A dynamic ordered logit model with fixed effects

Chris Muris, Pedro Raposo, and Sotiris Vandoros*

August 4, 2020

Abstract

We study a fixed-T panel data logit model for ordered outcomes that accommodates fixed effects and state dependence. We provide identification results for the autoregressive parameter, regression coefficients, and the threshold parameters in this model. Our results require only four observations on the outcome variable. We provide conditions under which a composite conditional maximum likelihood estimator is consistent and asymptotically normal. We use our estimator to explore the determinants of self-reported health in a panel of European countries over the period 2003-2016. We find that: (i) the autoregressive parameter is positive and analogous to a linear AR(1) coefficient of about 0.25, indicating persistence in health status; (ii) the association between income and health becomes insignificant once we control for unobserved heterogeneity and persistence.

*Muris: Department of Economics, McMaster University. Contact: muerisc@mcmaster.ca. Raposo: Católica Lisbon School of Business and Economics. Contact: pedro.raposo@ucp.pt. Vandoros: King's College London and Harvard University. Contact: s.vandoros@kcl.ac.uk. The data were made available by Eurostat (Contract RPP 132-2018-EU-SILC). We are grateful to Irene Botosaru and Krishna Pendakur for very helpful discussions.

		$P\left(Y_{i,t+1}=y' \middle| Y_{i,t}=y\right)$				
	y/y'	1	2	3	4	5
$P\left(Y_{i,t}=y\right)$	1	36.48	43.40	13.84	5.03	1.26
	2	10.23	44.44	35.38	8.77	1.17
	3	0.88	10.23	52.18	30.69	6.02
	4	0.15	1.08	14.87	59.74	24.16
	5	0.08	0.29	4.18	33.38	62.07

Table 1: Current and future self-reported health, United Kingdom.

1 Introduction

Certain individual-level conditions may tend to persist over time, in the sense that a condition has a memory of a previous period's state, or may involve an element of adaptation. Furthermore, the way individuals experience the same condition may vary, and they may also have a different understanding of how this is measured. A common example that fulfils these characteristics, and which is used extensively in the literature, is self-reported health status. Health status often depends on its value in the previous period, as health conditions may persist over time, given that recovery can take long, and that an illness may even have permanent effects. For example, Table 1 presents a transition matrix for self-reported health status in the United Kingdom for the period 2003-2016.[1] For individuals that report a value of current health in a given year (rows, on a 5-point scale with 5 being the highest), it shows the relative proportion of those that report a certain value in the subsequent year (columns). A striking feature of this transition matrix is that a lot of mass is on or near the main diagonal. This feature is found across all countries in our analysis. In other words, self-reported health status is persistent: individuals tend to stay in the same level of health.

There are at least two explanations for this observed persistence (Heckman, 1981; Honoré and Kyriazidou, 2000): unobserved heterogeneity and state dependence. Consider first *unobserved heterogeneity*. It refers to unobservable characteristics that affect the propensity to report higher health status. Unobserved heterogeneity is important in the literature on health status, because self-

[1]More information about the data and source is in Section 5, where we analyze this data using the methodology proposed in this paper.

reported health has been used extensively in the literature as a measure of health outcomes (see for example Bound and Waidmann, 1992; Banerjee et al., 2004; Gravelle and Sutton, 2009; McInerney and Mellor, 2012). It is often viewed as a limitation that self-reported measures are subjective. For example, reporting one's own health may depend on cultural factors (Jylhä et al., 1998; Baron-Epel et al., 2005; Jürges, 2007), and people may have a different understanding of reference points for health (Groot, 2000; Sen, 2002). Previous studies have used vignettes to address cross-country differences in reporting of health and disability (King et al., 2004; Salomon et al., 2004; Kapteyn et al., 2007). However, the issue with unobserved heterogeneity across individuals remains. As a result, it is important to take into account the role of unobserved heterogeneity when analyzing self-reported health data. The appropriate econometric approach to this is to allow for fixed effects.[2]

A number of studies have found a positive association between income and health (Carrieri and Jones, 2017; Ettner, 1996; Frijters et al., 2005; Mackenbach et al., 2005), but empirical evidence of a strong effect is sometimes limited (Larrimore, 2011; Gunasekara, 2011; Johnston et al., 2009).

[2]Studies have long debated the accuracy and reliability of subjective measures of health, such as self-reported health status (see for example Butler et al., 1987; Lindeboom and van Doorslaer, 2004; Johnston et al., 2009). As an alternative response to these concerns, a number of objective measures of health have been included in household surveys. These include blood pressure, BMI, the number of medicines taken (Health Survey England, 2019), the number of sick days off work, the number of days hospitalised (BHPS, 2019) etc. Some household surveys ask respondents to perform a task such as walking across the room or buttoning a shirt to capture any limitations (SHARE, 2019). Indexes such as the EQ-5D index are being used to cover different types of conditions and merge them into a single measure. The Euro-D scale measures mental health, and the CASP-12 index captures quality of life. However, these objective measures are often very specific to particular diseases, and even when creating a relevant index it may be impossible to include and accurately reflect all conditions. As such, while objective measures may accurately capture some health conditions, they have serious limitations in capturing the overall picture of one's health.

Nevertheless, the literature on the impact of economic downturns (which mean reduced income) has previously demonstrated positive effects of unemployment on health (Ruhm, 2000; Ruhm and Black, 2002), and more recently, no effect (Ruhm, 2015). What also appears to matter, at least in terms of happiness, apart from absolute income, is also relative income, i.e. how one's income compares to that of those around them (Frijters et al., 2008). The relationship between income and health is endogenous and complex, and both can be correlated with other factors, that are not always measured and included in empirical models. For example, Gunasekara et al. (2011) found that when controlling for unmeasured confounders, the association between the two becomes weaker. With regards to self-reported hypertension in particular, Johnston et al., (2009) found no link to income – something that did change when using objective measures. In our paper, using models that do not control for individual unobserved heterogeneity yields a positive and statistically significant coefficient (Table 3). However, this becomes insignificant when using fixed effects, suggesting that there are other factors that potentially drive the association between the two variables.

Consider now the second explanation for the observed persistence in health outcomes: *state dependence*. It refers to the possibility that past self-reported health status may be related to current self-reported health status even after conditioning on unobserved heterogeneity. State dependence arises if actual (as opposed to self-reported) health shocks are persistent, in the sense that a shock on health can have a long-lasting effect (a typical example is injury leading to disability). Contoyannis et al. (2004), using a random effects approach, found evidence for such persistence in respondents of the British Household Panel survey.

State dependence in self-reported health can also arise due to adaptation: self-reported health status may change over time for a person whose actual health has not changed. People tend to adapt to good or bad developments in life. According to the Global Adaptive Utility Model, individuals reallocate weights on various domains of life in order to maintain their previous level of utility (Bradford and Dolan, 2010). Similarly, the AREA model developed by Wilson and Gilbert (2008), suggests that attention is focused on a change, followed by reaction, explanation, and, finally, adaptation. This also applies to health, as health status tends to improve even when individuals' health has actually not experienced any objective change (Daltroy et al., 1999; Damschroder et al., 2005), and time since diagnosis is positively associated with self-reported health (Cubí-Mollá et al., 2017). Whether persistence or adaptation, or both, characterise a variable, this calls for a dynamic element in a model.

Overall, the challenges with studying self-reported health status is that (a) people with the same actual health status might be reporting different health levels; and (b) health shocks can have a lasting effect. Against this background, we propose and analyze a panel data ordered logit model that includes both fixed effects and a lagged dependent variable. This allows a researcher faced with panel data and an ordinal outcome variable to disentangle unobserved heterogeneity from state dependence, and to quantify state dependence. Thus, we address the limitations of using self-reported health as a proxy for individuals' health. Our contribution is important for studies using subjective health measures as it can help correct biases that naturally occur when using this type of measure.[3]

Specifically, we study the *dynamic ordered logit model with fixed effects:*

$$Y_{i,t}^* = \alpha_i + X_{i,t}\beta + \rho 1\left\{Y_{i,t-1} \geq k\right\} - U_{i,t},\ t = 1,2,3, \tag{1}$$

$$Y_{i,t} = \begin{cases} 1 & \text{if } Y_{i,t}^* < \gamma_2, \\ 2 & \text{if } \gamma_2 \leq Y_{i,t}^* < \gamma_3, \\ \vdots & \\ J & \text{if } Y_{i,t}^* \geq \gamma_J, \end{cases} \tag{2}$$

$$U_{i,t} \,|\, (\alpha_i, X_i, Y_{i,<t}) \sim LOG(0,1),\ t = 1,2,3, \tag{3}$$

[3] For example, happiness is perceived and reported differently across individuals and people adapt to things that make them happy (Layard, 2006), while shocks on happiness can have a scarring effect on next periods (Clark et al., 2001).

where $2 \leq k \leq J$ is a fixed and known cutoff for the lagged dependent variable. The person-specific parameter α_i captures unobserved heterogeneity, which we allow to be correlated with the other quantities in the model in an unrestricted way (fixed effects). The time-varying covariates $X_{i,t}$ are collected across time periods in $X_i = (X_{i,1}, X_{i,2}, X_{i,3})$, and the lagged dependent variables for period t are collected in $Y_{i,<t} = (Y_{i,0}, \cdots, Y_{i,t-1})$. The autoregressive parameter ρ is the regression coefficient on the lagged dependent variable $1\left\{Y_{i,t-1} \geq k\right\}$; β is the regression coefficient on the contemporaneous covariates; and the threshold parameters γ_j map the underlying latent variable $Y_{i,t}^*$ into the observed ordered outcome $Y_{i,t}$. Equation (3) restricts the error terms $U_{i,t}$ to be i.i.d. logistic, and is a strict exogeneity assumption on the regressors and past outcomes.[4]

This model combines a number of noteworthy features. First, it is a model for discrete ordered outcomes, and therefore a *nonlinear* model. Second, it is *dynamic*, in the sense that the current outcome depends directly on the outcome in the previous period. This feature, called *state dependence*, is governed by the autoregressive parameter ρ. Third, it allows for *unobserved heterogeneity* in an unrestricted way, i.e. it is a *fixed effects* model. Fourth, the model is only specified for a *small number of time periods*, $T = 3$. Period 0 is unmodelled, but an observation on the outcome variable in time 0 is required for identification.

