^0\gamma_1 + \eta_1\omega_{i,1,r}$$

$$\mu_{ij}^* = \mu_{i,j-1}^* + e^{(\lambda_j + z_i^0\gamma_j + \eta_j\omega_{i,j,r})}, \quad j = 2 \dots J-1,$$

and $\omega_{i,j,r}$ and v_{ir} are drawn from appropriate (in our case, normal) distributions. This is an extended case, similar to that estimated by Greene and Hensher (2010), but with the *case 2* form of the thresholds. Applications of ordered choice models of this complexity are rare (if present at all) in the literature.

8 Results

This section discusses the estimation results from each of the models, estimated using a single wave (year 2005) from the HILDA dataset. Although the focus of empirical research differs immensely among applications, we focus on three aspects - coefficient interpretation, overall model fit, and partial effects of variables on outcome probabilities at sample means. Coefficient interpretation is discussed in order to compare the effects of model choice (and thus economic frameworks) on how variables affect the different components of the model. For example, the direction of the effect of a variable on underlying health might change depending on the choice of framework for reporting heterogeneity. Overall model fit is assessed via the inspection of log-likelihood functions and, where appropriate, likelihood ratio tests. This gives some objective indication of the validity of model extensions and the existence of heterogeneity in general. Bayesian information criteria (BIC$=-2lnL + (lnN)k$) are also used to compare across non-nested models. Last, given the fact that interest exists for partial effects, rather than the coefficients, we investigate how partial effects of covariates change across models when different levels of heterogeneity are permitted.

8.1 The Standard Ordered Probit

As outlined earlier, the standard OP model implies quite strict assumptions about the structure of responses, with constant thresholds, constant variance, and uniform effects of covariates on every individual's underlying health. As seen in table 1, from the standard OP specification, it is estimated that education, household size, income and age have significant impacts on underlying health. Higher levels of education are shown to be associated with higher levels of underlying health, with the exception of *other tertiary* education, which has a point estimate close to that of a year 12 education. Large household size is estimated to have a negative effect on health, while income has the anticipated positive effect. Age appears to have a U shape effect on underlying health. The turning point of the age effect is 87 years indicating that, for the majority of the sample, health declines with age. None of the other covariates are considered significant at the 5 per cent significance level.

Insert table 1 about here

8.2 Linear Threshold Specifications ($x_i \equiv z_i$)

Relaxing the parallel regressions assumption appears to be merited with a model with a single source of threshold variation improving the log-likelihood by 19.3 points. With this specification using 14 additional parameters, the threshold parameters enjoy joint significance when assessed via the use of a likelihood ratio test. This suggests that the restrictive economic framework implied by the standard OP model may be inappropriate for this application. However, the large number of parameters estimated, compared to only two coefficients (for married and postgraduate education) as individually significant, suggests that the specification is quite inefficient. Although many of the latent regression covariates that were previously significant remain so, there is a large reduction in $z-$statistics for each of the coefficients. A specification along the lines of Boes and Winklemann (2006) (table 2) exacerbates this issue, with only 6 out of 42 additional parameters significant in this application. While tests for significant cut-point shift could be performed using the table 2 model in order to find which covariates exhibit index shift, this would not help to determine directional effect of covariates on underlying health, as the joint variation in the thresholds would not be identified from the latent health index.

Insert table 2 about here

8.3 Increased heterogeneity and partial effects

Partial effects of covariates on response outcomes are another subject of interest in many studies. We calculate the partial effect of Ln(income) for the models estimated thus far. For our purposes we calculate the partial effects at sample average values using numerical methods, and calculate standard errors using the delta method.

Insert figure 1 about here

Comparison of partial effects shows that model choice has significant implications when considering these effects. In particular, there are large changes in the partial effects in the transition from the Terza to the Boes and Winklemann model. In the case of the Good outcome category, the effect of income is halved, while there is a 45 per cent jump in the effect on Fair health. Full tables of partial effects are presented in appendix B.

8.4 Linear Threshold Specifications ($x_i \neq z_i$)

Seeking to avoid the interpretation issues related to the previous specifications, models with personality and *other adults* covariates in the thresholds are estimated. These variables are intended to account for threshold variation, as motivated in Section 4.2. Referring to Table 3, specification of the single varying-threshold model yields similar latent regression coefficient direction and significance to the standard OP model, albeit with *household size* losing significance. However, the improvement in the log-likelihood by 265.2 points reflects the great amount of explanatory power of the personality variables on individuals reporting of their health. In addition, the BIC indicates that these threshold variables help model fit to a far greater extent than the latent regression variables in thresholds. Interestingly, conscientious and emotionally stable (as opposed to neurotic) individuals are estimated to be more likely to rank themselves in higher categories of health than other adults when reporting their level of health. In contrast, the presence of other adults at the time of the survey is not estimated to have an effect on reporting in this specification.

Extending to the multiple threshold-varying model (the Pudney and Shields (2000) framework with only three thresholds parameterised) it is seen that most of the threshold covariate coefficients are significant, and the direction of the effects is largely consistent across the thresholds, which suggests the potential feasibility of a common threshold effect (index shift). The inclusion of the threshold covariates in the latent regression index (equivalent to incorporating variation within the first threshold under the assumption the covariates do not affect underlying health) results in many of the threshold coefficients losing significance. However, this may simply be the result of the *index shift* effect of covariates included in the latent regression. That is, each of the other threshold covariate coefficients can be interpreted as deviations from this index shift. Overall effects of each of the threshold covariates on each of the thresholds can be worked out using the results outlined in equation 1 and inference made in conjunction with appropriate transformations of the variance-covariance matrix of parameters.

Interestingly, there are only subtle changes in the latent regression coefficient estimates for the four varying-threshold model when compared to that of the standard OP model. Similar to the single varying threshold specification, household size loses significance. In addition, the turning point for the effect of age moves outside the sample ranges to 105 years. Overall, the results indicate the personality variables are congruent with our aim of finding factors that influence reporting but not underlying health. This therefore provides a strong motivation for the use of these variables in thresholds in future studies on SAH.

Insert table 3 about here

8.5 HOPIT specifications

The case 1 HOPIT (threshold scaling) model gives results similar to the *Terza* specification, as shown in Table 4. As with the *Terza* specification, *education*, *income* and *age* are the variables found to have a statistically significant effect on underlying health. In addition, the thresholds for relatively conscientious and emotionally stable individuals are estimated to be scaled down (i.e., closer to the first threshold), compared to individuals who are less conscientious and less emotionally stable. Taking into account the nesting of the standard OP model within the HOPIT models (when no threshold covariates are assumed), the log-likelihoods of this and the other models can be compared. The log-likelihood function is 212.9 points better than the standard OP model, indicating a substantial amount of additional heterogeneity has been accounted for, and clearly rejecting the null hypothesis of no significance of the threshold covariates when a likelihood ratio test is used.

The case 2 HOPIT model accounts for additional variation, improving the log-likelihood by 278.2 points from the standard OP model if only three thresholds are assumed to vary with covariates, and by 335.4 points if all thresholds are allowed to vary. The BIC values are also significantly lower than the standard OP model. Although the linear specifications would be preferred in this case, according to the BIC, the linear thresholds and HOPIT models are reasonably close in fit. The HOPIT form for the thresholds does, however, provide benefits by enforcing the ordering of the thresholds. In contrast to the linear thresholds model, rather than undertaking extensive post-estimation calculations to determine the estimated effect of each of the threshold covariates on each of the thresholds, direction and significance of these effects are much more straightforward with the four threshold HOPIT model. For the first threshold, the correct direction of these covariates results by simply taking the negative of the estimated coefficients. For example, emotional stability has a negative effect on the first threshold, resulting in more emotionally stable individuals being less likely to report poor health.

