

# feologit: A new command for fitting fixed-effects ordered logit models

Gregori Baetschmann  
University of Bern  
Bern, Switzerland  
gregori.baetschmann@soz.unibe.ch

Alexander Ballantyne  
University of Melbourne  
Melbourne, Australia  
ballantyne@student.unimelb.edu.au

Kevin E. Staub  
University of Melbourne  
Melbourne, Australia  
kevin.staub@unimelb.edu.au

Rainer Winkelmann  
University of Zurich  
Zurich, Switzerland  
rainer.winkelmann@econ.uzh.ch

**Abstract.** In this article, we describe how to fit panel-data ordered logit models with fixed effects using the new community-contributed command `feologit`. Fixed-effects models are increasingly popular for estimating causal effects in the social sciences because they flexibly control for unobserved time-invariant heterogeneity. The ordered logit model is the standard model for ordered dependent variables, and this command is the first in Stata specifically for this model with fixed effects. The command includes a choice between two estimators, the blow-up and cluster (BUC) estimator introduced in Baetschmann, Staub, and Winkelmann (2015, *Journal of the Royal Statistical Society, Series A* 178: 685–703) and the BUC- $\tau$  estimator in Baetschmann (2012, *Economics Letters* 115: 416–418). Baetschmann, Staub, and Winkelmann (2015) showed that the BUC estimator has good properties and is almost as efficient as more complex estimators such as generalized method-of-moments and empirical likelihood estimators. The command and model interpretations are illustrated with an analysis of the effect of parenthood on life satisfaction using data from the German Socio-Economic Panel.

**Keywords:** st0596, feologit, panel data, ordered dependent variables, logistic models, fixed effects, blow-up and cluster estimator

## 1 Introduction

While originating in the biometrics literature, regression models for ordered responses are now ubiquitous in the social sciences (Boes and Winkelmann 2006). One factor contributing to the widespread use of ordered responses is that Likert-type scales are the default way in which individual, household, and firm surveys collect information on issues that are otherwise difficult to measure, such as attitudes and beliefs. By far, the most common cross-sectional regression models for ordered responses are the ordered logit and ordered probit models. When analyzing ordinal panel data, researchers are frequently interested in applying extensions of these models that somehow account for the longitudinal nature of the data. The simplest approach, which we consider in this article, is specifying an additional unobservable individual-specific error term. Under

the assumptions that this error term is normally distributed and independent of the regressors, the models are known as the random-effects ordered logit or random-effects ordered probit (see, for example, Cameron and Trivedi [2005]), and they are implemented in Stata with the commands `xtologit` and `xtoprobit`, respectively.

Often, however, these distributional and independence assumptions on the individual-specific error term are undesirable. Fixed-effects models relax them: the distribution of the individual-specific error term and its dependence on the regressors are left completely unrestricted (compare, for example, Wooldridge [2010]). This feature of fixed-effects models is useful for the estimation of causal effects because it accounts for any potential endogeneity stemming from time-invariant characteristics. But because no default approach for fitting fixed-effects models for ordered responses exists, researchers are often faced with the choice of either fitting linear models, which are often inappropriate for ordinal data, or fitting random-effects ordered logit or probit models, which impose the strong assumptions mentioned above.

Because a fixed-effects estimator exists for the binary logit model, several different estimators for fixed-effects ordered logit models can be obtained using the binary logit model as a building block: the ordinal response variable can be transformed into binary responses, which then can be used for estimation and combined back differently to provide a single set of estimates. For the ordered probit model, in contrast, a similar approach is infeasible because no fixed-effects estimator for the binary probit model exists. Baetschmann, Staub, and Winkelmann (2015) studied several approaches available for the fixed-effects ordered logit model and showed that the so-called blow-up and cluster (BUC) estimator has good properties and is almost as efficient as more complex estimators such as generalized method-of-moments and empirical likelihood estimators. In this article, we discuss the BUC estimator as well as a more restricted version of it that makes it possible to fit additional model parameters—the BUC- $\tau$  estimator introduced in Baetschmann (2012)—and show how these estimators can be implemented in Stata using the community-contributed command `feologit`.

An integral part of our discussion focuses on the various potential objects of interest in this model, such as marginal effects (MEs), odds ratios, etc., and on whether they can or cannot be estimated, and if so, how. We further introduce a new specification test of the more restrictive assumptions relating to the additional threshold parameters estimated by BUC- $\tau$ . The test is simple to implement, and we show in an example using Stata syntax how BUC and BUC- $\tau$  estimates can be used to this end. Finally, we illustrate the use of `feologit` and the interpretation of the estimates in an application of the effect of motherhood on women's life satisfaction, which uses data from the German Socio-Economic Panel (SOEP).

We review the fixed-effects ordered logit model and the BUC and BUC- $\tau$  estimators in the next section. The syntax for `feologit` is presented in section 3.1. Next, section 4 gives a guide on how to interpret estimates. Section 5 exemplifies the use of `feologit` and possible interpretations of the estimates in the application to life satisfaction and provides a test of the BUC versus BUC- $\tau$  estimates. Section 5 offers some concluding remarks.

## 2 Fixed-effects ordered logit models

The fixed-effects ordered logit model uses the latent variable  $y^*$  to relate the observable characteristics  $\mathbf{x}$  to the observable ordered dependent variable  $y$ , which can take values  $1, \dots, K$ . The latent variable  $y_{it}^*$  for individual  $i$  at time  $t$  depends linearly on  $\mathbf{x}_{it}$  and the two unobservable characteristics  $\alpha_i$  and  $\varepsilon_{it}$ :

$$y_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T$$

The vector of covariates  $\mathbf{x}_{it}$  does not include an intercept because the  $\alpha_i$  act as individual-specific intercepts. We use a balanced panel for notational simplicity; but extending the model to imbalanced panels ( $t = 1, \dots, T_i$ ) is trivial, and the application in section 5 uses such an imbalanced panel. The time-invariant, individual-specific part of the unobservables ( $\alpha_i$ ) is called the fixed effect and can statistically depend on  $\mathbf{x}_{it}$ . The following observation rule ties the latent variable  $y_{it}^*$  to the observed ordered variable  $y_{it}$  through the thresholds  $\tau_{ik}$ :

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

In the most flexible version of the model, the thresholds can vary between individuals, as indicated by the subscript  $i$  in  $\tau_{ik}$ . Besides the stipulation of the lowest and highest thresholds as plus and minus infinity, the only assumption about the individual-specific thresholds is that they are increasing for each person:

$$\tau_{i1} = -\infty; \quad -\infty < \tau_{ik} < \tau_{ik+1} < \infty, \quad \forall k = 2, \dots, K-1; \quad \tau_{iK+1} = \infty \quad (1)$$

Moreover, the fixed-effects ordered logit model assumes that the time-varying unobservable terms,  $\varepsilon_{it}$ , are independent and identically distributed with standard logistic cumulative density function, hence the name of the model:

$$F(\varepsilon_{it}|\mathbf{x}_{it}, \alpha_i) = F(\varepsilon_{it}) = \frac{1}{1 + \exp(-\varepsilon_{it})} \equiv \Lambda(\varepsilon_{it})$$

The probability of observing outcome  $k$  for individual  $i$  at time  $t$  is therefore

$$\Pr(y_{it} = k|\mathbf{x}_{it}, \alpha_i) = \Lambda(\tau_{ik+1} - \mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_i) - \Lambda(\tau_{ik} - \mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_i) \quad (2)$$