We believe that we are the first to provide identification and estimation results for all common parameters in a dynamic ordered logit model with fixed effects and a fixed number of time periods. Using four time periods of data on the ordinal outcome variable, we identify the autoregressive coefficients on the lagged dependent variable, and the regression coefficients on the exogenous regressors. We also identify the threshold parameters, which makes it possible to interpret the magnitude of the estimated coefficients. This distinguishes the ordered choice model from the dynamic binary choice model with fixed effects, where such an interpretation is not available.

[4] The dynamics in our model are restricted to depend on $Y_{i,t}$ through $1\left\{Y_{i,t-1} \geq k\right\}$ only. An alternative model for which we can identify some features is one that is linear in its history, i.e. $Y_{i,t}^* = \alpha_i + X_{i,t}\beta + \rho Y_{i,t-1} - U_{i,t}$. We were unable to use our approach to obtain identification in the more general model with $Y_{i,t}^* = \alpha_i + X_{i,t}\beta + \sum_{j=2}^{J} \rho_j 1\left\{Y_{i,t-1} = j\right\} - U_{i,t}$.

Our identification result suggest a composite conditional maximum likelihood estimator for the parameters in our model. We establish conditions under which that estimator is consistent and asymptotically normal.

We use our estimator to investigate the determinants of self-reported health, focusing on the link between income and health in a panel of European countries over the period 2003-2016. We obtain two main findings. First, even after controlling for unobserved heterogeneity, persistence plays a positive and significant role in one's self-reported health. In other words, one's health is dependent on the health in the previous period, which is a reasonable thing to expect, as health problems may expand over a number of periods, or become permanent. Quantitatively, we estimate a persistence parameter that is analogous to an autoregressive parameter of about 0.25 in a linear AR(1) model. Second, we find that, when controlling for unobserved heterogeneity, the link between income and health becomes statistically insignificant, suggesting that other factors might explain the association between the two. This is in line with studies that have found a smaller or insignificant association when using fixed effects (Gunasekara, 2011; Larrimore, 2011).

2 Related literature in econometrics

We believe that our paper is the first to provide identification and estimation results for a panel data model with (i) ordered outcomes; (ii) a lagged dependent variable; (iii) fixed effects; and (iv) a fixed number of time periods. Our econometric contribution is related to several strands of literature, each of which features a subset of these features.

Most closely related to our paper is the literature on binary and multinomial choice models with fixed effects and lagged dependent variables, which features all but (i). The seminal work by Honoré and Kyriazidou (2000) builds on Cox (1958) and Chamberlain (1985) to estimate the parameters in dynamic binary choice logit model with fixed effects and time-varying regressors. Hahn (2001) discusses the information bound for a special case of their model. Honoré and Kyriazidou (2019) discuss identification of some closely related models. Honoré and Weidner (2020) construct moment conditions that shed light on identification in this and related models, and provide a $\sqrt{n}$-consistent estimator. Honoré and Tamer (2006), Aristodemou (2020) and Khan et al. (2020) obtain results for models that do not have logistic errors. For the static multinomial model, Chamberlain (1980) studies the logit case; Shi et al. (2008) provides results for the general static; and Magnac (2000) studies the dynamic version. We supplement these results by showing that, in an *ordered* choice model, the thresholds in the latent variable model can be identified along with the regression coefficients and the autoregressive parameter. This allows for a quantitative interpretation of true state dependence. Such an interpretation is not available in the binary and multinomial choice models.

The literature on static ordered logit models with fixed effects features all but (ii). This model was analyzed by Das and van Soest (1999), Baetschmann et al. (2015), and Muris (2017). Our result differs from the results in those papers, because we provide results for a *dynamic* version of the ordered logit model.

The literature on random effects dynamic ordered choice models features all but (iii). Random effects dynamic ordered choice models have been studied and applied extensively (Contoyannis et al., 2004; Albarran et al., 2019). Such approaches require strong restrictions on the relationship between the unobserved heterogeneity and the exogeneous variables in the model. Such restrictions are usually unappealing to economists, as evidenced by the fact that they are rarely used in linear models. Our approach does not impose random effects restrictions and is the first to provide a fixed effects approach for dynamic ordered choice models.

Note that our approach is fixed-T consistent. The difficulty of allowing for fixed effects is alleviated when one can assume that $T \to \infty$, referred to as "large-T". Large-T fixed effects dynamic ordered choice models have been studied by Carro and Traferri (2014) and Fernández-Val et al. (2017), see also Carro (2007) for the binary outcome case. In the large-T case, one can use techniques that correct for the bias that comes from including fixed effects in the nonlinear panel model. This approach does not feature (iv). These techniques are not appropriate for our empirical application, which is a rotating panel with $T = 4$.

One limitation of our approach is that we restrict the way in which the lagged dependent variable enters the model. The random effects and large-T approach can accommodate a richer dynamic specification. We leave for future work whether such an extension is possible with a fixed-T fixed-effects approach.

3 Identification

We normalize $\gamma_k = 0$, where k is as in equation (1). This scale normalization is without loss of generality because the scale of α_i is unrestricted. Our model implies that the binary variable $D_{i,t}(k) = 1\{Y_{i,t} \geq k\}$ follows the dynamic binary choice logit model in Honoré and Kyriazidou (2000), HK hereafter. Specifically, equation (3) in HK applies to the transformed model

$$D_{i,t}(k) = 1\left\{X_{i,t}\beta + \rho D_{i,t-1}(k) + \alpha_i - U_{i,t} \geq 0\right\},$$

i.e. the transformed model follows a dynamic binary choice logit model with fixed effects. The implied conditional probabilities relevant for our analysis are

$$P\left(D_{i,0}(k) = 1 \middle| X_i, \alpha_i\right) \equiv p_0(X_i, \alpha_i), \tag{4}$$

and, for $t = 1, 2, 3$,

$$P\left(D_{i,t}(k) = 1 | X_i, \alpha_i, D_{i,<t}(k)\right) = \frac{\exp\left(\alpha_i + X_{i,t}\beta + \rho D_{i,t-1}(k)\right)}{1 + \exp\left(\alpha_i + X_{i,t}\beta + \rho D_{i,t-1}(k)\right)}, \tag{5}$$

where we have let $D_{i,<t}(k) = \left(D_{i,0}(k), \cdots, D_{i,t-1}(k)\right)$. HK provide conditions that guarantee identification of β and ρ by constructing a conditional probability that features (β, ρ) but that is free of α_i.

If $Y_{i,t}$ has at least three points of support, there is information in Y_{it} beyond $D_{it}(k)$. In the remainder of this section, we show that this information can be used to identify the threshold parameters

$$\gamma \equiv (\gamma_2, \gamma_3, \cdots, \gamma_{k-1}, \gamma_{k+1}, \cdots, \gamma_J).$$

This leads to an interpretation of the magnitude of (β, ρ) that is not available for the dynamic binary choice model. Muris (2017, Section III.C) discusses this for the static panel data ordered choice models ($\rho = 0$).

We now construct a conditional probability that features (β, ρ, γ) but not the incidental parameters α_i. To this end, extend the definition

$$D_{i,t}(j) = 1\{Y_{i,t} \geq j\}, 2 \leq j \leq J,$$

to thresholds $j \neq k$, and abbreviate $D_{i,t} \equiv D_{i,t}(k)$. Define the events $\left(A_{j,l}, B_{j,l}, C_{j,l}\right)$, with $2 \leq j \leq k \leq l \leq J$,[5] as follows:

$$\begin{aligned}
A_{j,l} &= \left\{D_{i,0} = d_0, D_{i,1} = 0, D_{i,2}(l) = 1, D_{i,3}(j) = d_3\right\}, \\
B_{j,l} &= \left\{D_{i,0} = d_0, D_{i,1} = 1, D_{i,2}(j) = 0, D_{i,3}(l) = d_3\right\}, \\
C_{j,l} &= A_{j,l} \cup B_{j,l}.
\end{aligned}$$

For $d_0 = d_3 = 0$, the event $A_{j,l}$ corresponds to moving up in the middle periods $t = 1, 2$, starting below k to moving up to at least $l \geq k$. The event $B_{j,l}$ corresponds to moving down in the middle periods, starting from at least k and moving below $j \leq k$.