Insert table 4 about here

8.6 Specifications with Heteroskedasticity

Inspection of Table 5 shows heteroskedasticity in the scale of the underlying health variable is not supported with the selected covariates when non-varying thresholds are considered. However, when accounting for underlying health variance in the *case 2* heteroskedastic (HSK) HOPIT model, it is found that emotional stability and the presence of other adults does have an effect on the variance. Emotionally less stable individuals are estimated to give more extreme responses, and the presence of other adults appears to decrease the extremeness of health status. Note that a scaling of the utility function is not identifiable from a scaling of thresholds; however, these results might also be seen as implying shared variation in the thresholds. That is, it is not clear whether or not this result is indicative of underlying health variance or of threshold scaling. Jointly, the variance covariates are found to be significant, with the model exhibiting a likelihood ratio statistic of 15.6 points higher when compared to the four threshold HOPIT model ($LR_{crit} = 11.07$). Referring to information criteria yields a different conclusion, as the BIC indicates this specification is inferior to the less flexible linear threshold and non-HSK HOPIT models with four varying thresholds.

Controlling for these factors in the variance influences the significance and direction of many threshold covariates. For example, coefficients for μ_1 are found to be significant, whereas in the non-HSK model they were not. These results have implications for the overall effect of the covariates on the thresholds. Overall, nine threshold coefficients are found to be significant when heteroskedasticity is accounted for, compared to only six when it is not.

Although none of the latent regression coefficients has moved from being significant to being insignificant, the degree of significance has changed for many of the coefficients. For example, the $z-$statistics on the coefficients for *ln(income)* and *age* have decreased. In addition, the $z-$statistics of many previously insignificant variables have improved − including that of *separated* and *children 10−14*, which are now significant at the 10 per cent significance level.

Insert table 5 about here

8.7 Threshold and variance heterogeneity and partial effects

The partial effects of threshold and variance covariates are also considered. Once again, we use numerical methods to derive these effects. We compare the partial effects of emotional stability (a threshold variable that also appears in the variance function) across models.

Insert figure 2 about here

The choice of threshold structure makes a difference to partial effect estimation when the single varying threshold or the case 1 HOPIT model is considered. However, there is little difference in partial effects between specifications where every threshold is allowed to vary, regardless of the choice of functional form. In addition, although emotional stability is statistically significant in the variance function, its inclusion makes very little difference to partial effect estimation.

8.8 Stochastic Component specifications

The results from stochastic threshold (ST), stochastic variance (SV) and stochastic threshold and variance (STaV) models are interpreted from Table 6. For these models, covariates were removed from the thresholds and variance components, with the expectation that any variation would be integrated out under correct parametric assumptions. While noting the result by Train (1999) that 100 Halton draws can outperform 1000 random draws in the estimation of *Mixed Logit* models, a more conservative approach is used, utilising 200 Halton draws for each of the stochastic component models.

The ST model results in significance in the third and fourth stochastic threshold components, indicating that accounting for unobserved heterogeneity in this manner is justified. Corresponding improvement in the log-likelihood is expected, although this improvement is quite small compared to gains made by the use of personality and *other individual* variables. Interestingly, the significance for some previously significant latent regression coefficients has diminished (for example, *age* and *ln(income)*), indicating that stochastic thresholds are measuring factors similar to those measured by the personality variables. There are also small changes in the implied effects of these variables. For example, the turning point for the effect of age is at 82 years, compared to 87 years with the standard OP model.

Estimation of an SV model suggests there are factors that affect the extremeness of reported health levels. This is in contrast to the fully specified heteroskedastic OP model where, although the thresholds were not allowed to vary across individuals, personality and age variables were not found to be significant in the variance component.

Results from the STaV model reinforce the case for stochastic components when compared to the ST model. All stochastic components are significant, except for the first threshold/stochastic component. Thus, as with the HOPIT models estimated earlier, accounting for the variance and the thresholds simultaneously appears to uncover heterogeneity issues not accounted for using either of the components individually. The latent regression parameters exhibit similar direction and significance to previous estimations.

Although these specifications result in significant coefficient estimates, the BIC suggests that the use of actual covariates provides a better fit in this case. In fact, the variation accounted for by the stochastic thresholds and variance model does not outweigh the increase in model complexity from the standard ordered probit model, according to the BIC. As stated earlier, however, with the use of panel data, it is probable that the relative performance of these models would improve.

Insert table 6 about here

9 Conclusions

This paper investigated both the economic and the empirical implications of various extensions to the standard OP model as used to incorporate additional heterogeneity. Several conclusions have been drawn from the thorough discussion of the implications of each of the extensions. First, the discussion of the economic implications of the threshold extensions built a case for threshold covariates and latent regression covariates that are distinct from each other in order to help determine the effect of each of the variables on their component. Selection of such variables might prove to be problematic in practice because of the latency of these variables in some applications.

In this application the use of personality variables and interviewer perception questions was motivated by empirical evidence of the effect of these variables on reporting, but not underlying health. These variables appeared to be compatible with the aim of influencing reporting behaviour, rather than underlying health, as the inclusion of these covariates greatly improved the log-likelihood but did not dramatically influence the latent regression coefficients. In this context, tests for such assumptions are not possible because of the lack of identification between the thresholds and latent regression. A comparison of the linear and HOPIT forms of the thresholds indicated that the linear threshold specification fits these data better than the HOPIT form. Overall coefficient estimates and partial effects were very close, which suggests that the functional form of the thresholds had little effect on estimation. The latter specification provides the advantage of constraining the ordering of the thresholds.

An alternative was sought in treating the threshold covariates as unobserved and integrating out the threshold variation using maximum simulated likelihood techniques. Treating the threshold covariates as unobserved appeared to help account for heterogeneity, with reasonable levels of significance among latent regression coefficients. Similar improvements in the likelihood function to that made by the inclusion of personality covariates were not observed in this application. Comparison of BIC statistics across models reinforced this conclusion.

This review also looked at variance heterogeneity, which implies differences in the extremeness of respondent health. In the constant threshold case, deterministic variance heterogeneity was not precisely estimated. However, when considered in addition to personality threshold covariates, less neurotic individuals, as well as individuals completing the survey in the presence of other adults, were found to report less extreme levels of health. Treating the variance heterogeneity as unobserved meant that, even without individually varying thresholds, variance heterogeneity was seen to be important. This suggests that there may be factors other than those included in the previous models that influence extremeness in health status. Significance in variance heterogeneity was strengthened when stochastic thresholds were also allowed, reinforcing the results yielded when observable covariates were utilised.

References

O Baron-Epel and G Kaplan. General subjective health status or age-related subjective health status: does it make a difference? *Social Science and Medicine*, 53:1373–1381, 2001.

Y Benyamini, E Leventhal, and H Leventhal. Self-assessment of health: what do people know that predicts their mortality? *Research on Aging*, 21:477–500, 1999.

S Boes and R Winklemann. Ordered response models. *Allgemeines Statistisches Archiv*, 90 (2):165–180, 2006.

S Cameron and J Heckman. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *The Journal of Political Economy*, 106(2): 262–333, 1998.

J Carro and A Traferri. Correcting the bias in the estimation of dynamic ordered probit with fixed effects of self-assessed health status. Unpublished Mimeo, 2010.

F Cunha and S Navarro. The identification and economic content of ordered choice models with stochastic thresholds. University College Dublin, Gery Institute, Discussion Paper WP/26/2007, 2007.

N Eluru, C Bhat, and D Hensher. A mixed generalized ordered response models for examining pedestrian and bicyclist injury severity levels in traffic crashes. *Accident Analysis and Prevention*, 40(3):1033–1054, 2008.

W Greene. *Limdep Version 9.0: Reference Guide*. Plainview, New York, Econometric Software Inc., 2007.

W Greene. Ordered choices and heterogeneity in attribute processing. *Journal of Transport Economics and Policy*, 44(3):331–364, 2010.

W Greene and D Hensher. *Modeling Ordered Choices*. Cambridge University Press, 2010.

C Hernandez-Quevedo, A Jones, and N Rice. Reporting bias and heterogeneity in self-assessed health. evidence from the british household panel survey. *Cuadernos Econmicos de ICE*, 75:63–97, 2008.

E Idler and Y Benyamini. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behaviour*, 38(1):21–37, 1997.

A Jones. Panel data methods and applications to health economics. HEDG Working Paper 07/18, 2007.