This probability depends on  $\mathbf{x}_{it}$  and  $\boldsymbol{\beta}$ , the parameter of primary interest. However, it also depends on  $\alpha_i$ ,  $\tau_{ik}$ , and  $\tau_{ik+1}$ . As can be seen from (2), without further assumptions on the thresholds, only  $\tau_{ik} - \alpha_i \equiv \alpha_{ik}$  is identified because we can always define  $\tilde{\tau}_{ik} = \tau_{ik} + \eta$  and  $\tilde{\alpha}_i = \alpha_i - \eta$  for any  $\eta \in \mathbb{R}$ .<sup>1</sup>

1. On a more fundamental level, the parameters  $\boldsymbol{\beta}$  and  $\alpha_{ik}$  are identified only by the normalizing assumption that  $\varepsilon_{it}$  follows a standard (rather than any) logistic distribution, which fixes  $\varepsilon_{it}$ 's variance. Without this normalization, only  $\boldsymbol{\beta}/\text{var}(\varepsilon_{it})$  and  $\alpha_{ik}/\text{var}(\varepsilon_{it})$  are identified. Thus, identification and consistent estimation in this article refer to identification and consistency up to scale, as is the case for most standard latent-variable models, such as logit, probit, and their ordered and multinomial generalizations.

Direct estimation of  $\alpha_{ik}$  is difficult. Generally, estimation of  $\alpha_{ik}$  uses only information of the  $T$  observations of individual  $i$ . Thus, if the time dimension  $T$  is fixed—as is generally assumed in short panels—there are only a finite number of observations to estimate  $\alpha_{ik}$  even when the total number of observations  $NT$  grows to infinity. Consequently, the fixed effects  $\alpha_{ik}$  cannot be estimated consistently, and, in general, their inconsistency spills over to inconsistency of  $\beta$ , the parameters common to all observations. This situation is known as the incidental parameters problem (Neyman and Scott 1948; Lancaster 2000). In short panels, the resulting bias in  $\hat{\beta}$  can be substantial (Abrevaya 1997; Greene 2004). A consistent estimator of  $\beta$  can be obtained by collapsing  $y_{it}$  into a binary variable and then applying the well-known conditional maximum likelihood (CML) estimator (Andersen 1970; Chamberlain 1980).

## 2.1 CML estimator

The CML estimator is well known. But because the BUC estimator for the fixed-effects ordered logit model is based on the CML, we present it in some detail to fix notation. In Stata, this estimator is implemented in the command `clogit` and in the panel-data command `xtlogit` with the option `fe`, which relies on `clogit`. Similarly, `feologit` also relies on `clogit`.

Let  $d_{it}^k$  denote the binary variable that results from dichotomizing the ordered variable at the cutoff point  $k$ :  $d_{it}^k = \mathbb{1}(y_{it} \geq k)$ . This is the dependent variable of the CML estimator. Let

$$g_i^k = \sum_{t=1}^T d_{it}^k$$

be the observed number of ones of the dependent variable for individual  $i$ . Now, consider the probability of observing  $\mathbf{d}_i^k = (d_{i1}^k, \dots, d_{iT}^k)'$  conditional on observing  $g_i^k$  ones. It can be shown that this probability is

$$\mathcal{P}_i^k(\beta) \equiv \Pr \left( \mathbf{d}_i^k \mid \sum_{t=1}^T d_{it}^k = g_i^k \right) = \frac{\exp(\mathbf{d}_i^{k'} \mathbf{x}_i \beta)}{\sum_{\mathbf{j} \in \mathbf{B}_i} \exp(\mathbf{j}' \mathbf{x}_i \beta)} \quad (3)$$

where  $\mathbf{j}$  denotes a vector of dimension  $T$  with each element  $j_t$  equal to 0 or 1 and with  $\sum_{t=1}^T j_t = g_i^k$ . Further,  $\mathbf{B}_i$  denotes the set of all possible  $j$ -vectors with  $g_i^k$  ones and  $T - g_i^k$  zeros. There are  $\binom{T}{g_i^k} = T! / \{g_i^k! (T - g_i^k)!\}$  such combinations. Crucially, this probability (3) does not depend on  $\alpha_i$  and the thresholds. Chamberlain (1980) showed that maximizing the conditional log likelihood (LL)

$$\text{LL}^k(\mathbf{b}) = \sum_{i=1}^N \log \mathcal{P}_i^k(\mathbf{b}) \quad (4)$$

results in a consistent estimator for  $\beta$ . Therefore,  $\beta$  of the fixed-effects ordered logit model can be fit by first dichotomizing the ordered dependent variable into a binary one and then applying the standard CML estimator. However, different cutoff points

$k$  can be used, and using only one of them leads to loss of information and, therefore, to inefficiency. Further details of the CML estimator such as first-order conditions and asymptotic variance can be found, for example, in Baetschmann, Staub, and Winkelmann (2015).

## 2.2 BUC estimator

Several ideas exist to combine the information of the CML estimators obtained from dichotomizing samples at different cutoff points (see Baetschmann, Staub, and Winkelmann [2015]). The BUC estimator presented here combines the LL functions resulting from different cutoff points, leading to a one-step estimator of  $\beta$ . The LL function for this estimator is

$$\text{LL}^{\text{BUC}}(\mathbf{b}) = \sum_{k=2}^K \text{LL}^k(\mathbf{b}) \quad (5)$$

where  $\text{LL}^k(\mathbf{b})$  is defined as in (4) and the BUC estimator is the one that maximizes (5). It can also be regarded as a restricted CML estimator because it imposes the restriction that  $\hat{\beta}^2 = \dots = \hat{\beta}^K$ . We call this the BUC estimator because this describes how the estimator is implemented: first, every individual's observations in the sample are replaced with  $K-1$  copies or clones of itself ("blow up" the sample size); and, then, each clone is dichotomized at a different cutoff point. We then use the entire inflated sample to estimate  $\beta$  by applying the CML estimator. Because the clones of the same individual are not independent of each other, we have to compute standard errors that are clustered at the individual level. Baetschmann, Staub, and Winkelmann (2015) have shown that combining the likelihoods leads to a large efficiency gain compared with using only one cutoff. In addition, the BUC estimator has less convergence problems compared with efficient estimators like a two-step generalized method-of-moments estimator, and the efficiency loss in finite samples is negligible.

## 2.3 BUC- $\tau$ estimator assuming constant thresholds

The standard ordered logit model for cross-sectional data assumes that the thresholds are constant across individuals. The BUC estimator is conformable with a more general class of models because it is also consistent with models where each individual has different thresholds. If we are willing to make the additional assumption of constant thresholds, Baetschmann (2012) suggested a procedure based on the BUC estimator that allows us to estimate the thresholds, too. We will call this estimator BUC- $\tau$ . The additional assumption for the more restrictive model is that  $\tau_{ik} = \tau_{jk} = \tau_k$  for all individuals  $i$  and  $j$ . Because the rest of the model is unchanged, the probability of observing outcome  $k$  for individual  $i$  at time  $t$  is

$$\Pr(y_{it} = k | \mathbf{x}_{it}, \alpha_i) = \Lambda(\tau_{k+1} - \mathbf{x}'_{it}\beta - \alpha_i) - \Lambda(\tau_k - \mathbf{x}'_{it}\beta - \alpha_i)$$

As in the more flexible model described above, we cannot distinguish between terms that are constant within an individual. So if all thresholds ( $\tau$ ) increase by the same

amount as the individual fixed effect ( $\alpha$ ), the same probability results. We deal with this underidentification by restricting the second threshold to 0. Assumption (1) of the model therefore changes to

$$\tau_1 = -\infty; \quad \tau_2 = 0; \quad 0 < \tau_k < \tau_{k+1} < \infty, \quad k = 3, \dots, K-1; \quad \tau_{K+1} = \infty$$

Without the restriction  $\tau_2 = 0$ , only the differences between the thresholds are identified.