[5]Choosing $j \leq k$ guarantees that when $D_{i,2}(j) = 0$, the lagged dependent variable in period 3 is 0. The opposite is true for $l \geq k$ and $D_{i,2}(l) = 1$. There would be no gain from considering a threshold different from k in the first period. Using k as the threshold in the second period is the only way to cancel out the threshold parameters from period 1. Given that we are using the subpopulation $X_{i,2} = X_{i,3}$, the fact that j, l are used alternately in periods 2 and 3 does not create additional difficulties.

If $j=k=l$, the event $C_{k,k}$ corresponds to switchers (observations with $D_{i1}+D_{i2}=1$), as in HK. In the ordered model, it is possible to vary the cutoffs in the periods $t=2,3$ if the dependent variable has more than two points if support. Varying the cutoffs over time is what distinguishes our conditioning event from that in HK. It is what allows us to identify the threshold parameters.

The following sufficiency result shows that different choices of (j,l) reveal different combinations of thresholds in certain conditional probabilities that do not depend on the incidental parameters α_i. In what follows, the logistic function is denoted by $\Lambda(u)=\exp(u)/(1+\exp(u))$, and the change in the regressors from period 1 to 2 by $\Delta X_i = X_{i2}-X_{i1}$.

Theorem 1 (Sufficiency)**.** *For the dynamic ordered logit model with fixed effects, for any (j,l) such that $2\leq j\leq k\leq l\leq J$, and for any $d_0,d_3\in\{0,1\}$,*

$$P\left(A_{j,l}\middle|X_i,C_{j,l},X_{i,2}=X_{i,3}\right)=1-\Lambda\left(\Delta X_i\beta+\rho\left(d_0-d_3\right)+\left(1-d_3\right)\gamma_l+d_3\gamma_j\right) \tag{6}$$

$$P\left(B_{j,l}\middle|X_i,C_{j,l},X_{i,2}=X_{i,3}\right)=\Lambda\left(\Delta X_i\beta+\rho\left(d_0-d_3\right)+\left(1-d_3\right)\gamma_l+d_3\gamma_j\right). \tag{7}$$

Identification of the model parameters comes from considering all possible combinations of cutoffs. It is clear from Theorem 1 that different choices for (j,k,l,d_0,d_3) reveal information about distinct linear combinations of (ρ,γ). By considering multiple choices of (j,k,l,d_0,d_3), and then aggregating the resulting information, we can identify all the model parameters. We require an additional assumption before stating our main identification result.

Assumption 1. *For all (j,l) such that $2\leq j\leq k\leq l$, and for all $d_0,d_3\in\{0,1\}$*

$$Var\left(\Delta X_i|X_{i,2}=X_{i,3},C_{j,l}\right)$$

is invertible.

This assumption guarantees that for each choice of (j,l), there is sufficient variation in ΔX_i in the subpopulation of stayers $X_{i,2}=X_{i,3}$ to identify the regression coefficient. This assumption can be weakened: we only need sufficient variation for some (j,l). However, if it fails for sufficiently many (j,l), identification of some of the threshold parameters may fail.

Denote by $Y_i = (Y_{i,0}, Y_{i,1}, Y_{i,2}, Y_{i,3})$ the time series of dependent variables for a given individual.

Theorem 2 (Identification)**.** *If Assumption 2 holds, then (β, ρ, γ) can be identified from the joint distribution of the vector (X_i, Y_i) generated by the dynamic ordered logit model with fixed effects.*

4 Estimation

Theorem 1 suggests that, for each choice of $2 \le j \le k \le l$, we could use a conditional maximum likelihood estimator (CMLE) to estimate a linear combination of the model parameters. Theorem 2 suggests that a composite CMLE (CCMLE), based on the combination of conditional likelihoods across all choices of (j,k,l), may be used to estimate the model parameters (β, ρ, γ). In this section, we define that CCMLE and establish conditions under which it has desirable large sample properties. We focus on the discrete regressor case. Results for continuous regressors can be obtained by adapting Theorems 1 and 2 in HK to our case.

The binary random variable

$$C_{i,jl} = 1\left\{(D_{i,1} = 0, D_{i,2}(l) = 1) \text{ or } (D_{i,1} = 1, D_{i,2}(j) = 0)\right\} \times 1\left\{X_{i,2} = X_{i,3}\right\}.$$

indicates whether i's time series fits the description in $C_{j,l} = A_{j,l} \cup B_{j,l}$, and that it is also a "stayer" in the sense that $X_{i2} = X_{i3}$. Note that if $C_{i,jl} = 1$, then $D_{i,1} = 1$ implies that the individual time series is of the type $B_{j,l}$. Similarly, if $C_{i,jl} = 1$, then $D_{i,1} = 0$ implies that individual i is of type $A_{j,l}$.

In the log-likelihood contribution below, (8), we have substituted $D_{i,0}$ for d_0 in equation (7). The value to substitute for d_3 depends on whether we are in case A or B. To that end, define

$$D_{i,3,jl} = \begin{cases} D_{i,3}(j) & \text{if } D_{i,1} = 0, \\ D_{i,3}(l) & \text{if } D_{i,1} = 1. \end{cases}$$

The conditional log likelihood contribution for individual i, for cutoffs (j,l), $2 \leq j \leq k \leq l \leq J$, can then be written

$$l_{i,jl}(\beta,\rho,\gamma_j,\gamma_l) = C_{i,jl}\left[D_{i,1}\ln\left\{\Lambda\left(\Delta X_i\beta + \rho\left(D_{i,0}-D_{i,3,jl}\right)+\gamma_l\left(1-D_{i,3,jl}\right)+\gamma_j D_{i,3,jl}\right)\right\}+\right.$$
$$\left.+\left(1-D_{i,1}\right)\ln\left\{1-\Lambda\left(\Delta X_i\beta + \rho\left(D_{i,0}-D_{i,3,jl}\right)+\gamma_l\left(1-D_{i,3,jl}\right)+\gamma_j D_{i,3,jl}\right)\right\}\right]. \tag{8}$$

The CCMLE is

$$\widehat{\theta}_n = \left(\widehat{\beta}_n, \widehat{\rho}_n, \widehat{\gamma}_n\right) = \arg\max \frac{1}{n}\sum_{2\leq j\leq k\leq l}\sum_{i=1}^{n} l_{i,jl}(\beta,\rho,\gamma_j,\gamma_l), \tag{9}$$

where we have implicitly imposed $\gamma_k = 0$ in the definition of $l_{i,jl}$.

We maintain the following assumption to establish the asymptotic properties of the CCMLE.

Assumption 2 (Stayers). $P(X_{i,2} = X_{i,3}) > 0$.

With additional technical work, this assumption can be relaxed to the case where $X_{i,2} - X_{i,3}$ is continuously distributed with positive density around zero, see HK's Theorem 1 and 2.

Theorem 3. *Let $\{(Y_i, X_i), i = 1, \cdots n\}$ be a random sample of size n from the dynamic ordered logit model with fixed effects with true parameter values $\theta_0 = (\beta_0, \rho_0, \gamma_0)$. Under Assumptions 1 and 2, and for any value of θ_0,*

$$\widehat{\theta}_n \xrightarrow{p} \theta_0 \text{ as } n \to \infty.$$

Furthermore,

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, H^{-1}\Sigma H^{-1}\right) \text{ as } n \to \infty,$$

where Ω as the variance of the score of the composite likelihood, defined in (21), and H is the associated Hessian defined in (22).

Remark 1. The convexity of the summands in (9) means that the objective function is convex. We compute the CCMLE using the Newton-Raphson algorithm in R's nlm function (R Core Team, 2020). Supplying analytical gradients and Hessians speeds up the estimation.

5 Persistence in self-reported health status

Our analysis uses panel data for the period 2003-2016 from the European Union Statistics on Income and Living Conditions (EU-SILC), see Eurostat (2017) for detailed documentation. The microdata is publicly available upon request.[6] EU-SILC provides a set of indicators on income and poverty, social inclusion, living conditions and, importantly, health status. For each country in the European Union, plus Iceland, Norway, and Switzerland, EU-SILC contains data on a representative sample of the population of those 18 years and older.

EU-SILC is a rotating panel. Every individual is followed over a period of two to four years. The total number of individual-years for the period 2003-2016 is 1273877. Our identification result demands four observations per individual, so we restrict attention to individuals that report valid information on their health status for 4 consecutive years. This restriction, and the restriction that the explanatory variables that we use in the analysis below have non-missing information, leaves us with a sample of 260601 individuals, for 1042404 individual-years. The proportion of incomplete samples differs across countries. As a result, the sample we work with may not be representative of EU-SILC's population. For example, out of the 27 countries that contribute to our sample, the largest contributors are Italy (with 43385 individuals), Spain (25634), and Poland (22628); the smallest are Portugal (12), Iceland (1496), and Slovakia (1982).