A Jones and J Wildman. Health, income and relative deprivation: Evidence from the bhps. *Journal of Health Economics*, 27:308–324, 2008.

A Jones, N Rice, and J Roberts. Sick of work or too sick to work? evidence on self-reported health shocks and early retirement from the bhps. *Economic Modelling*, 27:866–880, 2010.

A Jorm, H Christensen, S Hendersen, A Korten, A Mackinnon, and R Scott. Neuroticism and self-reported health in an elderly community sample. *Personality and Individual Differences*, 15(5):515–521, 1993.

G King, C Murray, J Salomon, and A Tandon. Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98: 191–207, 2004.

D Korotkov and E Hannah. The five-factor model of personality: strengths and limitations in predicting health status, sick-role and illness behaviour. *Personality and Individual Differences*, 36:187–199, 2004.

J Lemp, K Kockelman, and A Unnikrishnan. Analysis of large truck crash severity using heteroskedastic ordered probit models. *Accident Analysis and Prevention*, 43:370–380, 2011.

M Lindeboom and E Doorslaer. Cut point shift and index shift in self-reported health. Ecuity III Project Working Paper number 2, 2004.

J Litchfield, B Reilly, and M Veneziani. An analysis of life satisfaction in albania: A heteroskedastic ordered probit model approach. Accepted for publication, 2010.

I Losonscz. Personality traits in hilda. HILDA survey research conference 2007, University of Melbourne, 2004.

R McKelvey and W Zavoina. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 4:106–120, 1975.

G Michel. The influence of neuroticism on concurrent symptom reporting; a multilevel modeling approach. *Personality and Individual Differences*, 41:549–560, 2006.

N Powdthavee. Ill-health as a household norm: Evidence from other peoples health problems. *Social Science and Medicine*, 68:251–259, 2009.

M Pudney and M Shields. Gender, race, pay and promotion in the british nursing profession: Estimation of a generalized ordered probit model. *Journal of Applied Econometrics*, 15: 367–399, 2000.

N Ritter and C Vance. The determinants of bicycle helmet use: Evidence from germany. *Accident Analysis and Prevention*, 43:95–100, 2011.

U Schneider, C Pfarr, B Schneider, and V Ulrich. I feel good! gender differences and reporting heterogeneity in self-assessed health. *European Journal of Health Economics*, 2011.

J Terza. Ordered probit: A generalization. *Communications in Statistics A. Theory and Methods*, 14:1–11, 1985.

K Train. Halton sequences for mixed logit. unpublished paper, 1999.

N Vaillant and F Wolff. Origin differences in self-reported health among older migrants living in france. *Public Health*, 124:90–98, 2010.

T Weterings, M Harris, and B Hollingsworth. Extending unobserved heterogeneity - a strategy for dealing with survey respondent perceptions in the absence of suitable data. Working Paper No. 13/12, http://www.buseco.monash.edu.au/ebs/pubs/wpapers/2012, 2012.

Figures

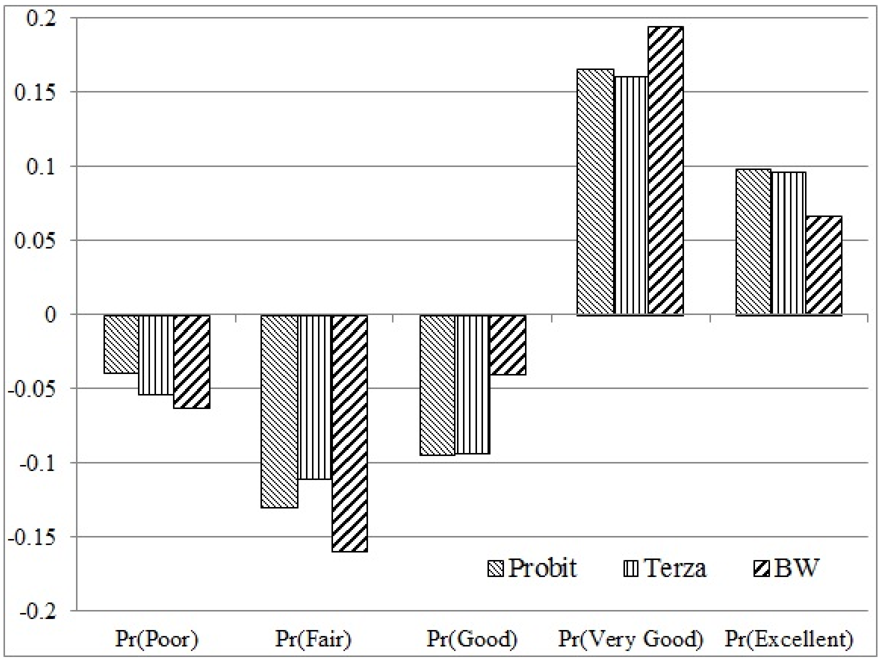

Figure 1: Comparison of partial effects of ln(income) across models

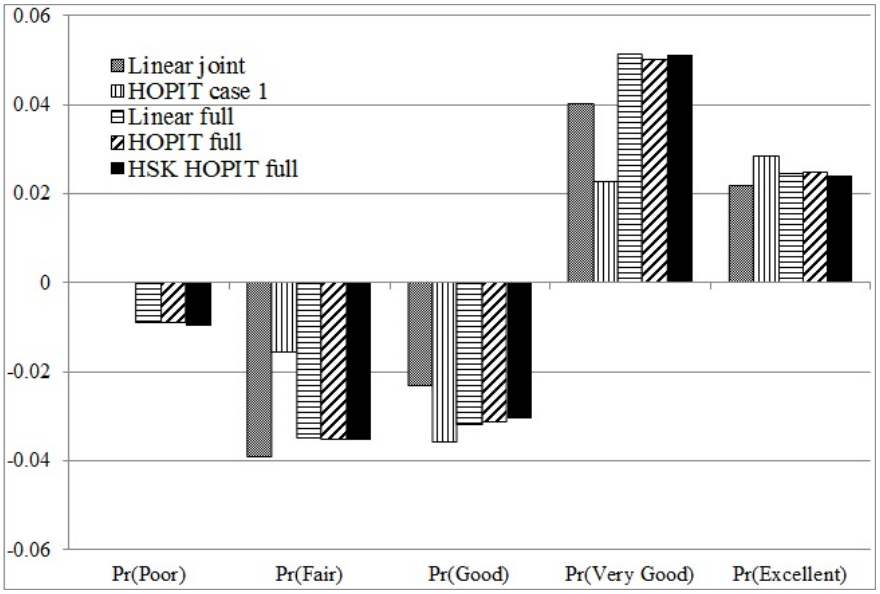

Figure 2: Comparison of partial effects of emotional stability across models

Tables

Table 1: Estimation results: Ordered Probit and single varying threshold specifications

Covariate	Standard OP	Model a[1] Regression	Model a[1] Thresholds
Constant	2.589**	2.516**	
	(0.078)	*(0.202)*	
Married	0.015	-0.165	-0.190*
	(0.033)	*(0.084)*	*(0.083)*
Separated	0.000	-0.095	-0.097
	(0.043)	*(0.094)*	*(0.092)*
Immigrant	-0.00	-0.091	-0.096
	(0.025)	*(0.056)*	*(0.054)*
Postgrad	0.434**	0.074	-0.372**
	(0.061)	*(0.148)*	*(0.141)*
Bachelors	0.305**	0.325**	0.025
	(0.035)	*(0.096)*	*(0.094)*
Other Tertiary	0.130**	0.157**	0.031
	(0.025)	*(0.056)*	*(0.054)*
Year 12	0.164**	0.201*	0.040
	(0.033)	*(0.086)*	*(0.084)*
Household Size	-0.025*	-0.021	0.003
	(0.012)	*(0.030)*	*(0.030)*
Children 0-4	0.028	0.108	0.090
	(0.035)	*(0.098)*	*(0.096)*
Children 5-9	0.019	-0.004	-0.020
	(0.034)	*(0.085)*	*(0.083)*
Children 10-14	0.050	0.048	0.001
	(0.034)	*(0.085)*	*(0.083)*
Ln(Income)	0.670**	0.835**	0.181
	(0.042)	*(0.120)*	*(0.118)*
Age	-3.217**	-3.477**	-0.433
	(0.333)	*(0.803)*	*(0.791)*
Age2	1.846**	2.740**	1.146
	(0.333)	*(0.742)*	*(0.733)*
μ_2	1.003**		0.948**
	(0.024)		*(0.199)*
μ_3	2.124**		2.069**
	(0.026)		*(0.199)*
μ_4	3.374**		3.319**
	(0.030)		*(0.199)*
Log-likelihood	-14399.0	-14379.7	
BIC	28965.8	29057.7	