The idea of the BUC- $\tau$  estimator is to dichotomize the observations within a person at different cutoff points and then to apply the standard CML estimator. This allows us to estimate the thresholds, too. Let  $\mathbf{d}_i$  denote the resulting vector of the dichotomized dependent variable for individual  $i$  and  $g_i$  the number of ones in  $\mathbf{d}_i$ . In addition, define  $\boldsymbol{\tau}_i^{\text{cut}}$  as the vector of thresholds used as cutoff points for person  $i$ . The conditional probability of observing  $\mathbf{d}_i$  conditional on  $g_i$ , when  $\mathbf{d}_i$  results from dichotomizing at different cutoff points, is

$$\Pr \left( \mathbf{d}_i \left| \sum_{t=1}^T d_{it} = g_i \right. \right) = \frac{\exp\{\mathbf{d}_i'(\mathbf{x}_i\boldsymbol{\beta} - \boldsymbol{\tau}_i^{\text{cut}})\}}{\sum_{\mathbf{j} \in \mathbf{B}_i} \exp\{\mathbf{j}'(\mathbf{x}_i\boldsymbol{\beta} - \boldsymbol{\tau}_i^{\text{cut}})\}} \quad (6)$$

where  $\mathbf{j}$  denotes again a vector with zeros and ones with  $\sum_{t=1}^T d_{it} = g_i$  and  $\mathbf{B}_i$  the set of all possible  $\mathbf{j}$ -vectors with  $g_i$  ones and  $T - g_i$  zeros.

As an example, consider a person who is observed for two time periods. For the BUC estimator, we would produce  $K - 1$  copies of this person's observations—that is,  $K - 1$  clones of the person—and dichotomize each clone at a different cutoff point. Thus, one of these clones, say,  $i$ , would be dichotomized at the cutoff point 3, resulting in  $\mathbf{d}_i = \{\mathbb{1}(y_{i1} \geq 3), \mathbb{1}(y_{i2} \geq 3)\}'$  with corresponding  $\boldsymbol{\tau}_i^{\text{cut}} = (\tau_3, \tau_3)'$ . The next clone,  $j$ , dichotomized at 4, would result in  $\mathbf{d}_j = \{\mathbb{1}(y_{j1} \geq 4), \mathbb{1}(y_{j2} \geq 4)\}'$  with corresponding  $\boldsymbol{\tau}_j^{\text{cut}} = (\tau_4, \tau_4)'$ .

In contrast, for the BUC- $\tau$  estimator, the first observation of clone  $i$  might be dichotomized at the cutoff point 3 and the second observation at the cutoff point 4, resulting in the vectors  $\mathbf{d}_i = \{\mathbb{1}(y_{i1} \geq 3), \mathbb{1}(y_{i2} \geq 4)\}'$  and  $\boldsymbol{\tau}_i^{\text{cut}} = (\tau_3, \tau_4)'$ . Thanks to this heterogeneity in the cutoff point within a conditional likelihood contribution (that is, within a clone), the expression depends also on the thresholds, as can be seen in (6).

In the case of the BUC estimator, we combine the  $(K - 1)$  possible clones of each person to estimate  $\boldsymbol{\beta}$ . However, if different cutoff points within a clone are allowed as with BUC- $\tau$ , the number of possible clones of each person is  $(K - 1)^T$ , and the sample size of the inflated dataset would be  $N(K - 1)^T$ . In standard applications, this would result in more observations than most of today's computers can handle. Therefore, not all possible clones can be included in the inflated estimation sample, and a selection has to be made. We propose to include all clones with no variation in the cutoff point (that is, the clones corresponding to the sample used by the BUC estimator) and use a limited number of clones with random variation in the cutoff points. The program **feologit**, **threshold** is implemented accordingly, where the default is to include 10 clones of each individual with randomly chosen cutoffs. The user can change the number of additional clones by using the option **clones()**. The process of randomly selecting cutoff points

can be influenced by the option `seed(##)`. With replicability of results in mind, the `feologit` command has been programmed so that running the BUC- $\tau$  estimator twice leads to the same results. This should not deflect from the fact that cutoff points are selected randomly.

## 3 The feologit estimation command

### 3.1 Syntax

The command `feologit` is called with the following syntax:

```
feologit depvar indepvars [if] [in] [weight], group(varname) [thresholds
    clones(##) keepsample seed(##) cluster(clustvar) or otheropts]
```

where *depvar* is an ordered categorical variable. Time-series operators are not allowed. `fweights`, `iweights`, and `pweights` are allowed (see [U] 11.1.6 **weight**), but they are interpreted to apply to groups as a whole, not to individual observations.

### 3.2 Description

`feologit` fits fixed-effects ordered logit models using the BUC estimator of Baetschmann, Staub, and Winkelmann (2015). It does so by replacing each observation in the dataset by  $K - 1$  copies of the observation (where  $K$  is the number of categories of the ordered dependent variable) and then applying the CML estimator `clogit`, clustering the standard errors at the level of the original panel unit. After estimation, the dataset is returned to its original form (unless the option `keepsample` is specified). With the option `threshold`, `feologit` applies the BUC- $\tau$  estimator of Baetschmann (2012), which assumes that thresholds are constant across panel units.

### 3.3 Options

`group(varname)` is required if `xtset panelvar` has not been specified; it specifies an identifier variable (numeric or string) for the matched groups. If a panel identifier has been set with `xtset`, the option `group(varname)` may be omitted; in this case, `feologit` will use the panel identifier and provide a warning. `strata(varname)` is a synonym for `group()`.

`thresholds` calls the BUC- $\tau$  estimator, which includes estimates of the thresholds.

`clones(##)` specifies the number of clones used in the estimation when `thresholds` has been specified. The default is `clones(10)`. A clone is a copy of all observations of a panel unit.

**keepsample** specifies that the estimation sample be kept. The estimation sample includes the original data as well as additional observations consisting of copies of the original data. The option **keepsample** generates the following new variables:

**dkdepvar**, the dichotomized dependent variable used in the **clogit** estimation step;

**dkthreshold**, a variable that indicates at which cutoff point each observation of the ordered dependent variable was dichotomized (to result in **dkdepvar**);

**bucsample**, a binary variable that indicates whether the observation forms part of the estimation sample of the BUC estimator—this variable exhibits variation only if the option **thresholds** has been specified;

**clonegroup**, an integer-valued variable that identifies observations corresponding to each panel unit and clone in the estimation sample; and

**clone**, a binary variable that indicates whether an observation is part of the original sample (**clone** = 0) or a copy (**clone** = 1).

For instance, after BUC- $\tau$  estimation of **feologit** with the option **keepsample**, the corresponding BUC estimates can be obtained by issuing the following command:

```
clogit dkdepvar indepvars if bucsample==1, group(clonegroup) cluster(clustvar)
```

where *indepvars* and *clustvar* are the variables that were used in the BUC- $\tau$  estimation.

**seed(#)** specifies the pseudo-random-number seed used in the estimation when the option **thresholds** has been specified. The default is **seed(79846512)**.

**cluster(*clustvar*)** sets the identifier variable for clustering standard errors. Standard errors are always clustered; specifying this option overrides the default clustering variable, which is the group identifier.