The outcome variable in our analysis is self-reported health status: self-perceived physical health, elicited during EU-SILC interviews. The person answers the question on how she perceives her physical health to be in general, at the date of the survey, by classifying it as one of: (1) bad and very bad (12% in our sample); (2) fair (26%); (3) good (44%); (4) very good (19%).[7] Out of $260601 \times 3 = 781803$ health transitions that we observe, most often there is no change in health

[6]The data were made available to us by Eurostat under Contract RPP 132-2018-EU-SILC.

[7]We have merged the separate categories "bad" and "very bad" in the original reported variable,

		Log income	
		mean	sd
health status	1	8.68	0.92
	2	8.95	0.94
	3	9.29	0.94
	4	9.51	0.90
		ΔLog income	
		mean	sd
	-3	0.05	0.47
Δhealth status	-2	0.06	0.43
	-1	0.07	0.40
	0	0.08	0.38
	1	0.08	0.40
	2	0.08	0.43
	3	0.07	0.49

Table 2: Health and income

status (65.6%). Decreases by one unit (16%) are slightly more frequent than increases by one unit (15%). Two-unit increases (1.2%) and decreases (1.5%) are infrequent, and three-unit increases (0.08%) and decreases (0.12%) are rare.

Table 2 relates the outcome variable, and changes to the outcome variable, to our main explanatory variable of interest, log income (total disposable household equivalised income). Household income was scaled using the composition and size of each household. This scale is based on the OECD modified equivalence scale, which gives a weight of 1.0 to the first adult in the household, 0.5 to other adults and 0.3 to each child (under 14 years old).

The table provides descriptive statistics for log income in our sample, grouped by health status. The top panel is in levels. Average income is increasing in health status. The bottom panel is in changes, which represents one way to control for unobserved heterogeneity. The implied increases for changes are close to zero, hinting at the imported role of unobserved heterogeneity.

Table 3 reports a set of descriptive statistics on income and other explanatory variables, described in the next few paragraphs. In our analysis below, we control for some time-varying variables that are standard in the literature. First, the number of children is measured as the number

because there is only a small fraction of observations with “very bad” health status.

of persons living in the private household that are age 14 or less, top-coded at 3 children. In our sample, 74% of respondents have no children, and the average number of children is 0.40. Second, marriage status is a dummy variable that indicates being married or living together. The majority of the individuals are married (61%). Third, we use a self-reported indicator for labor market status variable that we map onto 4 values: (1) employed, 51%; (2) unemployed, 5.2%; (3) retired, 12% and (4) other, 32%. The value "other" includes students, permanently disabled or unfit to work, and fulfilling domestic tasks and care responsibilities.

In our fixed effects results below, we do not further control for variables that do not change over the sample period for a given individual. However, we include a set of time-invariant explanatory variables when we obtain results for non-fixed effects estimators.[8] Table 3 provides descriptive statistics for such variables. The total sample contains slightly more females (54%) than males (46%). The proportion of individuals aged between 18 and 25 is 8.2%; 21% of individuals are aged 65 or more. With regards to education, 1.3% of the sample have no schooling (0); 14% have attended primary school (1); 19% have lower secondary education (3); 42% have upper secondary education (4); 3.8% have post-secondary education (5) and 20% have tertiary education (6). Geographically, 39% of individuals live in areas with a high degree of urbanisation and 39% live in areas with low levels of urbanisation.

We estimate the parameters in the dynamic ordered choice model with fixed effects, with latent variable outcome equation

$$\begin{aligned} SRH^*_{i,t} = \alpha_i + \rho 1\left\{SRH_{i,t-1} \geq 3\right\} + \beta_1 \log income_{it} + \beta_2 child_{it} + \beta_3 married_{it} + \\ + \beta_4 unemp_{it} + \beta_5 retired_{it} + \beta_6 other_{it} - U_{it}. \end{aligned} \tag{10}$$

Regression results are presented in Table 4. The first four columns (a-d, "DOLFE", for *d*ynamic *o*rdered *l*ogit with *f*ixed *e*ffects) presents the results for (10) using the estimator described in Sec-

[8]Coefficient estimates for these variables are omitted from the main text, and reported in Appendix B.

		mean	sd
health status	(1) bad and very bad	0.115	
	(2) fair	0.260	
	(3) good	0.437	
	(4) very good	0.187	
Time-varying explanatory variables			
log income		9.172	0.965
child		0.401	0.755
married		0.614	
employment status	employed	0.512	
	unemployed	0.052	
	retired	0.117	
	other	0.319	
Time-invariant explanatory variables			
age group	$[18;25]$	0.082	
	$]25;35]$	0.145	
	$]35;45]$	0.188	
	$]45;55]$	0.194	
	$]55;65]$	0.182	
	$]65;\infty]$	0.208	
urbanisation	high	0.388	
	middle	0.224	
	low	0.388	
male		0.460	
educ	no schooling	0.013	
	primary	0.143	
	lower secondary	0.188	
	upper secondary	0.420	
	post-secondary	0.038	
	tertiary	0.199	
n		260601	
T		4	
nT		1042404	

Table 3: Descriptive statistics

tion 4. Different values of h refer to a bandwidth parameter that we introduce because one of the explanatory variables is continuous, as in HK. Column (d) omits the employment variables, to check whether relationship between employment status and income matters for estimation of the effect of income on health.

We also present estimation results for different estimators. Results for the static ordered logit model with fixed effects, i.e. setting $\rho = 0$ in (10), are obtained using the estimator in Muris (2017), and presented in column (e) ("FEOL"). Column (f) ("DOL") estimates a dynamic ordered logit model without fixed effects, i.e. (10) with $\alpha_i = 0$. Column (g) ("OL") presents results for cross-sectional ordered logit estimator that does not take into account fixed effects or dynamics (i.e. $\alpha_i = \rho = 0$ in (10)). We also present results for a static linear model with (h, "FELM") and without (i, "LM") fixed effects. The standard errors for all estimators are obtained using the bootstrap (500 replications). For the estimators that are not of the fixed effects type, we additionally control for education, gender, education level and the level of urbanisation. DOLFE uses four periods of data, corresponding to $t = 0, 1, 2$. For comparability, the other dynamic estimator also uses periods 0,1,2; static estimators use periods 1,2.

Income. Our main explanatory variable of interest is income (log income, coefficient β_1). Across almost all specifications, we find a positive association between income and self-reported health. The only exception is column (b), where the point estimate is negative, and about the same magnitude as the standard error.

Controlling for unobserved heterogeneity leads to a very strong reduction in the magnitude of the association. For example, for the static case, a comparison of columns (e) and (g) says that, for the static case, controlling for unobserved heterogeneity reduces the coefficient on income by more than a factor 20. For this comparison, note that the threshold differences increase, suggesting that the scale increases; compare also the coefficients on the other variables, with an unchanged order of magnitude. We are not the first to observe a limited association between income and self-reported health. In a review of the literature, Gunasekara et al. (2011) found a small positive link between income and self-reported health, which is reduced when controlling for unmeasured confounders.

	(a) DOLFE $h=1$	(b) DOLFE $h=0.1$	(c) DOLFE $h=10$	(d) DOLFE $h=1$	(e) FEOL	(f) DOL	(g) OL	(h) FELM	(i) LM
log(income)	0.049	-0.047	0.059	0.061	0.020	0.340	0.492	0.003	0.194
	(0.033)	(0.056)	(0.029)	(0.029)	(0.019)	(0.004)	(0.004)	(0.003)	(0.002)
child	-0.030	0.006	-0.031	-0.026	0.021	0.060	0.089	0.002	0.033
	(0.051)	(0.069)	(0.050)	(0.049)	(0.032)	(0.005)	(0.005)	(0.005)	(0.002)
married	0.139	-0.041	0.157	0.130	0.164	0.073	0.141	0.029	0.062
	(0.087)	(0.119)	(0.086)	(0.088)	(0.053)	(0.007)	(0.008)	(0.009)	(0.003)
unemp	-0.188	-0.230	-0.178		-0.196	-0.242	-0.308	-0.033	-0.127
	(0.070)	(0.110)	(0.068)		(0.038)	(0.014)	(0.015)	(0.007)	(0.006)
retired	-0.132	-0.043	-0.139		-0.154	-0.050	-0.097	-0.027	-0.047
	(0.082)	(0.119)	(0.080)		(0.041)	(0.010)	(0.011)	(0.007)	(0.004)
other	-0.370	-0.207	-0.369		-0.473	-0.771	-1.087	-0.082	-0.460
	(0.061)	(0.087)	(0.061)		(0.040)	(0.010)	(0.012)	(0.007)	(0.005)
ρ	0.733	0.723	0.733	0.734		1.987			
	(0.020)	(0.025)	(0.020)	(0.017)		(0.023)			
γ_2	-3.275	-3.260	-3.272	-3.211	-3.487	-2.506	-1.992		
	(0.054)	(0.068)	(0.053)	(0.048)	(0.015)	(0.007)	(0.006)		
γ_4	3.326	3.356	3.329	3.321	3.997	3.089	2.603		
	(0.055)	(0.076)	(0.055)	(0.054)	(0.024)	(0.006)	(0.006)		

Table 4: Main results.

