Fixed effects ordered logit model for investigating adaptation in panel data

Anne de Hond

July/August 2017

1 Introduction

This is a study of adaptation to chronic infirmity. It is based on the discussion of the two adaptation studies conducted by Clark, D'Ambrosio and Ghislandi (2016) and Cubí-Mollà, Jofre-Boner and Serra-Sastre (2016) that I replicated. In short, I concluded that the study by Cubí-Mollà, Jofre-Boner and Serra-Sastre had merrit as it took the ordered nature of the dependent variable into account. However, their model heavily relied on the correct parameterization of the fixed effects, which can lead to biased estimates if done incorrectly. I therefore propose to use a fixed effects ordered logit model, which allows me to model the ordered dependent variable without having to make any assumptions regarding the fixed effects. The analysis in question makes use of the "blow-up and cluster" (BUC) estimator put forth by Baetschmann, Staub and Winkelmann (2015) and consistently estimates the parameters via a conditional maximum likelihood approach. The estimator is described in depth in the methodology section. I will be applying the respective methodologies to the SHARE (Survey of Health, Ageing and Retirement in Europe) dataset [1].

In addition, Clark, D'Ambrosio and Ghislandi (2016) set out to measure adaptation through duration represented by dummy variables as opposed to a continuous variable. In accordance, I will also construct dummy variables, since this does not put any restrictions on the effect of duration and allows it to be nonlinear. I measure adaptation in self-perceived health and life satisfaction. I find no support for adaptation in the analysis with self-perceived health, but there is a positive significant effect of duration on life satisfaction. A potential explanation for these discrepant results might be that life satisfaction consists of many factors next to self-perceived health and is therefore less sensitive to changes in health.

[1]DOIs: 10.6103/SHARE.w1.600, 10.6103/SHARE.w2.600, 10.6103/SHARE.w4.600, 10.6103/SHARE.w5.600, 10.6103/SHARE.w6.600, see Börsch-Supan et al. (2013) for methodological details.

2 Methodology

The consistent ordered logit estimation method utilized here is based on the "blow-up and cluster" (BUC) estimator proposed by Baetschmann, Staub and Winkelmann (2015). Before I will go into detail regarding this particular estimator, I will first revisit the theory regarding the conditional log likelihood estimator for the (dichotomized) logit model, since this will aid the explanation of the BUC estimator later on.

The ordered logit model assumes the existence of a latent dependent variable according to:

$$y_{it}^* = x_{it}'\beta + \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T. \tag{1}$$

Here, α_i is treated as a fixed effect. The observed dependent variable y_{it} is constructed from y_{it}^* as follows:

$$y_{it} = k \quad \text{if } \tau_{ik-1} < y_{it}^* \leq \tau_{ik}, \quad k = 1, \dots, K. \tag{2}$$

The thresholds between categories $k-1$ and k can be individual specific, with $\tau_{i0} = -\infty$ and $\tau_{iK} = \infty$. The model assumes a logisitc distribution for the error terms ε_{it}, yielding:

$$F(\varepsilon_{it}|x_{it}, \alpha_i) = \frac{1}{1 + exp(-\varepsilon_{it})} \equiv \Lambda(\varepsilon_{it}) \tag{3}$$

Thus, the probability for individual i at time t of reporting outcome k is given by

$$P(y_{it} = k|x_{it}, \alpha_i) = \Lambda(\tau_{ik} - x_{it}'\beta - \alpha_i) - \Lambda(\tau_{ik-1} - x_{it}'\beta - \alpha_i). \tag{4}$$

Clearly, equation 4 not only depends on x_{it} and β, but also on α_i, τ_{ik-1} and τ_{ik}. Hence, we are first of all faced with an identification problem, since only $\tau_{ik} - \alpha_i$ is identified. Secondly, under fixed T, α_i cannot be estimated consistently, even as NT goes to infinity. This is the incidental parameters problem (Neyman & Scott, 1948). In short panels, this can lead to severe bias in the estimation of β (Green, 2004).