**or** reports the estimated coefficients transformed to odds ratios, that is, **exp(b)** rather than **b**. Standard errors and confidence intervals are similarly transformed.

*otheropts*; see **help feologit**.



### 3.4 Stored results

Many of the results stored in `e()` are similar to `clogit` or `ologit`. Stored results specific to `feologit` are

#### Scalars

<code>e(N)</code>	number of observations including observations of clones
<code>e(N.true)</code>	number of authentic observations (without observations of clones)
<code>e(N.group)</code>	number of panel units specified by <code>group()</code>
<code>e(cut1)</code>	rescaling factor used for <code>predict</code> and <code>margins</code> postestimation commands (if option <code>thresholds</code> is specified; see sections 3.5 and 4.4 for details)
<code>e(clones)</code>	number of clones (additional to dichotomized sample)
<code>e(seed)</code>	seed for random-number generator

#### Macros

<code>e(estopt)</code>	estimation type option ( <code>basic</code> or <code>thresholds</code> )
<code>e(group)</code>	name of <code>group()</code> variable

### 3.5 Postestimation

The following postestimation commands are available after `feologit`:

- `logitmarg` calculates statistics of MES from sample averages. `logitmarg` uses a routine provided with the `feologit` installation. It uses sample averages to calculate MES (see section 4.3 for details). It provides standard errors using the Delta method. The following options allow users to modify the reported results:
  - `outcome(outcome)` displays estimated MES only for the category selected by `outcome`, which should be either one value of the dependent variable or an indicator of the ordered category (`#1`, `#2`, etc.).
  - `dydx(varlist)` displays estimated MES only for the variables listed by `varlist`.
  - `eretstore` stores estimates in `e()` instead of `r()`. Existing `e()` results will be lost.
- `predict` creates a new variable containing linear predictions (option `xb`) or predictions of probabilities (called using the same syntax as after `ologit`). Predictions of probabilities are available only after estimation with the option `thresholds`. Estimates of probabilities are calculated assuming all fixed effects are equal to `e(cut1)` (see section 4.4 for details).
- `margins` estimates margins of response for probabilities and linear predictions. Margins for probabilities are available only after estimation with the `thresholds` option. Margins of response for probabilities are calculated assuming a value of `e(cut1)` for all fixed effects (see section 4.4 for details).
- `test` and `testnl` conduct Wald tests of simple and composite linear hypotheses and tests of nonlinear hypotheses. These commands cannot be used on estimates of the second threshold, which is constrained ( $\tau_2 = 0$ ). The second threshold  $\tau_2$  is the first finite threshold and is called `/cut1` in the estimation output.

## 4 Interpretation

In empirical applications, the interest usually lies in the effect of the covariates  $\mathbf{x}$  on the dependent variable, and the interpretation of  $\beta$  is of primary interest. However, because the ordered logit model is a nonlinear model, this parameter does not reflect MES of  $\mathbf{x}$  on the ordered dependent variable  $y$ . There exist different possibilities of interpreting the regression results, some of which are discussed below.

### 4.1 Direction and compensating variation

The sign of  $\beta$  indicates the direction in which an increase of  $\mathbf{x}$  influences the cumulative distribution of the dependent variable. If  $\beta_l > 0$ , an increase of the regressor  $x_l$  will lead to an unambiguous decrease in the probability of the lowest category  $\Pr(y_{it} \geq 1 | \mathbf{x}_{it}, \alpha_i)$  and an increase in the probability of the highest category  $\Pr(y_{it} \geq K | \mathbf{x}_{it}, \alpha_i)$ . Moreover, the single crossing property of the ordered logit model implies that there will be exactly one change from the probabilities of lower categories, which decrease, to probabilities of higher categories, which increase (see Winkelmann and Boes [2010]). Without knowing the thresholds, one cannot determine at which category this switch from decrease to increase will take place.

The  $\beta$  can be interpreted as MES of  $\mathbf{x}$  on the latent variable  $y^*$ . Because the interest often lies on the ordered dependent variable  $y$  rather than the latent  $y^*$ , this interpretation is rarely used. Another simple interpretation of  $\beta$  with wider application is to compute the compensating variation between variables, for example, the change in two regressors such that the latent variable, and therefore the ordered dependent variable, remains unchanged. The compensating variation of two variables is given by the ratio of the corresponding  $\beta$ : an increase of  $x_l$  by 1 has the same effect as an increase of  $x_r$  by  $\beta_l/\beta_r$ .

### 4.2 Odds ratio

The effect size in logit models is often interpreted using odds, which refers to the ratio between the probability of a certain event and the complementary probability. In the case of ordered logit, the odds of individual  $i$  in period  $t$  having a  $y_{it}$  above category  $k$  relative to below or equal to  $k$  is

$$\text{Odds}(k, \mathbf{x}_{it}) \equiv \frac{\Pr(y_{it} > k | \mathbf{x}_{it})}{\Pr(y_{it} \leq k | \mathbf{x}_{it})} = \exp(\mathbf{x}'_{it}\beta - \tau_{ik})$$

The odds are independent of the fixed effects but still depend on the thresholds. However, the change in the odds if the  $l$ th regressor is modified solely depends on  $\beta$  and the shift of the regressor:

$$\text{Odds ratio}(k, \Delta \mathbf{x}_{itl}) \equiv \frac{\text{Odds}(k, \mathbf{x}_{it} + \Delta \mathbf{x}_{itl})}{\text{Odds}(k, \mathbf{x}_{it})} = \exp(\Delta \mathbf{x}'_{itl}\beta)$$

Therefore, an increase of  $x_l$  by 1 increases the odds ratio by  $\exp(\beta_l)$  for all categories except the first one, everything else being equal. Or in other words, a unit increase in  $x_l$  changes the odds by about  $\beta_l \times 100$  percent (for small  $\beta_l$ ) or by exactly  $\{\exp(\beta_l) - 1\} \times 100$  percent. The option `or` displays the results as  $\exp(\beta)$  as in the standard commands for logit models.

### 4.3 Marginal effects

In empirical applications, the interest often lies in the marginal probability effects, that is, the change in the probabilities of observing  $y_{it} = k$  if a covariate  $l$  is changed by a small amount:

$$\text{ME}_{itkl} \equiv \frac{\partial \Pr(y_{it} = k | \mathbf{x}_{it}, \alpha_i)}{\partial \mathbf{x}_{itl}} = \frac{\partial \Pr(y_{it} \leq k | \mathbf{x}_{it}, \alpha_i)}{\partial \mathbf{x}_{itl}} - \frac{\partial \Pr(y_{it} \leq k - 1 | \mathbf{x}_{it}, \alpha_i)}{\partial \mathbf{x}_{itl}}$$

Because the probabilities depend on the thresholds and the individual fixed effects, any marginal probability effects will generally also depend on these parameters. And because they are not estimated, estimates for MEs cannot generally be obtained either. For the ordered logit model, the ME has the specific form

$$\begin{aligned} \text{ME}_{itkl} &= [\Lambda(\tau_{ik+1} - \mathbf{x}_{it} - \alpha_i) \{1 - \Lambda(\tau_{ik+1} - \mathbf{x}_{it} - \alpha_i)\} \\ &\quad - \Lambda(\tau_{ik} - \mathbf{x}_{it} - \alpha_i) \{1 - \Lambda(\tau_{ik} - \mathbf{x}_{it} - \alpha_i)\}] \beta_l \\ &= [\Pr(y_{it} \leq k | \mathbf{x}_{it}, \alpha_i) \{1 - \Pr(y_{it} \leq k | \mathbf{x}_{it}, \alpha_i)\} \\ &\quad - \Pr(y_{it} \leq k - 1 | \mathbf{x}_{it}, \alpha_i) \{1 - \Pr(y_{it} \leq k - 1 | \mathbf{x}_{it}, \alpha_i)\}] \beta_l \end{aligned} \quad (7)$$

Thus, we see immediately that the relative size of the MEs of two covariates  $l$  and  $r$  is equal to the relative size of their  $\beta$  coefficients,

$$\text{ME}_{itkl} / \text{ME}_{itkr} = \beta_l / \beta_r$$

a quantity that can readily be estimated from  $\hat{\beta}_l / \hat{\beta}_r$ .