** *Significant at 1% significance level* * *Significant at 5% significance level*

Standard errors in parentheses

[1]Model a: $\mu_{i1} = 0, \mu_{ij} = \mu_j + x_i^0\gamma, \;\; j = 2 \ldots J-1$

Table 2: Estimation results: Full Varying Threshold Specification - Threshold covariates = x_i

Variable	Regression	μ_2	μ_3	μ_4
Constant	2.508**	0.950**	2.099**	3.206**
	(0.210)	*(0.201)*	*(0.217)*	*(0.237)*
Married	-0.088	-0.237**	-0.102	-0.029
	(0.088)	*(0.084)*	*(0.091)*	*(0.100)*
Separated	-0.091	-0.092	-0.129	-0.168
	(0.098)	*(0.095)*	*(0.104)*	*(0.121)*
Immigrant	-0.113	-0.082	-0.083	-0.232**
	(0.058)	*(0.055)*	*(0.061)*	*(0.070)*
Postgrad	0.150	-0.394**	-0.256	-0.288
	(0.161)	*(0.138)*	*(0.166)*	*(0.179)*
Bachelors	0.303**	0.038	-0.022	0.046
	(0.100)	*(0.096)*	*(0.104)*	*(0.112)*
Other Tertiary	0.184**	0.013	0.054	0.160*
	(0.058)	*(0.055)*	*(0.061)*	*(0.072)*
Year 12	0.220*	0.025	0.041	0.158
	(0.089)	*(0.085)*	*(0.092)*	*(0.101)*
Household Size	-0.049	0.018	-0.010	-0.071*
	(0.032)	*(0.031)*	*(0.034)*	*(0.036)*
Children 0-4	0.156	0.069	0.137	0.165
	(0.102)	*(0.097)*	*(0.105)*	*(0.114)*
Children 5-9	0.020	-0.031	-0.051	0.118
	(0.090)	*(0.085)*	*(0.094)*	*(0.103)*
Children 10-14	0.097	-0.021	0.028	0.157
	(0.090)	*(0.085)*	*(0.093)*	*(0.102)*
Ln(Income)	0.986**	0.092	0.322*	0.536**
	(0.123)	*(0.118)*	*(0.128)*	*(0.138)*
Age	-3.689**	-0.274	-0.862	-0.618
	(0.831)	*(0.804)*	*(0.880)*	*(1.007)*
Age^2	2.922**	1.014	1.441	1.323
	(0.767)	*(0.749)*	*(0.827)*	*(0.986)*
Log-likelihood	-14318.9			
BIC	29197.0			

** *Significant at 1% significance level*   * *Significant at 5% significance level*

Standard errors in parentheses

Threshold form: $\mu_{ij} = \mu_j + x_i^0\gamma_j, \quad j = 2 \ldots J-1$

Table 3: Linear Threshold Specifications

Latent Regression	Model 1	Model 2	Model 3	Threshold Variables	Model 1	Model 2	Model 3
Constant	2.854**	2.853**	1.684**	**Conscientiousness †**	-0.119**		0.070**
	(0.079)	*(0.079)*	*(0.158)*		*(0.010)*		*(0.024)*
Married	-0.009	-0.010	-0.002	**Emotional Stability †**	-0.158**		0.175**
	(0.034)	*(0.034)*	*(0.034)*		*(0.010)*		*(0.023)*
Separated	-0.044	-0.046	-0.062	**Other Adults †**	0.036**		-0.171**
	(0.044)	*(0.043)*	*(0.044)*		*(0.022)*		*(0.051)*
Immigrant	0.005	0.006	0.008	μ_2	2.447**	2.211**	1.322**
	(0.025)	*(0.025)*	*(0.025)*		*(0.071)*	*(0.089)*	*(0.135)*
Postgrad	0.410**	0.410**	0.402**	**Conscientiousness**		-0.096**	-0.045
	(0.062)	*(0.062)*	*(0.062)*			*(0.014)*	*(0.023)*
Bachelors	0.275**	0.274**	0.264**	**Emotional Stability**		-0.137**	-0.005
	(0.035)	*(0.035)*	*(0.035)*			*(0.013)*	*(0.022)*
Other Tertiary	0.109**	0.108**	0.102**	**Other Individuals**		0.026	-0.102*
	(0.025)	*(0.025)*	*(0.025)*			*(0.029)*	*(0.049)*
Year 12	0.144**	0.143**	0.137**	μ_3	3.595**	3.826**	2.758**
	(0.033)	*(0.033)*	*(0.033)*		*(0.073)*	*(0.086)*	*(0.151)*
Household Size	-0.019	-0.019	-0.016	**Conscientiousness**		-0.151**	-0.089**
	(0.012)	*(0.012)*	*(0.012)*			*(0.013)*	*(0.026)*
Children 0-4	0.009	0.008	0.001	**Emotional Stability**		-0.176**	-0.017
	(0.035)	*(0.036)*	*(0.035)*			*(0.012)*	*(0.024)*
Children 5-9	0.027	0.028	0.026	**Other Individuals**		0.074**	-0.081
	(0.034)	*(0.034)*	*(0.034)*			*(0.026)*	*(0.054)*
Children 10-14	0.052	0.049	0.048	μ_4	4.879**	4.871**	3.751**
	(0.034)	*(0.034)*	*(0.034)*		*(0.076)*	*(0.116)*	*(0.175)*
Ln(Income)	0.622**	0.620**	0.601**	**Conscientiousness**		-0.107**	-0.042
	(0.043)	*(0.043)*	*(0.043)*			*(0.017)*	*(0.029)*
Age	-3.572**	-3.618**	-3.562**	**Emotional Stability**		-0.167**	0.000
	(0.338)	*(0.338)*	*(0.339)*			*(0.017)*	*(0.027)*
Age^2	1.800**	1.842**	1.696**	**Other Individuals**		-0.016	-0.178**
	(0.337)	*(0.337)*	*(0.339)*			*(0.036)*	*(0.061)*
Log-likelihood	-14133.8	-14113.9	-14062.1				
BIC	28463.3	28479.5	28403.8				

† *Coefficients for all thresholds in model 1, and specified in latent regression for model 3*

** *Significant at 1% significance level*   * *Significant at 5% significance level*   *(Standard errors in parentheses)*

Model 1: Threshold form: constant μ_{i1} and $\mu_{ij} = \mu_j + z_i^0\gamma, \quad j = 2 \ldots J-1$

Model 2: Threshold form: constant μ_{i1} and $\mu_{ij} = \mu_j + z_i^0\gamma_j, \quad j = 2 \ldots J-1$

Model 3: Threshold form: $\mu_{ij} = \mu_j + z_i^0\gamma_j, \quad j = 1 \ldots J-1$