These two methodological concerns are addressed by means of conditional maximum likelihood estimation on a binary variable constructed from the original multinomial variable y_{it}. The binary variable, d_{it}^k, is constructed by dichotomizing the dependent variable at a cut-off point k: $d_{it}^k = \mathbb{1}(y_{it} \geq k)$. Here, the cut-off point can lie anywhere between 2 and K. Consequently, $P(d_{it}^k = 0) = P(y_{it} < k) = \Lambda(\tau_{ik} - x_{it}'\beta - \alpha_i)$ and $P(d_{it}^k = 1) = 1 - \Lambda(\tau_{ik} - x_{it}'\beta - \alpha_i)$. The joint probability of observing $d_i^k = (d_{i1}^k, \dots, d_{iT}^k)' = (j_{i1}, \dots, j_{iT})' = j_i$, where $j_{it} \in \{0, 1\}$, is given by

$$P_i^k(\beta) = P(d_i^k = j_i | \sum_{t=1}^{T} d_{it}^k = g_i) = \frac{exp(j_i' x_i \beta)}{\sum_{j \in B_i} exp(j' xi \beta)}. \tag{5}$$

Here, the sum of all the outcomes over time, $\sum_{t=1}^{T} d_{it}^k = g_i = \sum_{t=1}^{T} j_{it}$, is a sufficient statistic for α_i, since equation 5 is independent of α_i and the thresholds. Moreover, x_i is a $T \times M$ matrix, with M the number of regressors and row t equal to x_{it}. The sum in the denominator of equation 5 concerns the set B_i which consists of all vectors j of length T that have as many elements equal to one as the observed outcome of individual i, g_i:

$$B_i = \left\{ j \in \{0,1\}^T | \sum_{t=1}^{T} j_t = g_i \right\}.$$

The number of elements in the set B_i is equal to

$$\binom{T}{g_i} = \frac{T!}{g_i!(T-g_i)!}.$$

The resulting conditional log likelihood is given by

$$LL^k(b) = \sum_{i=1}^{N} log(P_i^k(b)). \tag{6}$$

The maximization of this likelihood function for a dichotomized dependent variable at any cut-off point k has been shown to be consistent by Chamberlain (1980) and will therefore be referred to as the Chamberlain estimator and denoted by $\hat{\beta}^k$. The first order derivatives of this likelihood function are as follows:

$$s_i^k = \frac{\partial log(P_i^k(b))}{\partial b} = x_i' \left\{ d_i^k - \sum_{j \in B_i} j \frac{exp(j'x_i b)}{\sum_{l \in B_i} exp(l'x_i b)} \right\}. \tag{7}$$

The individual Hessians are given by

$$\begin{aligned} H_i^k(b) = \frac{\partial^2 log(P_i^k(b))}{\partial b \partial b'} &= - \sum_{j \in B_i} \frac{exp(j'x_i b)}{\sum_{l \in B_i} exp(l'x_i b)} \times \\ \left(x_i' j - \sum_{m \in B_i} \frac{exp(m'x_i b)}{\sum_{l \in B_i} exp(l'x_i b} m'x_i \right) &\left(x_i' j - \sum_{m \in B_i} \frac{exp(m'x_i b)}{\sum_{l \in B_i} exp(l'x_i b} m'x_i \right)'. \end{aligned} \tag{8}$$

Note that individuals with constant d_{it}^k do not contribute to the conditional log likelihood, since $P(d_{it}^k = 1 | \sum_{t=1}^{T} d_{it}^k = T) = P(d_{it}^k = 0 | \sum_{t=1}^{T} d_{it}^k = 0) = 1$. Hence, it is worthwhile to obtain β estimates acquired through different cut-off points k, since the group of individuals contributing to the likelihood is likely to change for different cut-off points. In fact, if we employ all possible $K-1$ Chamberlain estimators of β, each individual will contribute at least once to a likelihood function, as long as the observed y_{it}'s of the individual in question are not constant.

The BUC estimator proposed by Baetschmann, Staub and Winkelmann (2015) is exploiting this variation of the ordered variable apparent in the changing selection of individuals contributing to the different Chamberlain estimators by summing over the likelihood functions of all the different Chamberlain estimators. The estimated β is then based on the sum of these conditional log likelihoods. This estimator is more efficient than the Chamerlain estimator since it incorporates more information regarding the variation in y_{it}. The BUC log likelihood is given by

$$LL^{BUC}(b) = \sum_{k=2}^{K} LL^{k}(b), \tag{9}$$

where $LL^k(b)$ is defined in equation 6. The BUC estimator, $\hat{\beta}^{BUC}$, maximizes the likelihood in equation 9 under the restriction that $\hat{\beta}^2 = \ldots = \hat{\beta}^K$. Since the individual Chamberlain estimators are consistent, it is easy to verify the consistency of the BUC estimator. The first order derivatives of the Chamberlain estimators converge to 0 at the true parameter, which means that the sum of the derivatives - equalling the first order derivative of the BUC log likelihood - will converge to 0 as well at β. Given the concavity of the objective function, this ensures that $\hat{\beta}^{BUC}$ converges to β.