Interestingly, Johnston et al. (2009) found no link between self-reported hypertension and income; an association that, however, became positive when using objective measures of hypertension.

The estimated effect of income also changes when we control for state dependence. Comparing columns (f) and (g), we see that controlling for state dependence in a model without unobserved heterogeneity reduces the association between income and self-reported health. So, individually controlling for unobserved heterogeneity or for dynamics reduces the magnitude of the association between health and income.

Finally, a comparison between columns (a) and (d) shows that the estimate for income association is robust to controlling for employment status,

State dependence. We estimate an autoregressive parameter of around 0.75, with threshold differences of about 3. The estimated ratio of ρ to the thresholds (which measure the distance from category 3) are much lower than for column (f). This confirms the importance of controlling for unobserved heterogeneity, which reduces the estimated magnitude of persistence by a factor 3. Nevertheless, even when controlling for unobserved (and observed) heterogeneity, we find strong evidence for large, positive persistence in self-reported health.

There are at least two ways to get a sense of the magnitude of persistence. The first approach, also available for binary choice methods, is to compare estimates of ρ to estimates of regression coefficients. For example, in our preferred specification in column (a), a health shock that lifts you from any category below 3, to category 3 or 4, has an impact on future health that is almost 4 times that of becoming unemployed. The impact is more than 5 times that of marrying.

The second approach to interpreting estimates of ρ uses the estimated thresholds to obtain an estimate similar to a linear autoregressive model.[9] Differences between the thresholds are a measure of the distances between two categories. If $\gamma_2 = -\gamma_4$, then categories 2 and 3 are as far apart as categories 3 and 4. In such a case, a linear model may yield similar results in terms of partial effects. In this case, $-\rho/\gamma_2$ and ρ/γ_3 can be interpreted as linear regression coefficients for

[9] This approach is not available for binary choice models because threshold parameters are not available.

that category; we find that they are about 0.25. Said differently, we find the analog of an AR(1) coefficient of 0.25 in a linear model.

Other time-varying covariates. The literature so far has been inconclusive on how retirement is associated with health. On one hand, retiring allows more time for health-promoting activities, and reduces work-related stress. On the other hand, people may lose traction and motivation and may become less active. Therefore, while Coe and Zamarro (2011) find that retirement improves health, Behncke (2012) finds an increase in the likelihood of disease following retirement. In our DOLFE model, the coefficient is statistically insignificant. This suggests that the association between retirement and health may not be as strong as previously thought. Compared to the FEOL model, the effect of retirement on health disappears when controlling for state dependence.

The extensive literature on the link between unemployment and health in particular, and economic conditions and health more generally, is broadly inconclusive. Some studies have suggested a protective role of unemployment on health (Ruhm, 2000), while others suggest that unemployment is detrimental for health (McInerney and Mellor, 2012). Our results appear to be more in line with the findings of Ruhm (2015) and Böckerman and Ilmakunnas (2009). In DOLFE, the coefficient of being unemployed is negative and statistically significant. Controlling for state dependence does not change things compared to the FEOL model.

Having children is insignificant in our DOLFE model, while previous studies have provided mixed findings on this question (Mckenzie and Carter, 2013; Evenson and Simon, 2005). This is also insignificant in the FEOL model, suggesting that previous findings on having children might have been driven by unobserved heterogeneity.

Being married is generally considered a protective factor for health (Kaplan and Kronick, 2006; Molloy et al., 2009). In our model, however, it is statistically insignificant – as opposed to the FEOL model where it was positive and significant. Thus, controlling for state dependence appears to be important for this variable.

6 Conclusion

This paper studies a fixed$-T$ dynamic ordered logit model with fixed effects (DOLFE) and is the first to provide identification and estimation results for all common parameters in a dynamic ordered logit model with fixed effects and a fixed number of time periods. The results require only four time periods of data on the ordinal outcome variable. We demonstrate identification of the autoregressive coefficients on the lagged dependent variable, the regression coefficients on the exogenous regressors, and differences of the threshold parameters. The latter makes it possible to interpret the magnitude of the coefficients.

Including fixed effects and state dependence in the model is particularly relevant for self-reported health, a measure that is widely used in the literature. Future research using self-reporting health can benefit from our model for two main reasons. First, controlling for fixed effects, one can take into account unobserved heterogeneity (Carro and Traferri, 2014; Halliday, 2008; Fernández-Val et al., 2017), which is especially important due to differences in understanding and reporting health status (Groot, 2000; Sen, 2002; Jylhä et al., 1998; Baron-Epel et al., 2005; Jürges, 2007). Second, it incorporates elements of persistence (Contoyannis et al., 2004; Ohrnberger et al., 2017; Hernández-Quevedo et al., 2008; Roy and Schurer, 2013) or adaptation (Cubí-Mollá et al., 2017; Daltroy et al., 1999; Damschroder et al., 2005; Heiss et al., 2014) by controlling for state dependence (Carro and Traferri, 2014; Fernández-Val et al., 2017; Halliday, 2008). Thus, using our estimator addresses such biases often present in studies using self-rated health (Davillas et al., 2017).

We thus applied the new dynamic ordered logit model with fixed effects to investigate the determinants of self-reported health, focusing on the link between income and health in a panel of European countries. We found that when controlling for unobserved heterogeneity, the association between income and health becomes statistically insignificant. This is in line with studies that have found a smaller or insignificant association when using fixed effects (Gunasekara, 2011; Larrimore, 2011) – while other studies have suggested a positive association between the two (Carrieri and Jones, 2017; Ettner, 1996; Frijters et al., 2005; Mackenbach et al., 2005). Being retired or married also becomes statistically insignificant in our model when controlling for state dependence. Being unemployed or having children does not appear to be associated with self-reported health in our model.

Our empirical results suggest that persistence plays a positive and significant role in one's self-reported health. In other words, one's health is dependent on the health in the previous period, which is a reasonable thing to expect, as health problems may expand over a number of periods, or become permanent. This element reflects persistence of health status over time (Contoyannis et al., 2004). Furthermore, what is particularly interesting is that in our data, self-reported health tends to improve, on average, over time - even as people become four years older during the study period - and being older is typically associated with worse health outcomes. Therefore, it is reasonable to believe that this improvement in self-reported health is often subjective, and does not necessarily reflect one's objective health level. This element might reflect adaptation to health problems: Even though one's health does not improve, they adapt to their situation and therefore report better health (Cubí-Mollá et al 2017; Daltroy et al., 1999; Damschroder et al., 2005). This second element is a typical bias when using self-reported health outcomes, and our model helps correct such biases by introducing the dynamic element to a fixed effects ordered model. Another interesting finding is that, when controlling for unobserved heterogeneity, the link between income and health becomes statistically insignificant, suggesting that other factors might explain the association between the two.

Overall, measurement bias in studies using self-reported outcomes often poses challenges to research, that may discourage the use of such variables. Our model addresses these biases, and thus provides a basis for more choices in conducting research with databases that provide such variables.

References

[1] Albarran, P., R. Carrasco, and J.M. Carro, 2019. Estimation of Dynamic Nonlinear Random Effects Models with Unbalanced Panels. Oxford Bulletin of Economics and Statistics 81 (6), 1424–1441.

[2] Aristodemou, E., 2020. Semiparametric identification in panel data discrete response models. Journal of Econometrics, forthcoming.

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

[4] Banerjee, A., A. Deaton, and E. Duflo, 2004. Wealth, health, and health services in rural Rajasthan. American Economic Review, 94 (2), 326–330.

[5] Baron-Epel, O., G. Kaplan, A. Haviv-Messika, J. Tarabeia, M.S. Green and D.N. Kaluski, 2005. Self-reported health as a cultural health determinant in Arab and Jewish Israelis: MABAT—National Health and Nutrition Survey 1999–2001. Social Science and Medicine, 61 (6), 1256–1266.

[6] Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M. and O'Donnell, O., 2008. Does reporting heterogeneity bias the measurement of health disparities?. Health Economics, 17(3), 351–375.

[7] Behncke, S., 2012. Does retirement trigger ill health? Health Economics, 21(3), 282–300.

[8] Böckerman, P. and P. Ilmakunnas, 2009. Unemployment and self-assessed health: evidence from panel data. Health Economics, 18 (2), 161–179.

[9] Bound, J. and T. Waidmann, 1992. Disability transfers, self-reported health, and the labor force attachment of older men: Evidence from the historical record. The Quarterly Journal of Economics, 107 (4), 1393–1419.

[10] Bradford, W.D. and P. Dolan, 2010. Getting used to it: the adaptive global utility model. Journal of Health Economics, 29 (6), 811–820.

[11] Butler, J.S., R.V. Burkhauser, J.M. Mitchell, and T.P. Pincus, 1987. Measurement error in self-reported health variables. Review of Economics and Statistics, 69(4), 644–650.

[12] Caroli, E. and Weber-Baghdiguian, L., 2016. Self-reported health and gender: The role of social norms. Social Science and Medicine, 153, 220–229.