Table 4: Hierarchical Ordered Probit (HOPIT) Specifications

Latent Regression Variables	Case 1	Case 2 $\mu_{i1} = 0$	Case 2 $\mu_{i1} = \mathbf{z_i^0}\gamma_1$	Threshold Variables	Case 1	case 2 $\mu_{i1} = 0$	Case 2 $\mu_{i1} = \mathbf{z_i^0}\gamma_1$
Constant	2.837**	2.849**	1.646**	**Conscientiousness** †	-0.045**		0.078**
	(0.080)	*(0.079)*	*(0.153)*		*(0.004)*		*(0.023)*
Married	-0.008	-0.010	-0.002	**Emotional Stability** †	-0.060**		0.174**
	(0.034)	*(0.034)*	*(0.033)*		*(0.004)*		*(0.021)*
Separated	-0.035	-0.045	-0.062	**Other Adults** †	0.012		-0.162**
	(0.043)	*(0.043)*	*(0.043)*		*(0.009)*		*(0.052)*
Immigrant	0.006	0.006	0.007	μ_2	0.568**	1.076**	0.277*
	(0.025)	*(0.025)*	*(0.025)*		*(0.034)*	*(0.068)*	*(0.120)*
Postgrad	0.412**	0.408**	0.402**	**Conscientiousness**		-0.090**	-0.036
	(0.062)	*(0.062)*	*(0.062)*			*(0.012)*	*(0.021)*
Bachelors	0.276**	0.272**	0.264**	**Emotional Stability**		-0.120**	-0.006
	(0.035)	*(0.035)*	*(0.035)*			*(0.011)*	*(0.019)*
Other Tertiary	0.112**	0.106**	0.102**	**Other Adults**		0.032	-0.091
	(0.025)	*(0.025)*	*(0.025)*			*(0.027)*	*(0.048)*
Year 12	0.145**	0.142**	0.138**	μ_3	0.664**	0.538**	0.362**
	(0.033)	*(0.033)*	*(0.033)*		*(0.029)*	*(0.071)*	*(0.076)*
Household Size	-0.020	-0.019	-0.016	**Conscientiousness**		-0.045**	-0.035**
	(0.012)	*(0.012)*	*(0.012)*			*(0.012)*	*(0.013)*
Children 0-4	0.009	0.007	0.001	**Emotional Stability**		-0.036**	-0.009
	(0.035)	*(0.036)*	*(0.035)*			*(0.011)*	*(0.012)*
Children 5-9	0.026	0.027	0.025	**Other Adults**		0.039	0.019
	(0.034)	*(0.034)*	*(0.034)*			*(0.027)*	*(0.028)*
Children 10-14	0.050	0.048	0.048	μ_4	0.787**	0.119	0.059
	(0.034)	*(0.034)*	*(0.034)*		*(0.031)*	*(0.083)*	*(0.084)*
Ln(Income)	0.633**	0.622**	0.601**	**Conscientiousness**		0.029*	0.033*
	(0.043)	*(0.043)*	*(0.043)*			*(0.013)*	*(0.013)*
Age	-3.691**	-3.644**	-3.562**	**Emotional Stability**		0.001	0.010
	(0.338)	*(0.338)*	*(0.338)*			*(0.013)*	*(0.013)*
Age2	2.021**	1.898**	1.697**	**Other Adults**		-0.072**	-0.078**
	(0.336)	*(0.337)*	*(0.339)*			*(0.029)*	*(0.029)*
Log-likelihood	-14186.1	-14120.8	-14063.6				
BIC	28567.9	28493.3	28406.8				

† *Scaling coefficients for all thresholds in HOPIT model Case 1, γ_1 - included in latent regression for Case 2*

** *Significant at 1% significance level*   * *Significant at 5% significance level (Standard errors in parentheses)*

Case 1: Threshold form: $\mu_{i1} = 0, \mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma)}$ $\quad j = 2 \ldots J-1$

Case 2: Threshold form: $\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma_j)}, \quad j = 2 \ldots J-1$

Table 5: Heteroskedastic Variance Specifications

Latent Regression Variables	**HSK Probit**	**HSK HOPIT**	**Threshold Variables**	**HSK Probit**	**HSK HOPIT**	**Variance Variables**	**HSK Probit**	**HSK HOPIT**
Constant	2.429** *(0.167)*	1.602** *(0.144)*	**Conscientiousness**		0.097* *(0.040)*	**Age**	-0.208 *(0.224)*	-0.130 *(0.225)*
Married	0.016 *(0.031)*	0.000 *(0.027)*	**Emotional Stability**		0.064** *(0.034)*	**Age^2**	0.182 *(0.232)*	0.095 *(0.235)*
Separated	0.002 *(0.04)*	-0.066 *(0.037)*	**Other Adults**		-0.343** *(0.088)*	**Conscientiousness**	0.005 *(0.008)*	0.026 *(0.020)*
Immigrant	-0.004 *(0.024)*	0.000 *(0.020)*	μ_2	0.939** *(0.063)*	0.217 *(0.133)*	**Emotional Stability**	-0.01 *(0.008)*	-0.053** *(0.019)*
Postgrad	0.407** *(0.063)*	0.324** *(0.067)*	**Conscientiousness**		-0.015 *(0.026)*	**Other Adults**	0.03 *(0.017)*	-0.144** *(0.049)*
Bachelors	0.286** *(0.037)*	0.210** *(0.040)*	**Emotional Stability**		-0.051* *(0.024)*			
Other Tertiary	0.123** *(0.025)*	0.080** *(0.023)*	**Other Adults**		-0.229** *(0.066)*			
Year 12	0.154** *(0.032)*	0.116** *(0.030)*	μ_3	1.988** *(0.128)*	0.319** *(0.09)*			
Household Size	-0.023* *(0.011)*	-0.014 *(0.010)*	**Conscientiousness**		-0.014 *(0.022)*			
Children 0-4	0.025 *(0.033)*	0.004 *(0.028)*	**Emotional Stability**		-0.057** *(0.020)*			
Children 5-9	0.017 *(0.032)*	0.017 *(0.027)*	**Other Adults**		-0.110* *(0.052)*			
Children 10-14	0.044 *(0.032)*	0.046 *(0.028)*	μ_4	3.16** *(0.203)*	0.028 *(0.091)*			
Ln(inc)	0.626** *(0.056)*	0.474** *(0.075)*	**Conscientiousness**		0.054* *(0.022)*			
Age	-3.025** *(0.362)*	-2.874** *(0.447)*	**Emotional Stability**		-0.039 *(0.021)*			
Age^2	1.738** *(0.329)*	1.432** *(0.308)*	**Other Adults**		-0.206** *(0.053)*			
Log-likelihood	-14395.7	-14055.8						
BIC	29005.8	28437.8						

** *Significant at 1% significance level*   * *Significant at 5% significance level*   *(Standard errors in parentheses)*

HOPIT threshold form: $\mu_{i1} = z_i^0\gamma_1$ and $\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j)}$, $j = 2 \dots J-1$

HSK form: $\sigma_i = e^{w_i^0\theta}$

Table 6: Stochastic Component Specifications

Latent Regression Variables	ST Model	SV Model	STaV Model	Other Parameters	ST Model	SV Model	STaV Model
Constant	3.707**	2.624**	5.691**	μ_2	0.324	0.032	0.082
	(1.353)	*(0.084)*	*(1.084)*		*(0.366)*	*(0.034)*	*(0.104)*
Married	0.038	0.017	0.098	μ_3	0.336	0.116**	0.587**
	(0.055)	*(0.034)*	*(0.113)*		*(0.369)*	*(0.014)*	*(0.152)*
Separated	-0.003	0.001	0.069	μ_4	0.741*	0.229**	1.683**
	(0.068)	*(0.044)*	*(0.124)*		*(0.369)*	*(0.015)*	*(0.295)*
Immigrant	-0.015	-0.004	-0.013	η_1	0.933		0.057
	(0.039)	*(0.025)*	*(0.073)*		*(0.733)*		*(0.278)*
Postgrad	0.679**	0.438**	1.419**	η_2	0.011		0.887**
	(0.261)	*(0.062)*	*(0.425)*		*(0.095)*		*(0.107)*
Bachelors	0.461**	0.307**	0.876**	η_3	0.593**		1.430**
	(0.175)	*(0.035)*	*(0.214)*		*(0.093)*		*(0.169)*
Other Tertiary	0.200*	0.132**	0.316**	η_4	0.449**		1.276**
	(0.082)	*(0.026)*	*(0.084)*		*(0.127)*		*(0.224)*
Year 12	0.256*	0.167**	0.474**	η_v		0.116*	0.693**
	(0.106)	*(0.033)*	*(0.139)*			*(0.049)*	*(0.117)*
Household Size	-0.047	-0.026*	-0.120*				
	(0.026)	*(0.012)*	*(0.056)*				
Children 0-4	0.057	0.026	0.102				
	(0.059)	*(0.035)*	*(0.126)*				
Children 5-9	0.030	0.022	0.091				
	(0.054)	*(0.034)*	*(0.121)*				
Children 10-14	0.088	0.052	0.118				
	(0.062)	*(0.034)*	*(0.117)*				
Ln(Income)	1.096**	0.677**	2.502**				
	(0.402)	*(0.043)*	*(0.596)*				
Age	-5.262**	-3.238**	-12.418**				
	(1.960)	*(0.339)*	*(3.125)*				
Age^2	3.220*	1.848**	8.542**				
	(1.257)	*(0.338)*	*(2.35)*				
Log-likelihood	-14387.3	-14398.2	-14377.7				
BIC	28979.6	28973.5	28969.8				