Moreover, from the second equality in (7), it is clear that for given probabilities of the dependent variable  $y_{it}$ , the ME is just a function of  $\beta_l$  and straightforward to calculate with an estimate  $\hat{\beta}_l$ . One can therefore calculate MEs for any interesting probabilities of the dependent variable. An obvious choice for such probabilities is the sample proportions. We called this the ME at the average, and it can be computed by the `feologit` postestimation command `logitmarg` for each regressor and each possible outcome category. Then, an estimate of the ME of regressor  $l$  for category  $k$  is

$$\overline{\text{ME}}_{kl} = \left\{ \bar{d}^{k+1} \left( 1 - \bar{d}^{k+1} \right) - \bar{d}^k \left( 1 - \bar{d}^k \right) \right\} \hat{\beta}_l$$

where  $\bar{d}^k$  is the sample average of  $d_{it}^k$ . Standard errors for  $\overline{\text{ME}}_{kl}$  can be obtained via the Delta method. Depending on whether one is interested in  $\overline{\text{ME}}_{kl}$  as an estimate of the ME at the sample average or as an estimate of the ME at the population mean, standard errors need to account for sampling variation from estimation of only  $\beta$  or, in addition,

from estimation in  $\bar{d}^{k+1}$  and  $\bar{d}^k$ . The command `logitmarg` provides standard errors for the ME at the sample average.

This is a simple and arguably useful object of interest. However, it is different from the average ME,  $E(\text{ME}_{itkl})$ , which is infeasible. It is also different from the ME at the average of the regressors, defined as (7) evaluated at  $\bar{\mathbf{x}} = \sum_i \sum_t \mathbf{x}_{it} / (NT)$ , which is also infeasible. Both of these more widely used objects of interest depend on the unavailable individual thresholds and fixed effects [see the expression after the first equality of (7)], while the ME at the average of the dependent variable that we propose circumvents this problem by focusing instead on the expression after the second equality of (7).

Finally, another potentially useful quantity that is identified and easily estimable is the average semielasticity of the continuation probability at category  $k$  with respect to regressor  $l$ . The identification and estimation of the average semielasticity in binary fixed-effects logit models was demonstrated by Kitazawa (2012) (see also Santos Silva and Kemp [2016]), and here we generalize it to the ordered case:

$$E\left(\frac{\partial \ln P(y_{it} \geq k | \mathbf{x}_{it}, \alpha_i)}{\partial \mathbf{x}_{itl}}\right) = \beta_l (1 - \bar{d}^k)$$

## 4.4 Thresholds

The assumption of constant thresholds for all individuals used by the BUC- $\tau$  estimator allows for additional interpretations. A model with constant thresholds can be fit with `feologit` by using the option `threshold`. For the interpretations that follow, we change our conceptual perspective: up to this point, all interpretations were cast in terms of probabilities. The  $\varepsilon_{it}$  was treated as a random variable. Now, we keep  $\varepsilon_{it}$  fixed, and a change of a regressor leads to a deterministic effect of either pushing the ordered variable to a different category or staying in the same category.

If the spaces between adjacent thresholds are almost equal, a change of the regressors has similar effects independently of the specific category. This does not apply, however, for the tails of the distribution, that is, the lowest and the highest category. If the spaces are unequal, a change of a regressor has a smaller effect for categories where the thresholds are far apart.

Regarding the effects of covariates, even a marginal change can lead to a switch of category because we do not know the exact value of the latent variable. However, because an estimate of the differences between the thresholds is available, we can also compute the change in the regressors, which surely leads to a switch of category. For example, a person in the third category will surely rise up to the next higher category if  $\mathbf{x}_{itl}$  increases by  $(\tau_3 - \tau_4)/\beta_l$ , everything else being equal.

When both the thresholds and  $\beta$  are known, the only unknown parameter in the formula for the ME (7) is  $\alpha_i$ . For such situations in fixed-effects models, Stata's `margins` postestimation command assumes  $\alpha_i = 0$  for all  $i$ . This is the case, for instance, when using `margins` after `xtlogit` with the option `fe` (the binary fixed-effects logit).

If  $\alpha_i \neq 0$  for all  $i$ , the object estimated assuming  $\alpha_i = 0$  for all  $i$  will, in general, not be consistent for the average ME. Nevertheless, we have equipped `feologit` with a similar postestimation capability. When `margins` is called after `feologit` with the option `thresholds`, an average ME is computed assuming that  $\alpha_i$  is constant across individuals. Because of the underidentification of  $\tau_{ik}$  and  $\alpha_i$ , and the normalization  $\tau_2 = 0$ , calculating probabilities at  $\alpha_i = 0$  for all  $i$  is often a particularly poor choice. Therefore, we use  $\alpha_i = \tilde{\alpha}$  for all  $i$  instead, which is defined as the estimate of the constant in a binary (cross-sectional) logit of the dichotomized dependent variable at the first cutoff ( $d_{it}^2$ ) with a single regressor  $z_{it} \equiv \mathbf{x}_{it}\hat{\beta}$  whose coefficient is restricted to 1:

$$\tilde{\alpha} = \max_a \sum_i \sum_t d_{it}^2 \log \Lambda(a + z_{it}) + (1 - d_{it}^2) \{1 - \Lambda(a + z_{it})\}$$

with first-order condition

$$\bar{d}^2 = (NT)^{-1} \sum_i \sum_t \Lambda(\tilde{\alpha} + z_{it})$$

The estimate  $\tilde{\alpha}$  is stored in `e(cut1)` after `feologit` with the option `thresholds`. A potentially better estimate for  $\tilde{\alpha}$  could be obtained by using all  $d^k$  instead of basing it only on  $d^2$ . However, our intention here is to provide such an estimate only as a suggestive result, and we caution against relying on these MEs (see also, for example, Santos Silva and Kemp [2016], who argue persuasively against using MEs based on  $\alpha_i = 0$ ).

## 5 Application: Effect of motherhood on life satisfaction

To illustrate the empirical application of the `feologit` command, we analyze the effect of the birth of the first child on his or her mother's life satisfaction using the SOEP, a large representative household survey (Wagner, Frick, and Schupp 2007). The data were collected yearly between 1984 and 2009, and it is therefore possible to follow a person up to 26 years. The sample includes women between the ages of 20 and 60 who either were mothers or became mothers during the observation period. The application is based on data from Baetschmann, Staub, and Studer (2016), where additional information about data and estimation of causal effects can be found.<sup>2</sup> The dependent variable is life satisfaction, an ordinal variable that ranges from 0 (completely dissatisfied) to 10 (completely satisfied). Below, we tabulate its distribution.