We need to cluster the standard errors at the individual level, due to the constructed dependency between the observations. Hence, the information matrix equality is not valid and a cluster robust variance estimator should be used based on the following assymptotic variance (the limiting variance of $\sqrt{n}(\hat{\beta}^{BUC} - \beta)$):

$$Avar(\hat{\beta}^{BUC}) = \left\{ \sum_{k=2}^{K} E(H_i^k(\beta)) \right\}^{-1} \left[\sum_{k=2}^{K} \sum_{l=2}^{K} E(s_i^k(\beta) s_i^l(\beta)') \right] \left\{ \sum_{k=2}^{K} E(H_i^k(\beta)) \right\}^{-1}. \tag{10}$$

In the analysis, the expectations are replaced by their sample analogs and the parameters by their estimated values.

My model aims to measure adaptation to chronic disability or disease. Hence, y_{it} represents either life satisfaction obtained by the question: "*On a scale from 0 to 10 where 0 means completely dissatisfied and 10 means completely satisfied, how satisfied are you with your life?*"[2] or self-perceived health on a 5-point scale. Both dependent variables are considered, since the adaptation process might differ for self-perceived health versus life satisfaction and both regressions could be interesting for further research or policy makers. Life satisfaction is transformed from an 11-point scale into an 8-point scale, since too few subjects in the sample fall into the categories below 3 points on the scale, which incapacitates the BUC estimator. Collapsing the the categories 0, 1, 2 and 3 solves this issue.

The number of limitations with instrumental activities of daily living (IADL) is used as an indicator of chronic infirmity. This measure was chosen, since it reflects the presence of an underlying health problem, without having to specify the cause of the chronic condition itself. Therefore, it is believed to be a more meaningful comparison across different diseases and disabilities as opposed to for example an indication if somebody is chronically ill in and off itself. Moreover, the IADL measure allows one to control for the severity of the disability or disease. Hence, in addition to indicating whether somebody has a chronic condition, the IADL measure is also added to the subsequent analyses to control for the intensity of that condition.

[2]Since life satisfaction is only administered in this manner for waves 2 and up, the analysis with life satisfaction is restricted to waves 2, 4, 5 and 6.

Adaptation is measured by means of a duration specification. If adaptation is indeed occurring within the sample, this ameliorating effect of time should be reflected in the coefficient on the duration variable. Since there is no duration measure available in the SHARE data since the first indication of a limitation with IADL, I construct the duration variable myself. Duration is administered from the moment the IADL measure indicates one or more limitations for the first time. The average of the time between the two adjacent waves is taken as an estimate of the onset of infirmity. The time between two consecutive waves is based on the difference in age of a subject between those respective waves. In the current specification, duration is measured by dummy variables. This was done to allow for a nonlinear effect of time on life satisfaction or self-perceived health. Moreover, my duration specification is prone to error, since the onset of limitations with IADL is measured by averages between waves. The dummy variables will smooth out this error somewhat by clustering duration into intervals. The dummy variables indicate whether the onset of the chronic infirmity is reported within the past 2 years, between 2 and 5.5 years or more than 5.5 years ago. This seemingly arbitrary division was chosen in order to spread out the subjects as much as possible across all dummy groups. The reference group for the regression analysis is chosen to be an indication of chronic infirmity reported within the past 2 years. This allows me to assess how much better or worse subjects score on self-perceived health or life satisfaction with respect to the time since the onset of the chronic infirmity.

The covariates included are age, marital status, labour-force status, years of education and number of children. The reference categories for the first four covariates are aged below 51, married, employed and high school education respectively.

I select all individuals with valid information on either the life satisfaction question or self-perceived health. Moreover, I restrict my sample to those individuals for whom the first reported limitations with IADL are known, thereby creating a prospective study. Finally, the included subjects must remain limited with their IADL until their last observed wave and report to be chronically ill from their first observed limitation with one or more IADL onward. This ensures that the included sample indeed experience an underlying health problem that is of a chronic nature.