[13] Carrieri, V. and Jones, A.M., 2017. The income–health relationship 'beyond the mean': New evidence from biomarkers. Health Economics, 26(7), 937–956.

[14] Carro, J.M., 2007. Estimating Dynamic Panel Data Discrete Choice Models with Fixed Effects. Journal of Econometrics, 140 (2), 503–528.

[15] Carro, J.M. and A. Traferri, 2014. State dependence and heterogeneity in health using a bias-corrected fixed-effects estimator. Journal of Applied Econometrics, 29 (2), 181–207.

[16] Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. The Review of Economic Studies, 47 (1), 225–238.

[17] Chamberlain, G., 1985. Heterogeneity, omitted variable bias, and duration dependence." In: "Longitudinal Analysis of Labor Market Data", edited by J. J. Heckman and B.S. Singer, 3–38. Econometric Society Monographs. Cambridge University Press.

[18] Clark, A.E., Frijters, P. and Shields, M.A., 2008. Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. Journal of Economic Literature, 46(1), 95–144.

[19] Clark, A., Y. Georgellis, and P. Sanfey, 2001. Scarring: The psychological impact of past unemployment. Economica, 68 (270), 221–241.

[20] Coe, N.B. and G. Zamarro, 2011. Retirement effects on health in Europe. Journal of Health Economics, 30 (1), 77–86.

[21] Contoyannis, P., A.M. Jones, and N. Rice, 2004. The dynamics of health in the British Household Panel Survey. Journal of Applied Econometrics, 19(4), 473–503.

[22] Cox, D. R., 1958. The regression analysis of binary sequences. Journal of the Royal Statistical Society, 20, 215–242.

[23] Cubí-Mollá, P., M. Jofre-Bonet, and V. Serra-Sastre, 2017. Adaptation to health states: Sick yet better off?. Health Economics, 26 (12), 1826–1843.

[24] Daltroy, L.H., M.G. Larson, H.M. Eaton, C.B. Phillips, and M.H. Liang, 1999. Discrepancies between self-reported and observed physical function in the elderly: the influence of response shift and other factors. Social Science and Medicine, 48 (11), 1549–1561.

[25] Damschroder, L.J., B.J. Zikmund-Fisher, and P.A. Ubel, 2005. The impact of considering adaptation in health state valuation. Social Science and Medicine, 61 (2), 267–277.

[26] Das, M., and A. van Soest, 1999. A panel data model for subjective information on household income growth. Journal of Economic Behavior and Organization, 40 (4), 409–426.

[27] Davillas, A., A.M. Jones, and M. Benzeval, 2017. The income-health gradient: Evidence from self-reported health and biomarkers using longitudinal data on income. ISER Working Paper Series 2017-03.

[28] Ettner, S.L., 1996. New evidence on the relationship between income and health. Journal of Health Economics, 15(1), 67–85.

[29] Eurostat (2017). Methodological Guidelines and Description of EU-SILC Target Variables. 2016 operation.

[30] Evenson, R.J. and R.W. Simon, 2005. Clarifying the relationship between parenthood and depression. Journal of Health and Social Behavior, 46(4), 341–358.

[31] Fernández-Val, I., Y. Savchenko, and F. Vella, 2017. Evaluating the Role of Income, State Dependence and Individual Specific Heterogeneity in the Determination of Subjective Health Assessments. Economics and Human Biology, 25, 85–98.

[32] Frijters, P., Haisken-DeNew, J.P. and Shields, M.A., 2005. The causal effect of income on health: Evidence from German reunification. Journal of Health Economics, 24(5), 997–1017.

[33] Gravelle, H. and M. Sutton, 2009. Income, relative income, and self-reported health in Britain 1979–2000. Health Economics, 18 (2), 125–145.

[34] Groot, W., 2000. Adaptation and scale of reference bias in self-assessments of quality of life. Journal of Health Economics, 19 (3), 403–420.

[35] Gunasekara, F., K. Carter, and T. Blakely, 2011. Change in income and change in self-rated health: Systematic review of studies using repeated measures to control for confounding bias. Social Science and Medicine, 72 (2), 193–201.

[36] Hahn, J., 2001. The Information Bound of a Dynamic Panel Logit Model with Fixed Effects. Econometric Theory 17 (5), 913–932.

[37] Halliday, T.J., 2008. Heterogeneity, state dependence and health. The Econometrics Journal, 11 (3), 499–516.

[38] Heckman, J.J., 1981. The incidental parameters problem and the problem of initial conditions in estimating a discrete time–discrete data stochastic process. In: Structural Analysis of Discrete Data with Econometric Applications, Manski CF, McFadden D (eds). MIT Press: Cambridge, MA, 179–195.

[39] Heiss, Florian, Steven F Venti, and David A Wise. The Persistence and Heterogeneity of Health among Older Americans, NBER Working Paper 20306.

[40] Hernández-Quevedo, C., A.M. Jones, and N. Rice, 2008. Persistence in health limitations: A European comparative analysis. Journal of Health Economics, 27 (6), 1472–88.

[41] Honoré, B. E., and E. Kyriazidou, 2000. Panel Data Discrete Choice Models with Lagged Dependent Variables. Econometrica, 68 (4), 839–874.

[42] Honoré, B. E., and E. Kyriazidou, 2019. Identification in Binary Response Panel Data Models: Is Point-Identification More Common Than We Thought? Annals of Economics and Statistics, 134, 207–226.

[43] Honoré, B. E., and Martin Weidner, 2020. Moment Conditions for Dynamic Panel Logit Models with Fixed Effects. ArXiv:2005.05942.

[44] Johnston, D.W., C. Propper, and M.A. Shields, 2009. Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. Journal of Health Economics, 28 (3), 540–552.

[45] Jürges, H., 2007. True health vs response styles: exploring cross-country differences in self-reported health. Health Economics, 16 (2), 163–178.

[46] Jylhä, M., J.M. Guralnik, L. Ferrucci, J. Jokela, and E. Heikkinen, 1998. Is self-rated health comparable across cultures and genders? The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 53 (3), S144–S152.

[47] Kaplan, R.M. and R.G. Kronick, 2006. Marital status and longevity in the United States population. Journal of Epidemiology and Community Health, 60 (9), 760–765.

[48] Kapteyn, A., J.P. Smith, and A. van Soest, 2007. Vignettes and self-reports of work disability in the United States and the Netherlands. American Economic Review, 97 (1), 461–473.

[49] King, G., C.J. Murray, J.A. Salomon, and A. Tandon, 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. American Political Science Review, 98 (1), 191–207.

[50] Khan, S., M. Ponomareva, and E. Tamer, 2020. Identification of dynamic binary response models. Working paper.

[51] Larrimore, J., 2011. Does a higher income have positive health effects? Using the earned income tax credit to explore the income-health gradient. The Milbank Quarterly, 89(4), 694–727.

[52] Layard, R., 2006. Happiness and public policy: A challenge to the profession. The Economic Journal, 116 (510), C24–C33.

[53] Lindeboom, M. and E. van Doorslaer, 2004. Cut-point shift and index shift in self-reported health. Journal of Health Economics, 23 (6), 1083–1099.

[54] Mackenbach, J.P., Martikainen, P., Looman, C.W., Dalstra, J.A., Kunst, A.E. and Lahelma, E., 2005. The shape of the relationship between income and self-assessed health: an international study. International Journal of Epidemiology, 34(2), 286–293.

[55] Magnac, T., 2000. Subsidised Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories. The Economic Journal, 110 (466), 805-837.

[56] McInerney, M. and J.M. Mellor, 2012. Recessions and seniors' health, health behaviors, and healthcare use: Analysis of the Medicare Current Beneficiary Survey. Journal of Health Economics, 31 (5), 744–751.

[57] Mckenzie, S.K., and K. Carter, 2013. Does transition into parenthood lead to changes in mental health? Findings from three waves of a population based panel study. Journal of Epidemiological Community Health, 67 (4), 339–345.

[58] Molloy, G.J., E. Stamatakis, G. Randall, and M. Hamer, 2009. Marital status, gender and cardiovascular mortality: behavioural, psychological distress and metabolic explanations. Social Science and Medicine, 69 (2), 223–228.

[59] Muris, C., 2017. Estimation in the fixed-effects ordered logit model. Review of Economics and Statistics, 99 (3), 465–477.

[60] Ohrnberger, J., E. Fichera, and M. Sutton, 2017. The dynamics of physical and mental health in the older population. The Journal of the Economics of Ageing, 9, 52–62.

[61] R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

[62] Roy, J., and S. Schurer, 2013. Getting stuck in the blues: Persistence of mental health problems in Australia. Health Economics, 22 (9), 1139–1157.

[63] Ruhm, C.J., 2000. Are recessions good for your health? The Quarterly Journal of Economics, 115 (2), 617–650.

[64] Ruhm, C.J., 2015. Recessions, healthy no more? Journal of Health Economics, 42, 17–28.

[65] Salomon, J.A., A. Tandon, and C.J. Murray, 2004. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. BMJ, 328 (7434), 258.