** *Significant at 1% significance level* * *Significant at 5% significance level* *(Standard errors in parentheses)*

ST Model: $\mu^*_{ij} = \mu^*_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j + \eta_j\omega_{ij})}, \ \ j = 2 \dots J-1$

SV Model: $\sigma_i^2 = [e^{(w_i^0\theta + \eta_v v_i)}]^2$

STaV Model: $\mu^*_{ij} = \mu^*_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j + \eta_j\omega_{ij})}, \ \ j = 2 \dots J-1$ and $\sigma_i^2 = [e^{(w_i^0\theta + \eta_v v_i)}]^2$

A Definition of Variables

Table 7: Descriptive statistics and definitions for model covariates

Variable	Average	St. Dev	Min	Max	Description
SAH	3.361	0.954	1.000	5.000	Self-Assessed Health Variable
Married	0.629	0.483	0.000	1.000	1 if Married, 0 otherwise
Separated	0.139	0.346	0.000	1.000	1 if Separated, 0 otherwise
Immigrant	0.209	0.407	0.000	1.000	1 if Immigrant, 0 otherwise
Postgraduate	0.032	0.176	0.000	1.000	1 if highest level of education is Postgraduate, 0 otherwise
Bachelors	0.127	0.333	0.000	1.000	1 if highest level of education is a Bachelors degree , 0 otherwise
Other Tertiary	0.346	0.476	0.000	1.000	1 if highest level of education is another tertiary qualification, 0 otherwise
Year 12	0.144	0.351	0.000	1.000	1 if highest level of education is Year 12, 0 otherwise
Household Size	2.882	1.468	1.000	13.000	Number of individuals living in the household
Children 0-4	0.136	0.342	0.000	1.000	1 if have children between 0 and 4 years, 0 otherwise
Children 5-9	0.145	0.352	0.000	1.000	1 if have children between 5 and 9 years, 0 otherwise
Children 10-14	0.188	0.391	0.000	1.000	1 if have children between 10 and 14 years, 0 otherwise
Ln(Income)	0.780	0.656	0.000	8.176	Ln(Total Household Income+1)/100000
Age	0.440	0.180	0.150	0.930	Age of respondent (in 100 years)
Conscientiousness	5.084	1.038	1.000	7.000	Personality measure for conscientiousness
Emotional Stability	5.174	1.090	1.000	7.000	Personality measure for Emotional Stability
Other Adults	0.352	0.478	0.000	1.000	1 if other adults were present at time of survey, 0 otherwise

B Full Partial Effect Tables

Table B.1: Partial Effects: Standard Ordered Probit Model

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	-0.001	-0.003	-0.002	0.004	0.002
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Separated	0.000	0.000	0.000	0.000	0.000
	(0.003)	*(0.009)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	0.001	0.001	-0.001	-0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.025**	-0.084**	-0.061**	0.107**	0.063**
	(0.004)	*(0.012)*	*(0.009)*	*(0.015)*	*(0.009)*
Bachelors	-0.018**	-0.059**	-0.043**	0.075**	0.045**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.008**	-0.025**	-0.018**	0.032**	0.019**
	(0.002)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.010**	-0.032**	-0.023**	0.041**	0.024**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001*	0.005*	0.004*	-0.006*	-0.004*
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	-0.002	-0.005	-0.004	0.007	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.004	-0.003	0.005	0.003
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Children 10-14	-0.003	-0.010	-0.007	0.012	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Ln(Income)	-0.039**	-0.130**	-0.094**	0.166**	0.098**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.019**	0.062**	0.045**	-0.079**	-0.047**
	(0.001)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent
(Standard errors in parentheses)

Table B.2: Partial Effects: Single varying-threshold (Threshold covariates = z_i)

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.011*	-0.017	-0.004	0.006	0.004
	(0.005)	*(0.009)*	*(0.005)*	*(0.008)*	*(0.005)*
Separated	0.006	-0.007	0.000	0.001	0.000
	(0.006)	*(0.011)*	*(0.006)*	*(0.011)*	*(0.007)*
Immigrant	0.006	-0.007	-0.001	0.001	0.001
	(0.004)	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.005	-0.107**	-0.064**	0.110**	0.066**
	(0.01)	*(0.016)*	*(0.009)*	*(0.015)*	*(0.009)*
Bachelors	-0.021**	-0.054**	-0.043**	0.074**	0.044**
	(0.006)	*(0.01)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.010**	-0.022**	-0.018**	0.031**	0.019**
	(0.004)	*(0.006)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.013*	-0.027**	-0.023**	0.039**	0.024**
	(0.005)	*(0.009)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.005	0.003*	-0.006*	-0.004*
	(0.002)	*(0.003)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	-0.007	0.002	-0.003	0.005	0.003
	(0.006)	*(0.01)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	0.000	-0.004	-0.002	0.004	0.002
	(0.006)	*(0.009)*	*(0.005)*	*(0.008)*	*(0.005)*
Children 10-14	-0.003	-0.009	-0.007	0.012	0.007
	(0.005)	*(0.009)*	*(0.005)*	*(0.008)*	*(0.005)*
Ln(Income)	-0.054**	-0.111**	-0.093**	0.161**	0.097**
	(0.008)	*(0.012)*	*(0.007)*	*(0.011)*	*(0.007)*
Age	0.014**	0.069**	0.047**	-0.081**	-0.049**
	(0.003)	*(0.005)*	*(0.003)*	*(0.005)*	*(0.003)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Thresholds form: $\mu_{ij} = \mu_j + x_i^0\gamma, \;\; j = 2 \ldots J-1$

Table B.3: Partial Effects: Fully varying thresholds (Threshold covariates = z_i)

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.005	-0.043**	0.032*	0.014	-0.009
	(0.005)	*(0.012)*	*(0.016)*	*(0.015)*	*(0.008)*
Separated	0.006	-0.006	-0.015	0.003	0.012
	(0.006)	*(0.014)*	*(0.02)*	*(0.02)*	*(0.011)*
Immigrant	0.007	0.000	0.004	-0.030*	0.018**
	(0.004)	*(0.008)*	*(0.012)*	*(0.011)*	*(0.006)*
Postgrad	-0.009	-0.126**	-0.025	0.095**	0.065**
	(0.01)	*(0.024)*	*(0.033)*	*(0.027)*	*(0.013)*
Bachelors	-0.019**	-0.047**	-0.063**	0.090**	0.039**
	(0.006)	*(0.013)*	*(0.017)*	*(0.016)*	*(0.008)*
Other Tertiary	-0.012**	-0.031**	-0.009	0.048**	0.004
	(0.004)	*(0.008)*	*(0.012)*	*(0.012)*	*(0.007)*
Year 12	-0.014*	-0.034**	-0.022	0.061**	0.009
	(0.006)	*(0.012)*	*(0.016)*	*(0.015)*	*(0.008)*
Household Size	0.003	0.013**	-0.001	-0.019**	0.003
	(0.002)	*(0.004)*	*(0.006)*	*(0.005)*	*(0.003)*
Children 0-4	-0.010	-0.012	0.014	0.009	-0.001
	(0.006)	*(0.013)*	*(0.017)*	*(0.016)*	*(0.008)*
Children 5-9	-0.001	-0.011	-0.015	0.043**	-0.015
	(0.006)	*(0.012)*	*(0.016)*	*(0.015)*	*(0.008)*
Children 10-14	-0.006	-0.023	0.002	0.036*	-0.009
	(0.006)	*(0.012)*	*(0.016)*	*(0.015)*	*(0.008)*
Ln(Income)	-0.062**	-0.160**	-0.040	0.195**	0.067**
	(0.008)	*(0.016)*	*(0.022)*	*(0.02)*	*(0.01)*
Age	0.014**	0.072**	0.034**	-0.070**	-0.050**
	(0.003)	*(0.007)*	*(0.009)*	*(0.008)*	*(0.004)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: $\mu_{ij} = \mu_j + x_i^0\gamma_j, \quad j = 2 \ldots J-1$