---

2. While all results shown below were obtained using the original SOEP data, the accompanying dataset (`bbsw.dta`) (Baetschmann, Staub, and Studer 2016) has been modified by artificially adding errors for anonymization purposes and reducing it slightly so that permission could be obtained for it to be shared publicly without restrictions. Consequently, results using the dataset differ somewhat from the ones presented here. For reference, the results obtained with the dataset can be obtained by running the `application_bbsw.do` file. SOEP data are publicly available—access can be requested through DIW Berlin.

```
. tabulate lifesat
```

life satisfaction		Freq.	Percent	Cum.
0. Completely dissatisfied	0	650	0.52	0.52
	1	498	0.40	0.92
	2	1,485	1.19	2.11
	3	3,332	2.67	4.79
	4	4,657	3.74	8.53
	5	16,464	13.22	21.74
	6	13,979	11.22	32.97
	7	26,793	21.51	54.48
	8	36,127	29.00	83.48
10. Completely satisfied	9	13,669	10.97	94.45
	10	6,908	5.55	100.00
Total		124,562	100.00	

About two-thirds of the responses lie in the upper part of the distribution (seven or higher). The modal answer is category 8 with a proportion of around 29%. We want to analyze the effect of the first child on a mother's life satisfaction and are especially interested in the evolution of a woman's general life satisfaction in the first years after the birth of her first child. Our specification also includes a small set of additional regressors: age, logarithm of household income, and a dummy indicating whether the respondent is working. While controlling for potentially endogenous factors such as income and labor-force participation can lead to biases in the estimates of the effects of interest, we include these variables here only to illustrate the regression output of the `feologit` command.

```
. describe lifesat kidage01 age lhinc work
```

variable name	storage type	display format	value label	variable label
lifesat	byte	%31.0g	p1110109	life satisfaction
kidage01	byte	%9.0g		age of first child
age	byte	%9.0g		age
lhinc	float	%9.0g		log of monthly household income
work	byte	%8.0g		worked last week

To estimate the effect of having the first child on life satisfaction and the following dynamics, we include five dummy variables, representing the age of the first child in years. For example, the variable `kidage01_2` is equal to 1 when the first child's age is 2 and 0 otherwise. Such a flexible form approach is common in the literature on adaptation to life events (for example, Clark et al. [2008]). Women who are happier might be more likely to get married and have children (Stutzer and Frey 2006). To control for this selection into motherhood, we want to control for time-invariant characteristics like a genetic disposition to happiness. And because the dependent variable is ordered, we use a fixed-effects ordered logit model fit with the BUC estimator. First, we declare the panel variable with the `xtset` command,

```
. xtset idpers
      panel variable:  idpers (unbalanced)
```

and then fit the model with the BUC estimator using `feologit`. We issued the following command:

```
. feologit lifesat kidade01_0-kidade01_4 age lhinc work
note: group() not specified; assuming group(idpers) from panel identifier
note: multiple positive outcomes within groups encountered.
Iteration 0:  log conditional likelihood = -172931.52
Iteration 1:  log conditional likelihood = -170560.88
Iteration 2:  log conditional likelihood = -170558.47
Iteration 3:  log conditional likelihood = -170558.47
Fixed-effects ordered logistic regression

                                N. of obs. (inc. copies) =    469318
                                N. of observations      =    115257
                                N. of panel units       =     11247
                                Wald chi2(8)           =    1356.60
                                Prob > chi2            =     0.0000
Log conditional likelihood = -170558.47  Pseudo R2      =     0.0222
                                (Std. Err. adjusted for 11,247 clusters in idpers)
```

lifesat	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
kidade01_0	.7795942	.0526895	14.80	0.000	.6763246	.8828637
kidade01_1	.6458431	.0514869	12.54	0.000	.5449306	.7467555
kidade01_2	.2123843	.0488413	4.35	0.000	.1166571	.3081115
kidade01_3	.1570779	.0457228	3.44	0.001	.0674628	.246693
kidade01_4	.0118047	.0426905	0.28	0.782	-.0718671	.0954765
age	-.059859	.0022636	-26.44	0.000	-.0642956	-.0554224
lhinc	.5426902	.0262043	20.71	0.000	.4913307	.5940497
work	.2002279	.0212617	9.42	0.000	.1585558	.2419

The output shows that the algorithm for maximizing the log conditional likelihood converged after three steps. For fitting the model parameters, only individuals (panel units) who have variation in their dependent variables are informative. Individuals who are observed only once or have always the same life satisfaction scores over time are excluded by the program (because their LL contribution is zero). This condition is met by 11,247 individuals, which results in 115,257 observations. On average, people in the estimation sample are therefore observed about 10 times. The ordered dependent variable has 11 categories, so 10 different dichotomizations are possible. However, because not all dichotomizations lead to copies with variation in the binary dependent variable, we end up with 469,318 copies that contribute to the estimation procedure. Because the copies are not independent of each other, `feologit` calculates cluster-adjusted standard errors at the individual level (11,247 individuals).

The Wald test indicates that all eight included variables are jointly statistically significant. Regarding the effect of having the first child, the effect is highest in the year of birth (coefficient on `kidade01_0`). Thereafter, the effect decreases and reaches a nonsignificant level after four years. Age has a negative effect after controlling for time-invariant characteristics and the other variables in the model. As expected, household income and working have a positive effect on life satisfaction. The compensating variation between work and log household income is about 0.37, meaning that log household

income has to increase by 0.37 to compensate for not working. This is equivalent to saying that household income has to increase by 45% [ $\exp(0.37) - 1$ ] to offset a nonworking status.

Odds ratios for interpreting the effect sizes can be obtained by using the option `or`. Below are the code and an excerpt of the output:

```
. feologit, or
(output omitted)
```

lifesat	Odds Ratio	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
kidage01_0	2.180587	.1148941	14.80	0.000	1.966636	2.417814
kidage01_1	1.907595	.0982161	12.54	0.000	1.724489	2.110143
kidage01_2	1.236623	.0603983	4.35	0.000	1.123734	1.360853
kidage01_3	1.170087	.0534997	3.44	0.001	1.06979	1.279786
kidage01_4	1.011875	.0431974	0.28	0.782	.9306546	1.100183
age	.9418973	.0021321	-26.44	0.000	.9377277	.9460854
lhinc	1.720629	.0450879	20.71	0.000	1.63449	1.811309
work	1.221681	.025975	9.42	0.000	1.171817	1.273667

Having the first child increases the odds ratio by about 118% in the year of birth, about 91% in the first, about 24% in the second, and about 17% in the third year thereafter. In the fourth year, the effect is essentially no longer present.

Marginal effects at the average can be obtained by using the postestimation command `logitmarg` after fitting the model. They are computed using the relative frequencies of the corresponding categories in the estimation sample. Below are the code and the output. We specified the option `dydx(kidage01_0)` to limit the output to the MES of the year of birth.

```
. logitmarg, dydx(kidage01_0)
```

Marginal effects at the average

N. of observations= 115257

N. of panel units = 11247

	Margin	Std. Err.	z	P> z	[95% Conf. Interval]	
1	-.0039435	.0002665	-14.80	0.000	-.0044659	-.0034211
2	-.0031271	.0002113	-14.80	0.000	-.0035414	-.0027129
3	-.0090451	.0006113	-14.80	0.000	-.0102433	-.007847
4	-.0196532	.0013283	-14.80	0.000	-.0222566	-.0170498
5	-.0253752	.001715	-14.80	0.000	-.0287365	-.0220139
6	-.0718074	.0048532	-14.80	0.000	-.0813194	-.0622954
7	-.0401357	.0027126	-14.80	0.000	-.0454523	-.0348191
8	-.0197687	.0013361	-14.80	0.000	-.0223874	-.0171501
9	.0881895	.0059604	14.80	0.000	.0765074	.0998716
10	.0661833	.0044731	14.80	0.000	.0574163	.0749503
11	.0384832	.0026009	14.80	0.000	.0333855	.0435809

Note that `logitmarg` enumerates categories starting at 1 and ignores the actual (arbitrary) labels of the dependent variable, which in our case starts at 0. Because having a child has a positive effect on life satisfaction in the first year, the marginal



probability effects at the average are negative for the lower categories and positive for the ninth and higher categories. For example, in the first year having a child decreases the probability of falling into the sixth category by 7.2% points and increases the probability of having the highest rating by 3.8% points for this average person, everything else being equal. The effects for the lowest categories are small because only a few people have such a low life-satisfaction status.