[66] Sen, A., 2002. Health: perception versus observation. BMJ, 324 (7342), 860–861.

[67] Shi, X., M. Shum, and W. Song, 2018. Estimating semi-parametric panel multinomial choice models using cyclic monotonicity. Econometrica, 86 (2), 737–761.

[68] Wilson, T.D. and D.T. Gilbert, 2008. Explaining away: A model of affective adaptation. Perspectives on Psychological Science, 3 (5), 370–386.

A Proofs

Proof of Theorem 1. Recall the definition of the events in the main text, for $2 \leq j \leq k \leq l \leq J$,

$$\begin{aligned}
A_{j,l} &= \left\{D_{i,0}(k) = d_0, D_{i,1}(k) = 0, D_{i,2}(l) = 1, D_{i,3}(j) = d_3\right\} \\
B_{j,l} &= \left\{D_{i,0}(k) = d_0, D_{i,1}(k) = 1, D_{i,2}(j) = 0, D_{i,3}(l) = d_3\right\}, \\
C_{j,l} &= A_{j,l} \cup B_{j,l}.
\end{aligned}$$

The following derivations modifies the development in HK (p. 843-844). The probability of the event $B_{j,l}$, conditional on the covariates X_i and the unobserved heterogeneity α_i is given by

$$
\begin{aligned}
P\left(B_{j,l}\middle|X_i,\alpha_i\right) = {} & p_0\left(X_i,\alpha_i\right)^{d_0}\left[1-p_0\left(X_i,\alpha_i\right)\right]^{1-d_0} \\
& \times \frac{\exp\left(\alpha_i+X_{i,1}\beta+\rho d_0\right)}{1+\exp\left(\alpha_i+X_{i,1}\beta+\rho d_0\right)} \\
& \times \frac{1}{1+\exp\left(\alpha_i+X_{i,2}\beta+\rho-\gamma_j\right)} \\
& \times \frac{\left[\exp\left(\alpha_i+X_{i,3}\beta-\gamma_l\right)\right]^{d_3}}{1+\exp\left(\alpha_i+X_{i,3}\beta-\gamma_l\right)}.
\end{aligned} \tag{11}
$$

Similarly, for $A_{j,l}$,

$$
\begin{aligned}
P\left(A_{j,l}\middle|X_i,\alpha_i\right) = {} & p_0\left(X_i,\alpha_i\right)^{d_0}\left[1-p_0\left(X_i,\alpha_i\right)\right]^{1-d_0} \\
& \times \frac{1}{1+\exp\left(\alpha_i+X_{i,1}\beta+\rho d_0\right)} \\
& \times \frac{\exp\left(\alpha_i+X_{i,2}\beta-\gamma_l\right)}{1+\exp\left(\alpha_i+X_{i,2}\beta-\gamma_l\right)} \\
& \times \frac{\left[\exp\left(\alpha_i+X_{i,3}\beta+\rho-\gamma_j\right)\right]^{d_3}}{1+\exp\left(\alpha_i+X_{i,3}\beta+\rho-\gamma_j\right)}.
\end{aligned} \tag{12}
$$

The probability of event $A_{j,l}$, conditional on the event $C_{j,l}=A_{j,l}\cup B_{j,l}$ and on $X_{i,2}=X_{i,3}$ is given by

$$
\begin{aligned}
P\left(A_{j,l}\middle|X_i,\alpha_i,C_{j,l},X_{i,2}=X_{i,3}\right) &= \frac{P\left(A_{j,l},A_{j,l}\cup B_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)}{P\left(A_{j,l}\cup B_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)} \\
&= \frac{P\left(A_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)}{P\left(A_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)+P\left(B_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)} \\
&= \frac{1}{1+P\left(B_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)/P\left(A_{j,l}\middle|X_i,\alpha_i,X_{i,2}=X_{i,3}\right)}.
\end{aligned} \tag{13}
$$

where the first step follows from the definition of conditional probability; the second follows from the fact that $A_{j,l}$ and $B_{j,l}$ are disjoint; the third from division by the probability of $A_{j,l}$. Plugging the conditional probabilities (11) and (12) into the final expression (13) in the display below obtains our final sufficiency result:

$$P\left(A_{j,l} \middle| X_i, C_{j,l}, X_{i,2} = X_{i,3}\right) = \frac{1}{1+\exp\left(\Delta X_i \beta + \rho\left(d_0 - d_3\right) + \left(1-d_3\right)\gamma_l + d_3 \gamma_j\right)}, \tag{14}$$

$$P\left(B_{j,l} \middle| X_i, C_{j,l}, X_{i,2} = X_{i,3}\right) = \frac{\exp\left(\Delta X_i \beta + \rho\left(d_0 - d_3\right) + \left(1-d_3\right)\gamma_l + d_3 \gamma_j\right)}{1+\exp\left(\Delta X_i \beta + \rho\left(d_0 - d_3\right) + \left(1-d_3\right)\gamma_l + d_3 \gamma_j\right)}. \tag{15}$$

□

Proof of Theorem 2. For notational convenience, we refer to the conditional probability obtained in our sufficiency result, Theorem 1, equation (15), as

$$\begin{aligned} p_{jl}\left(X_i, d_0, d_3\right) &= P\left(B_{j,l} \middle| X_i, C_{j,l}, X_{i,2} = X_{i,3}\right) \\ &= \Lambda\left(\Delta X_i \beta + \rho\left(d_0 - d_3\right) + \left(1-d_3\right)\gamma_l + d_3 \gamma_j\right). \end{aligned} \tag{16}$$

Evaluate this for the case of $j = k = l$ and for $d_0 = d_3 = 0$,

$$p_{kk}\left(X_i, 0, 0\right) = \Lambda\left(X_i \beta\right),$$

which is a simplification of (16) because the second and fourth term are zero due to $d_0 = d_3 = 0$ and the third term is zero because of the choice of l and the scale normalization, $\gamma_l = \gamma_k = 0$. Then

$$\beta = E_{kk,00}\left[\Delta X_i' \Delta X_i\right]^{-1} E_{kk,00}\left[\Delta X_i' \Lambda^{-1}\left(p_{kk}\left(X_i, 0, 0\right)\right)\right],$$

where $E_{jl,00}$ is the expectation conditional on $\left(X_i, C_{j,l}, X_{i,2} = X_{i,3}\right)$ and using the cutoffs $j = l = k$ and starting and ending values $d_0 = d_3 = 0$. The invertibility of the first term is due to Assumption 1, and the second term is well-defined because p_{kk} is bounded away from 0 and 1 because of the logistic errors. This obtains identification of β.

Next, note that

$$p_{kk}\left(X_i, 1, 0\right) = \Lambda\left(\Delta X_i \beta + \rho\right).$$

From this we obtain

$$\rho = E_{kk,10}\left[\Lambda^{-1}\left(p_{kk}\left(X_i,1,0\right)\right) - \Delta X_i\beta\right],$$

where $E_{kk,10}$ now uses starting and ending values $d_0 = 1$ and $d_3 = 0$. This identifies the persistence parameter, since β was identified previously.

To identify the thresholds γ_l, $l > k$, consider that for all $l > k$:

$$p_{kl}\left(X_i,0,0\right) = \Lambda\left(\Delta X_i\beta + \gamma_l\right),$$
$$\gamma_l = E_{kl,00}\left[\Lambda^{-1}\left(p_{kl}\left(X_i,0,0\right)\right) - \Delta X_i\beta\right].$$

Finally, to identify the thresholds γ_j, $j < k$, consider that for all $j < k$,

$$p_{jk}\left(X_i,1,1\right) = \Lambda\left(\Delta X_i\beta + \gamma_j\right),$$
$$\gamma_j = E_{jk,11}\left[\Lambda^{-1}\left(p_{jk}\left(X_i,1,1\right)\right) - \Delta X_i\beta\right].$$

$\square$

Proof of Theorem 3. **Consistency.** We will use the fact that the objective function is concave (demonstrated in the next paragraph). We can therefore use Theorem 2.7 in Newey and McFadden. That condition (i, identification) holds is suggested by our identification result in Theorem 2. The information inequality and Assumption 1 ensure that identification is not lost when moving to the composite conditional likelihood function, see also the Hessian below. Condition (iii, pointwise convergence) follows from a law of large numbers for i.i.d. data.