Table B.4: Partial Effects: Single varying threshold (Threshold covariates = z_i)

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.000	0.002	0.001	-0.002	-0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.002	0.009	0.006	-0.011	-0.006
	(0.002)	*(0.009)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.001	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.022**	-0.080**	-0.060**	0.104**	0.057**
	(0.003)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.014**	-0.054**	-0.040**	0.070**	0.038**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.006**	-0.021**	-0.016**	0.028**	0.015**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.003)*
Year 12	-0.008**	-0.028**	-0.021**	0.037**	0.020**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.004	0.003	-0.005	-0.003
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	-0.002	-0.001	0.002	0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.005	-0.004	0.007	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.003	-0.010	-0.008	0.013	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.033**	-0.121**	-0.091**	0.158**	0.086**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.021**	0.078**	0.058**	-0.101**	-0.055**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	0.000	-0.029**	-0.017**	0.030**	0.017**
		(0.003)	*(0.002)*	*(0.003)*	*(0.001)*
Emotional Stability	0.000	-0.039**	-0.023**	0.040**	0.022**
		(0.002)	*(0.002)*	*(0.003)*	*(0.001)*
Other Adults	0.000	0.009	0.005	-0.009	-0.005
		(0.005)	*(0.003)*	*(0.005)*	*(0.003)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: constant μ_{i1} and $\mu_{ij} = \mu_j + z_i^0\gamma, \;\; j = 2 \ldots J-1$

Table B.5: Partial Effects: Three varying thresholds (Threshold covariates = z_i)

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.001	0.002	0.001	-0.003	-0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.003	0.009	0.007	-0.012	-0.006
	(0.002)	*(0.008)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.002	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.022**	-0.080**	-0.059**	0.104**	0.057**
	(0.004)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.015**	-0.053**	-0.040**	0.070**	0.038**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.006**	-0.021**	-0.016**	0.027**	0.015**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.003)*
Year 12	-0.008**	-0.028**	-0.021**	0.036**	0.020**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.004	0.003	-0.005	-0.003
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	-0.001	-0.001	0.002	0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.002	-0.005	-0.004	0.007	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.003	-0.010	-0.007	0.013	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.033**	-0.121**	-0.090**	0.158**	0.086**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.022**	0.078**	0.058**	-0.102**	-0.055**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	0.000	-0.024**	-0.035**	0.044**	0.015**
		(0.003)	*(0.005)*	*(0.005)*	*(0.002)*
Emotional Stability	0.000	-0.034**	-0.035**	0.046**	0.023**
		(0.003)	*(0.005)*	*(0.005)*	*(0.002)*
Other Adults	0.000	0.006	0.023*	-0.031**	0.002
		(0.007)	*(0.01)*	*(0.01)*	*(0.005)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: constant μ_{i1} and $\mu_{ij} = \mu_j + z_i^0\gamma_j, \quad j = 2 \ldots J-1$

Table B.6: Partial Effects: Fully varying thresholds (Threshold covariates = z_i)

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.003	0.012	0.009	-0.016	-0.009
	(0.002)	*(0.008)*	*(0.007)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.002	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.021**	-0.077**	-0.060**	0.102**	0.056**
	(0.003)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.014**	-0.051**	-0.040**	0.067**	0.037**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.005**	-0.020**	-0.015**	0.026**	0.014**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.007**	-0.026**	-0.021**	0.035**	0.019**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.003	0.002	-0.004	-0.002
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.005	-0.004	0.006	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.002	-0.009	-0.007	0.012	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.031**	-0.116**	-0.090**	0.153**	0.084**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.021**	0.080**	0.062**	-0.105**	-0.058**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	-0.004**	-0.024**	-0.034**	0.047**	0.015**
	(0.001)	*(0.003)*	*(0.005)*	*(0.005)*	*(0.002)*
Emotional Stability	-0.009**	-0.035**	-0.032**	0.051**	0.024**
	(0.001)	*(0.003)*	*(0.005)*	*(0.005)*	*(0.002)*
Other Adults	0.009**	0.008	0.019	-0.037**	0.001
	(0.003)	*(0.007)*	*(0.01)*	*(0.01)*	*(0.005)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: $\mu_{ij} = \mu_j + z_i^0 \gamma_j, \quad j = 1 \ldots J-1$

Table B.7: Partial Effects: HOPIT Case 1

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.000	0.002	0.001	-0.002	-0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.002	0.007	0.005	-0.009	-0.005
	(0.002)	*(0.009)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.002	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.003)*
Postgrad	-0.023**	-0.081**	-0.058**	0.105**	0.057**
	(0.004)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.015**	-0.054**	-0.039**	0.071**	0.038**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.006**	-0.022**	-0.016**	0.029**	0.015**
	(0.001)	*(0.005)*	*(0.004)*	*(0.007)*	*(0.003)*
Year 12	-0.008**	-0.029**	-0.021**	0.037**	0.020**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.004	0.003	-0.005	-0.003
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	-0.001	-0.002	-0.001	0.002	0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.005	-0.004	0.007	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.003	-0.010	-0.007	0.013	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.035**	-0.124**	-0.090**	0.162**	0.087**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.021**	0.075**	0.054**	-0.098**	-0.053**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.002)*
Conscientiousness	0.000	-0.012**	-0.027**	0.017**	0.021**
		(0.001)	*(0.003)*	*(0.002)*	*(0.002)*
Emotional Stability	0.000	-0.016**	-0.036**	0.023**	0.029**
		(0.001)	*(0.002)*	*(0.002)*	*(0.002)*
Other Adults	0.000	0.003	0.007	-0.004	-0.006
		(0.002)	*(0.005)*	*(0.003)*	*(0.004)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: $\mu_{i1} = 0, \mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma)} \quad j = 2 \ldots J - 1$

Table B.8: Partial Effects: HOPIT case 2, $\mu_{i1} = 0$

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.001	0.002	0.002	-0.003	-0.001
	(0.002)	*(0.006)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.002	0.009	0.007	-0.011	-0.006
	(0.002)	*(0.008)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.002	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.022**	-0.078**	-0.061**	0.103**	0.058**
	(0.004)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.015**	-0.052**	-0.041**	0.068**	0.039**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.006**	-0.020**	-0.016**	0.027**	0.015**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.008**	-0.027**	-0.021**	0.036**	0.020**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.004	0.003	-0.005	-0.003
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	-0.001	-0.001	0.002	0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.005	-0.004	0.007	0.004
	(0.002)	*(0.006)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.003	-0.009	-0.007	0.012	0.007
	(0.002)	*(0.006)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.034**	-0.118**	-0.093**	0.156**	0.089**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.022**	0.075**	0.059**	-0.099**	-0.056**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	0.000	-0.022**	-0.034**	0.041**	0.015**
		(0.003)	*(0.005)*	*(0.004)*	*(0.002)*
Emotional Stability	0.000	-0.030**	-0.034**	0.041**	0.023**
		(0.003)	*(0.004)*	*(0.004)*	*(0.002)*
Other Adults	0.000	0.008	0.022*	-0.032**	0.002
		(0.007)	*(0.01)*	*(0.01)*	*(0.005)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: $\mu_{i1} = 0, \mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0 \gamma_j)}, \quad j = 2 \dots J-1$