Estimating the thresholds is possible by using the option `thresholds`, which calls the BUC- $\tau$  estimator. This requires the assumption that the spacing between the thresholds is the same for all individuals. Below is the output:

```
. feologit lifesat kidge01_0-kidge01_4 age lhinc work, group(idpers) threshold
note: multiple positive outcomes within groups encountered.
Iteration 0:  log conditional likelihood = -335358.34
Iteration 1:  log conditional likelihood = -317080.65
Iteration 2:  log conditional likelihood = -316611.97
Iteration 3:  log conditional likelihood = -316611.28
Iteration 4:  log conditional likelihood = -316611.28
Fixed-effects ordered logistic regression
                                N. of obs. (inc. copies) =    1544471
                                N. of observations      =    1166665
                                N. of panel units       =     11725
                                Wald chi2(17)          =   33921.94
                                Prob > chi2            =     0.0000
Log conditional likelihood = -316611.28 Pseudo R2      =     0.5250
                                (Std. Err. adjusted for 11,725 clusters in idpers)
```

lifesat	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
kidge01_0	.7835162	.0537423	14.58	0.000	.6781832	.8888491
kidge01_1	.6396636	.052942	12.08	0.000	.5358992	.743428
kidge01_2	.2046826	.0501304	4.08	0.000	.1064288	.3029365
kidge01_3	.1624746	.047537	3.42	0.001	.0693037	.2556455
kidge01_4	.0107345	.0432988	0.25	0.804	-.0741297	.0955986
age	-.0578395	.002264	-25.55	0.000	-.0622769	-.0534021
lhinc	.5471535	.0269294	20.32	0.000	.4943729	.5999342
work	.200715	.021935	9.15	0.000	.1577232	.2437068
/cut1	0 (constrained)					
/cut2	.7284609	.0615201			.6078836	.8490381
/cut3	1.637104	.0677637			1.504289	1.769918
/cut4	2.666144	.0709232			2.527137	2.805151
/cut5	3.424063	.0726216			3.281727	3.566399
/cut6	4.872104	.0734364			4.728172	5.016037
/cut7	5.696417	.0749376			5.549542	5.843292
/cut8	7.036473	.0766951			6.886153	7.186792
/cut9	9.097256	.0807096			8.939068	9.255444
/cut10	10.65767	.0872651			10.48663	10.8287

In contrast with the procedure without thresholds, the cutoff point within a likelihood contribution (clone) can change. Therefore, even when the ordered dependent variable is constant, the resulting dichotomized dependent variable with different cutoff points can vary. This increases the number of individuals in the estimation sample

slightly to 11,725 and the number of observations to 116,665. The estimator includes all contributions of the BUC estimator plus 10 copies of each individual with random variation in the cutoff point. Therefore, the number of included copies increases noticeably to over 1.5 million.

The regression coefficient  $\beta$  has the same interpretation as before. One can also display the odds ratios. From the output, we see that the exact estimates changed only slightly. Because life satisfaction can take 11 different values and the first finite threshold `/cut1` (corresponding to  $\tau_2$ ) is normalized to zero, the output shows estimates for the 2nd to the 10th threshold. Careful inspection of the spaces shows that there is a moderate tendency that differences increase toward the top. Where the difference between the first and second, and second and third, is 0.728 and 0.909, respectively, the two spaces at the upper end are 2.06 and 1.56. This implies that changes of regressors have a larger effect on the observed ordered dependent variable for unhappy individuals compared with happy individuals.

The spaces range from 0.73 to 2.06, where the second smallest difference is 0.76. The size of the MEs of the different regressors are rather small compared with these spaces. Except for the unhappiest people (lowest category), having a child can never increase life satisfaction by more than one point, everything else being equal. The same is true for doubling the income [ $\exp(0.547) = 1.73$ ]. Regarding age, for a person with a life satisfaction score of 9, 26 years need to pass for her to surely change to the next lower category  $[(10.35 - 8.91)/0.057]$ .

The BUC and BUC- $\tau$  estimator are both consistent under the additional assumption of constant thresholds. Therefore, we can test this assumption in the form of a generalized Hausman specification test by comparing the two estimates. If they are statistically different, we can reject the null hypothesis of constant thresholds.<sup>3</sup> There are different possibilities to implement such a test in Stata. The most obvious way would be to use the command `suest`. Using this command requires fitting the model without clustered standard errors and, in a second step, computing a joint covariance matrix of both estimators adjusting for clusters. However, because computing the BUC estimator without clusters can be misleading and adjusting them after using `clogit` can lead to error messages if the group variable is not properly nested within the cluster variable, we decided to work with interaction terms. In a first step, the command `feologit, threshold` with the additional option `keepsample` is used:

```
. quietly: feologit lifesat kidage01_0-kidage01_4 age lhinc work, group(idpers)
> threshold keep
```

The option `keepsample`, here shortened to `keep`, keeps the inflated dataset of clones that is used to compute the BUC- $\tau$  estimator after the estimation. In addition, the dichotomized dependent variable is stored under `dkdepvar`, the corresponding cutoff point under `dkthreshold`, and the variable `bucsample` indicates whether the cutoff varies within the clone (`bucsample = 0`) or not (`bucsample = 1`). Observations with `bucsample = 1` constitute the sample of the BUC estimator. The remaining observations

---

3. The precision of the test might be improved by increasing the number of clones used in estimation.

are primarily used to estimate the thresholds but contribute to the estimates of  $\beta$ , too. Under the null hypothesis of constant thresholds, both samples (`bucsample` = 0 and `bucsample` = 1) can be used to estimate  $\beta$ . Statistical differences between the two estimates indicate that the thresholds are not constant. Thus, after having used `feologit` with the options `thresholds` and `keepsample`, we can implement this test by interacting the variable `bucsample` with all regressors and fitting the model with `clogit` while grouping on the clone variable and clustering on the individual level. Finally, we test whether the differences of the interaction terms are jointly equal to zero.