To see that the objective function is concave, denote $Z_{ijl} = \left(\Delta X_i, D_{i,0} - D_{i,3,jl}, \left(1 - D_{i,3,jl}\right), D_{i,3,jl}\right)$ and $\theta_{jl} = \left(\beta, \rho, \gamma_l, \gamma_j\right)$, so that

$$l_{i,jl}\left(\theta_{jl}\right) = C_{i,jl}\left[D_{i,1}\ln\Lambda\left(Z_{i,jl}\theta_{jl}\right) + \left(1 - D_{i,1}\right)\ln\left[1 - \Lambda\left(Z_{i,jl}\theta_{jl}\right)\right]\right], \tag{17}$$

so that the score contribution is

$$s_{i,jl}\left(\theta_{jl}\right)=C_{i,jl}\left[D_{i,1}-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right]Z_{i,jl}^{'} \tag{18}$$

and the contribution to the Hessian is

$$H_{i,jl}\left(\theta_{jl}\right)=-C_{i,jl}\Lambda\left(Z_{i,jl}\theta_{jl}\right)\left(1-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right)Z_{i,jl}^{'}Z_{i,jl}. \tag{19}$$

It can be seen immediately from (19) that $l_{i,jl}\left(\theta_{jl}\right)$ is concave: $C_{i,jl}\in\{0,1\}$, $\Lambda\left(Z_{i,jl}\theta_{jl}\right)\left(1-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right)\in(0,1)$, and $Z_{i,jl}^{'}Z_{i,jl}$ is positive semi-definite. Because sums of concave functions are concave, the objective function

$$l_n(\beta,\rho,\gamma)=\sum_{2\le j\le k\le l}l_{n,jl}\left(\beta,\rho,\gamma_j,\gamma_l\right)$$

is concave. This completes the proof of concavity.

Asymptotic normality. To demonstrate asymptotic normality of the estimator, we will verify the conditions in Theorem 3.1 of NM94. Condition (i, interior) holds by construction. The fact that condition (ii, twice CD) holds can be seen by inspecting the expression of the second derivative in (19). Since the composite conditional likelihood function is a sum of functions of that form, it is also twice continuously differentiable. Condition (iii, CLT for score) holds because standard central limit theorems for i.i.d. data apply to (18). To see this, note that

$$\begin{aligned}
Var\left[s_{i,jl}\left(\theta_{jl,0}\right)\right]&=E\left[s_{i,jl}\left(\theta_{jl,0}\right)s_{i,jl}\left(\theta_{jl,0}\right)^{'}\right]\\
&=E\left[C_{i,jl}\left[D_{i,1}-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right]^2Z_{i,jl}^{'}Z_{i,jl}\right]\\
&=E\left[E\left[C_{i,jl}\left[D_{i,1}-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right]^2Z_{i,jl}^{'}Z_{i,jl}\right]\middle|Z_{i,jl},C_{i,jl}\right]\\
&=E\left[C_{i,jl}E\left[D_{i,1}-\Lambda\left(Z_{i,jl}\theta_{jl}\right)^2\middle|Z_{i,jl},C_{i,jl}\right]Z_{i,jl}^{'}Z_{i,jl}\right]\\
&=E\left[C_{i,jl}\Lambda\left(Z_{i,jl}\theta_{jl}\right)\left[1-\Lambda\left(Z_{i,jl}\theta_{jl}\right)\right]Z_{i,jl}^{'}Z_{i,jl}\right]\\
&\equiv\Sigma_{jl}.
\end{aligned} \tag{20}$$

The score for the composite likelihood function reuiqres more notation. First, note that the score in (20) is for the parameter θ_{jl}, and is therefore a matrix of dimensions at most $(K+1+2)\times(K+1+2)$ matrix. The score contribution for the composite likelihood is necessarily a $(K+1+(J-2))\times(K+1+(J-2))$ matrix, with rows and columns of zeros inserted into the location where parameters in θ are absent from θ_{jl} is called Ω_{jl}. Formally,

$$\tilde{s}_{i,jl}(\theta)=\frac{\partial l_{i,jl}(\theta_{jl})}{\partial\theta},$$
$$\Omega_{jl}\equiv E\left[\tilde{s}_{i,jl}(\theta_0)\,\tilde{s}_{i,jl}(\theta_0)'\right].$$

The variance of the score of the composite conditional log likelihood function is

$$\begin{aligned}\Omega &\equiv E\left[\left(\sum_{j,l}\tilde{s}_{i,jl}(\theta_0)\right)\left(\sum_{j,l}\tilde{s}_{i,jl}(\theta_0)\right)'\right]\\ &=\sum_{jl}\Omega_{jl}+\sum_{(j,l)\neq(j',l')}E\left[\tilde{s}_{i,jl}(\theta_0)\,\tilde{s}_{i,j'l'}(\theta_0)'\right].\end{aligned}\tag{21}$$

That the conditions for a CLT (cf. condition iii in NM94) are satisfied then follows from the boundedness of C and Λ, and Assumption 1.

Furthermore, note that the Hessian of the (j,l) contribution is given by

$$E\left[H_{i,jl}(\theta_{jl})\right]=-\Sigma_{jl},$$

which follows immediately from comparing (20) and (19). To obtain a Hessian for the composite likelihood, we enlarge the dimension of that Hessian by defining

$$\tilde{H}_{i,jl}(\theta)=\frac{\partial^2 l_{i,jl}(\theta_{jl})}{\partial\theta\partial\theta'}.$$

It follows that $E\left[\tilde{H}_{i,jl}\left(\theta_0\right)\right] = -\Omega_{jl}$ and

$$H = -\sum_{j,l} \Omega_{jl}. \tag{22}$$

Conditions (iv, v, Hessian) then follow from Assumption 1. All conditions in Theorem 3.1 of NM94 hold, and Theorem 3 therefore holds. $\square$

B Additional empirical results

We present a version of our main results in Table 4 that incldues the coefficients on the time-invariant variables, in Table 5. In the main text, we established that controlling for unobserved heterogeneity is important, so we should be careful in interpreting results from models without fixed effects (columns f, g, i). We find that men demonstrate higher levels of self-reported health (Caroli and Weber-Baghdiguian, 2016; Bago d'Uva et al., 2008), and those living in rural areas are also more likely to report better health (Lindeboom and van Doorslaer, 2004). Results also show that age is negatively related to self-reported health (Bago d'Uva et al., 2008; Lindeboom and van Doorslaer, 2004); and that individuals with higher levels of education report better health (Conti et al., 2010).

	(a) DOLFE $h=1$	(b) DOLFE $h=0.1$	(c) DOLFE $h=10$	(d) DOLFE $h=1$	(e) FEOL	(f) DOL	(g) OL	(h) FELM	(i) LM
log(income)	0.049	-0.047	0.059	0.061	0.020	0.340	0.492	0.003	0.194
	(0.033)	(0.056)	(0.029)	(0.029)	(0.019)	(0.004)	(0.004)	(0.003)	(0.002)
child	-0.030	0.006	-0.031	-0.026	0.021	0.060	0.089	0.002	0.033
	(0.051)	(0.069)	(0.050)	(0.049)	(0.032)	(0.005)	(0.005)	(0.005)	(0.002)
married	0.139	-0.041	0.157	0.130	0.164	0.073	0.141	0.029	0.062
	(0.087)	(0.119)	(0.086)	(0.088)	(0.053)	(0.007)	(0.008)	(0.009)	(0.003)
unemp	-0.188	-0.230	-0.178		-0.196	-0.242	-0.308	-0.033	-0.127
	(0.070)	(0.110)	(0.068)		(0.038)	(0.014)	(0.015)	(0.007)	(0.006)
retired	-0.132	-0.043	-0.139		-0.154	-0.050	-0.097	-0.027	-0.047
	(0.082)	(0.119)	(0.080)		(0.041)	(0.010)	(0.011)	(0.007)	(0.004)
other	-0.370	-0.207	-0.369		-0.473	-0.771	-1.087	-0.082	-0.460
	(0.061)	(0.087)	(0.061)		(0.040)	(0.010)	(0.012)	(0.007)	(0.005)
ρ	0.733	0.723	0.733	0.734		1.987			
	(0.020)	(0.025)	(0.020)	(0.017)		(0.023)			
γ_2	-3.275	-3.260	-3.272	-3.211	-3.487	-2.506	-1.992		
	(0.054)	(0.068)	(0.053)	(0.048)	(0.015)	(0.007)	(0.006)		
γ_4	3.326	3.356	3.329	3.321	3.997	3.089	2.603		
	(0.055)	(0.076)	(0.055)	(0.054)	(0.024)	(0.006)	(0.006)		
male						0.126	0.180		0.070
						(0.006)	(0.007)		(0.003)
urban_mid						-0.005	-0.018		-0.006
						(0.008)	(0.009)		(0.004)
urban_low						0.048	0.021		0.006
						(0.007)	(0.008)		(0.003)
age $]25;35]$						-0.613	-0.772		-0.282
						(0.015)	(0.016)		(0.005)
age $]35;45]$						-1.047	-1.349		-0.496
						(0.015)	(0.016)		(0.005)
age $]45;55]$						-1.423	-1.938		-0.722
						(0.015)	(0.015)		(0.005)
age $]55;65]$						-1.392	-2.016		-0.747
						(0.016)	(0.017)		(0.006)
age $]65;\infty[$						-1.549	-2.278		-0.854
						(0.017)	(0.019)		(0.007)
Primary schooling						0.385	0.510		0.204
						(0.028)	(0.031)		(0.013)
Lower secondary						0.516	0.737		0.302
						(0.029)	(0.031)		(0.013)
Upper secondary						0.694	0.952		0.384
						(0.028)	(0.031)		(0.012)
Post-secondary						0.717	0.931		0.377
						(0.031)	(0.034)		(0.014)
Tertiary						0.890	1.233		0.490
						(0.029)	(0.031)		(0.013)

Table 5: Main results.