Table B.9: Partial Effects: HOPIT case 2, $\mu_{i1} = z_i^0\gamma_1$

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.003	0.012	0.009	-0.016	-0.009
	(0.002)	*(0.008)*	*(0.007)*	*(0.011)*	*(0.006)*
Immigrant	0.000	-0.001	-0.001	0.002	0.001
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.021**	-0.077**	-0.060**	0.102**	0.056**
	(0.003)	*(0.012)*	*(0.009)*	*(0.016)*	*(0.009)*
Bachelors	-0.014**	-0.051**	-0.040**	0.067**	0.037**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.005**	-0.020**	-0.015**	0.026**	0.014**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.007**	-0.026**	-0.021**	0.035**	0.019**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.003	0.002	-0.004	-0.002
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.005	-0.004	0.006	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.002	-0.009	-0.007	0.012	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.031**	-0.115**	-0.090**	0.152**	0.084**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Age	0.021**	0.079**	0.062**	-0.105**	-0.058**
	(0.002)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	-0.004**	-0.024**	-0.033**	0.045**	0.016**
	(0.001)	*(0.003)*	*(0.005)*	*(0.005)*	*(0.002)*
Emotional Stability	-0.009**	-0.035**	-0.031**	0.050**	0.025**
	(0.001)	*(0.003)*	*(0.005)*	*(0.005)*	*(0.002)*
Other Adults	0.008**	0.008	0.019	-0.037**	0.001
	(0.003)	*(0.007)*	*(0.01)*	*(0.01)*	*(0.005)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

Threshold form: $\mu_{i1} = z_i^0\gamma_1, \mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j)},\ \ j = 2 \ldots J-1$

Table B.10: Partial Effects: HSK Ordered Probit Model

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	-0.001	-0.003	-0.002	0.004	0.002
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Separated	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.008)*	*(0.006)*	*(0.011)*	*(0.006)*
Immigrant	0.000	0.001	0.000	-0.001	0.000
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.025**	-0.084**	-0.062**	0.107**	0.064**
	(0.004)	*(0.012)*	*(0.009)*	*(0.015)*	*(0.009)*
Bachelors	-0.018**	-0.059**	-0.043**	0.075**	0.045**
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.008**	-0.025**	-0.018**	0.032**	0.019**
	(0.002)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Year 12	-0.009**	-0.032**	-0.023**	0.040**	0.024**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001*	0.005*	0.003*	-0.006*	-0.004*
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	-0.002	-0.005	-0.004	0.006	0.004
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.003	-0.003	0.004	0.003
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Children 10-14	-0.003	-0.009	-0.007	0.012	0.007
	(0.002)	*(0.007)*	*(0.005)*	*(0.008)*	*(0.005)*
Ln(Income)	-0.038**	-0.129**	-0.094**	0.164**	0.098**
	(0.003)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.007)*
Age	0.018**	0.062**	0.045**	-0.079**	-0.047**
	(0.001)	*(0.004)*	*(0.003)*	*(0.005)*	*(0.003)*
Conscientiousness	0.000	0.000	-0.001	0.000	0.001
	(0.001)	*(0.001)*	*(0.002)*	*(0.001)*	*(0.002)*
Emotional Stability	-0.001	-0.001	0.003	0.002	-0.002
	(0.001)	*(0.001)*	*(0.002)*	*(0.001)*	*(0.002)*
Other Adults	0.004	0.004	-0.010	-0.004	0.006
	(0.002)	*(0.002)*	*(0.005)*	*(0.002)*	*(0.003)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

HSK form: $\sigma_i = e^{w_i^0 \theta}$

Table B.11: Partial Effects: HSK HOPIT Model - health covariates

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
Married	0.000	0.000	0.000	0.000	0.000
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Separated	0.004	0.016	0.013	-0.021	-0.012
	(0.002)	*(0.009)*	*(0.007)*	*(0.011)*	*(0.006)*
Immigrant	0.000	0.000	0.000	0.000	0.000
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.004)*
Postgrad	-0.021**	-0.078**	-0.062**	0.104**	0.057**
	(0.003)	*(0.012)*	*(0.01)*	*(0.016)*	*(0.009)*
Bachelors	-0.013**	-0.051**	-0.040**	0.068**	0.037**
	(0.002)	*(0.007)*	*(0.006)*	*(0.009)*	*(0.005)*
Other Tertiary	-0.005**	-0.019**	-0.015**	0.026**	0.014**
	(0.001)	*(0.005)*	*(0.004)*	*(0.006)*	*(0.003)*
Year 12	-0.007**	-0.028**	-0.022**	0.037**	0.020**
	(0.002)	*(0.006)*	*(0.005)*	*(0.008)*	*(0.005)*
Household Size	0.001	0.003	0.003	-0.004	-0.002
	(0.001)	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*
Children 0-4	0.000	-0.001	-0.001	0.001	0.001
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 5-9	-0.001	-0.004	-0.003	0.005	0.003
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Children 10-14	-0.003	-0.011	-0.009	0.015	0.008
	(0.002)	*(0.007)*	*(0.005)*	*(0.009)*	*(0.005)*
Ln(Income)	-0.030**	-0.114**	-0.091**	0.152**	0.083**
	(0.003)	*(0.009)*	*(0.007)*	*(0.012)*	*(0.006)*
Age	0.020**	0.076**	0.065**	-0.102**	-0.059**
	(0.002)	*(0.004)*	*(0.005)*	*(0.006)*	*(0.003)*
Age (in variance)	-0.001	-0.001	0.003	0.001	-0.002
	(0.001)	*(0.001)*	*(0.003)*	*(0.001)*	*(0.002)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

HOPIT threshold form: $\mu_{i1} = z_i^0\gamma_1$ and $\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j)}, \;\; j = 2 \ldots J-1$

HSK form: $\sigma_i = e^{w_i^0\theta}$

Table B.12: Partial Effects: Threshold and variance covariates

	$Pr(y_i = 1)$	$Pr(y_i = 2)$	$Pr(y_i = 3)$	$Pr(y_i = 4)$	$Pr(y_i = 5)$
			Overall Effect		
Conscientiousness	-0.004**	-0.023**	-0.035**	0.045**	0.017**
	(0.001)	*(0.003)*	*(0.005)*	*(0.005)*	*(0.002)*
Emotional Stability	-0.010**	-0.035**	-0.030**	0.051**	0.024**
	(0.001)	(0.003)	(0.005)	(0.005)	(0.002)
Other Adults	0.007**	0.006	0.023*	-0.035**	-0.002
	(0.003)	(0.007)	(0.01)	(0.01)	(0.005)
			Thresholds		
Conscientiousness	0.000	-0.004	-0.008	0.018*	-0.005
		(0.006)	*(0.012)*	*(0.009)*	*(0.009)*
Emotional Stability	0.000	-0.013*	-0.034**	0.023**	0.024**
		(0.006)	*(0.011)*	*(0.008)*	*(0.009)*
Other Adults	0.000	-0.057**	-0.086**	0.055*	0.088**
		(0.016)	*(0.031)*	*(0.023)*	*(0.024)*
			Variance		
Conscientiousness	0.003	0.004	-0.008	-0.004	0.005
	(0.002)	*(0.003)*	*(0.006)*	*(0.003)*	*(0.004)*
Emotional Stability	-0.005**	-0.007**	0.016**	0.007**	-0.011**
	(0.002)	*(0.003)*	*(0.006)*	*(0.003)*	*(0.004)*
Other Adults	-0.015**	-0.020**	0.044**	0.020**	-0.029**
	(0.005)	*(0.007)*	*(0.015)*	*(0.007)*	*(0.01)*

Outcomes are as follows: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very Good, 5 = Excellent

(Standard errors in parentheses)

HOPIT threshold form: $\mu_{i1} = z_i^0\gamma_1$ and $\mu_{ij} = \mu_{i,j-1} + e^{(\lambda_j + z_i^0\gamma_j)}$, $j = 2 \ldots J-1$

HSK form: $\sigma_i = e^{w_i^0\theta}$