3 Results

The results of the fixed effects ordered logit regressions with the BUC estimator can be found in table 1. The age categories clearly have an increasingly negative effect on self-perceived health with respect to the subjects under 50. This is in accordance with the expectation that health deteriorates with age. Moreover, all labour-force status groups are negatively related to self-perceived health with respect to employed subjects. This is not surprising either, since all categories either consist of older subjects (those in retirement) or are likely to contain a disproportionate amount of indisposed subjects (unemployed and inactive). As far as adaptation goes, there appears to be no significant effect of the duration dummies on self-perceived health with respect to the onset of the chronic infirmity. Potential explanations for this lack of effect can be found in the discussion section. There is however a significant effect of those not having any chronic infirmity as opposed to those coping with chronic infirmity for less than two years. Not surprisingly, the subjects reporting no limitations with IADL have a higher self-reported health on average.

If we look at the regression on life satisfaction, we see a significant effect of labour-force status. In accordance with the literature, not having a job negatively affects life satisfaction. Moreover, there is a significant effect of duration, albeit only for the dummy corresponding to an onset of a chronic infirmity of more than 5.5 years ago. This effect is positive, implying there is indeed adaptation after more than 5.5 years of coping with limitations with IADL.

Table 1: FE ordered logit regression

	Life satisfaction	Self-reported health
Age: 51-60	-0.246	-1.156**
	(0.375)	(0.345)
Age: 61-70	0.036	-1.374***
	(0.397)	(0.367)
Age: 71-80	0.071	-1.941***
	(0.415)	(0.385)
Age: 80+	-0.060	-2.621***
	(0.433)	(0.403)
Separated/Single/	-0.243	0.001
Divorced/Widowed	(0.157)	(0.152)
Retired	-0.538***	-0.711***
	(0.142)	(0.149)
Unemployed	-0.720**	-0.615**
	(0.226)	(0.224)
Inactive	-0.598***	-0.786***
	(0.147)	(0.148)
Educ < high school	0.144	0.201
	(0.182)	(0.130)
Educ > high school	0.223	-0.035
	(0.180)	(0.127)
No. of children	0.043	-0.027
	(0.037)	(0.028)
No IADL limit.	0.003	0.816***
	(0.142)	(0.180)
2-5.5 years IADL limit.	0.140	0.071
	(0.075)	(0.081)
> 5.5 years IADL limit.	0.469**	0.223
	(0.147)	(0.151)
No. of IADL limitations	-0.161***	-0.160***
	(0.030)	(0.028)
No. of subjects	5341	5341
No. of observations	13328	15802

Here, the significant codes are as follows: *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$

4 Discussion

something about IADL, dummies nonlinear, model I analyzed adaptation with a fixed effects ordered logit model making use of the BUC estimator proposed by Baetschmann, Staub and Winkelmann (2015). I did not find any proof of adaptation to self-perceived health, but I did find some indication of adaptation in the regression on life satisfaction. The obvious discrepancy between these two results might first of all be due to the nature of the dependent variables. It is likely that people take many factors into account when answering a question about life satisfaction, which makes it into a more holistic measure of well-being. Hence, adaptation to chronic infirmity might be more apparent on the life satisfaction scale, since other factors like social support also affect this measure over time. In contrast, self-reported health is a much more narrowly defined construct, which presumably focuses people's attention on their health situation.

A limitation of this study is that its prospective nature (I can only include individuals whose onset of chronic infirmity is known) severely limits the effect of duration I can measure. My maximum duration time is 9.5 years and only a very small chronically indisposed sub-sample is observed for that period. Taking into account that Cubí-Mollà, Jofre-Bonet, and Serra-Sastre (2016) only found a significant effect of duration after 20 years, the observed period in the SHARE data is simply not long enough. Hence, future research should pay attention to the time scope of their data before conducting the analysis.

Furthermore, the SHARE data set is restricted to individuals of age 50 and up. It is possible that the adaptation process of the SHARE subjects differs from that of younger individuals. For example, it is imaginable that younger subjects find it easier to adapt to a chronic disability or disease, since they can more easily change their occupation or are less reliant on social support. Thus, the differences in the adaptation process for different age categories is a question for future research.

REFERENCES

Baetschmann, G., Staub, K.E. & Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society, A, 178(3),* 685-703.

Clark, A.E., D'Ambrosio, C. & Ghislandi, S. (2016). Adaptation to poverty in long-run panel data. *Review of Economics and Statistics, 98(3),* 591-600.

Cubí-Mollà, P., Jofre-Bonet, M. & Serra-Sastre, V. (2016). Adaptation to health states: Sick yet better off? *Health Economics,* 1-18.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal, Vol 7(1)*, 98-119.

Neyman, J. & Scott E.L. (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica, 16(1),* 1-32.