As an auxiliary step, we first use a small loop to create the interaction variables (prefixed `tau_`) and store them in the local macro `interact`:

```
. local interact=""
. foreach i of var kidage01_0-kidage01_4 age lhinc work {
2. quietly: generate tau_`i' = `i'*(bucsample==0)
3. local interact="`interact' tau_`i'"
4. }
```

We then fit the model with interactions using the `clogit` command, grouping on the clone variable and clustering on the individual level:

```
. clogit dkdepvar kidade01_0-kidade01_4 age lhinc work `interact' i.dkthreshold,
> group(clonegroup) cluster(idpers)
note: multiple positive outcomes within groups encountered.
note: 132,242 groups (827,909 obs) dropped because of all positive or
      all negative outcomes.
Iteration 0:  log pseudolikelihood = -335337.34
Iteration 1:  log pseudolikelihood = -317228.3
Iteration 2:  log pseudolikelihood = -316599
Iteration 3:  log pseudolikelihood = -316594.58
Iteration 4:  log pseudolikelihood = -316594.58
Conditional (fixed-effects) logistic regression

                                Number of obs    =   1,544,470
                                Wald chi2(25)      =   34242.88
                                Prob > chi2        =    0.0000
                                Pseudo R2          =    0.5251
Log pseudolikelihood = -316594.58
                                (Std. Err. adjusted for 11,725 clusters in idpers)
```

dkdepvar	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
kidade01_0	.7795942	.0526894	14.80	0.000	.6763248	.8828635
kidade01_1	.6458431	.0514868	12.54	0.000	.5449308	.7467553
kidade01_2	.2123843	.0488412	4.35	0.000	.1166573	.3081113
kidade01_3	.1570779	.0457228	3.44	0.001	.0674629	.2466928
kidade01_4	.0118047	.0426904	0.28	0.782	-.0718669	.0954764
age	-.059859	.0022636	-26.44	0.000	-.0642956	-.0554225
lhinc	.5426902	.0262043	20.71	0.000	.4913307	.5940496
work	.2002279	.0212616	9.42	0.000	.1585559	.2418999
tau_kidade01_0	.008167	.0306995	0.27	0.790	-.052003	.068337
tau_kidade01_1	-.013622	.029474	-0.46	0.644	-.07139	.0441459
tau_kidade01_2	-.0168844	.0268302	-0.63	0.529	-.0694707	.0357018
tau_kidade01_3	.0119503	.0251759	0.47	0.635	-.0373935	.0612941
tau_kidade01_4	-.0023741	.024214	-0.10	0.922	-.0498327	.0450844
tau_age	.0040727	.0008054	5.06	0.000	.0024942	.0056513
tau_lhinc	.0093463	.0121449	0.77	0.442	-.0144573	.0331498
tau_work	.0009298	.0102628	0.09	0.928	-.0191849	.0210444
dkthreshold						
2	-.7305085	.061547	-11.87	0.000	-.8511384	-.6098786
3	-1.638868	.0678167	-24.17	0.000	-1.771786	-1.505949
4	-2.667425	.0709832	-37.58	0.000	-2.80655	-2.5283
5	-3.424848	.072681	-47.12	0.000	-3.567301	-3.282396
6	-4.871923	.0734845	-66.30	0.000	-5.01595	-4.727896
7	-5.695513	.0749785	-75.96	0.000	-5.842468	-5.548557
8	-7.034464	.0767388	-91.67	0.000	-7.184869	-6.884058
9	-9.093563	.0807186	-112.66	0.000	-9.251769	-8.935358
10	-10.65302	.0873042	-122.02	0.000	-10.82414	-10.48191

As can be seen, the estimates of the noninteracted regressors are numerically equivalent to the previous BUC estimates. The estimates of the interaction terms mirror the differences between the BUC and the BUC- $\tau$  estimates but are not identical because the thresholds adjust to the additional regressors. The differences are small, and most of them are not individually significant. To formally test whether they are jointly different, we can use Stata's `test` command:

```
. test `interact'
( 1) [dkdepvar]tau_kidage01_0 = 0
( 2) [dkdepvar]tau_kidage01_1 = 0
( 3) [dkdepvar]tau_kidage01_2 = 0
( 4) [dkdepvar]tau_kidage01_3 = 0
( 5) [dkdepvar]tau_kidage01_4 = 0
( 6) [dkdepvar]tau_age = 0
( 7) [dkdepvar]tau_lhinc = 0
( 8) [dkdepvar]tau_work = 0

      chi2( 8) =    39.13
      Prob > chi2 =    0.0000
```

Despite the differences being small, the joint hypothesis that all these terms are equal to zero can be rejected at the 0.1% level [ $\chi^2(8) = 39.13$ ] because of the large sample. Therefore, formally we reject the null hypothesis of constant thresholds and should use the BUC instead of the BUC- $\tau$  estimator to interpret the results. However, given that differences in the estimates in this case are so small, the thresholds might still shed some light on the data-generating process but should not be overinterpreted.

## 6 Conclusion

In this article, we presented and discussed the BUC and BUC- $\tau$  estimators of the fixed-effects ordered logit model and introduced a new community-contributed command that implements these estimators in Stata, `feologit`.

BUC and BUC- $\tau$  both offer consistent estimates of the slope parameters  $\beta$ . In addition, BUC- $\tau$  obtains consistent estimates of the thresholds  $\tau$ , under the slightly more restrictive assumption that they do not vary between individuals. While the calculation of average MEs is not possible with these estimators, useful identified objects of interest include odds ratios, compensating variation, and other quantities. Of particular interest is a particular ME at the average, which we proposed in this article and for which a dedicated postestimation command is available with `feologit`. Finally, we also presented a new specification test that can be used to evaluate the assumption of constant thresholds by comparing BUC and BUC- $\tau$  estimates and that can be easily implemented with a few lines of code in Stata.

## 7 Acknowledgments

The authors thank Johannes Kunz for comments on an early version of this article. Staub gratefully acknowledges financial support from the Australian Research Council through grant DE170100644.

## 8 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 20-2
. net install st0596      (to install program files, if available)
. net get st0596          (to install ancillary files, if available)
```

## 9 References

- Abrevaya, J. 1997. The equivalence of two estimators of the fixed-effects logit model. *Economics Letters* 55: 41–43. [https://doi.org/10.1016/S0165-1765\(97\)00044-X](https://doi.org/10.1016/S0165-1765(97)00044-X).
- Andersen, E. B. 1970. Asymptotic properties of conditional maximum-likelihood estimators. *Journal of the Royal Statistical Society, Series B* 32: 283–301. <https://doi.org/10.1111/j.2517-6161.1970.tb00842.x>.
- Baetschmann, G. 2012. Identification and estimation of thresholds in the fixed effects ordered logit model. *Economics Letters* 115: 416–418. <https://doi.org/10.1016/j.econlet.2011.12.100>.
- Baetschmann, G., K. E. Staub, and R. Studer. 2016. Does the stork deliver happiness? Parenthood and life satisfaction. *Journal of Economic Behavior & Organization* 130: 242–260. <https://doi.org/10.1016/j.jebo.2016.07.021>.
- Baetschmann, G., K. E. Staub, and R. Winkelmann. 2015. Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society, Series A* 178: 685–703. <https://doi.org/10.1111/rssa.12090>.
- Boes, S., and R. Winkelmann. 2006. Ordered response models. *Allgemeines Statistisches Archiv* 90: 167–181. <https://doi.org/10.1007/s10182-006-0228-y>.
- Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Chamberlain, G. 1980. Analysis of covariance with qualitative data. *Review of Economic Studies* 47: 225–238. <https://doi.org/10.2307/2297110>.
- Clark, A. E., E. Diener, Y. Georgellis, and R. E. Lucas. 2008. Lags and leads in life satisfaction: A test of the baseline hypothesis. *Economic Journal* 118: F222–F243. <https://doi.org/10.1111/j.1468-0297.2008.02150.x>.

- Greene, W. 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal* 7: 98–119. <https://doi.org/10.1111/j.1368-423X.2004.00123.x>.
- Kitazawa, Y. 2012. Hyperbolic transformation and average elasticity in the framework of the fixed effects logit model. *Theoretical Economics Letters* 2: 192–199. <https://doi.org/10.4236/tel.2012.22034>.
- Lancaster, T. 2000. The incidental parameter problem since 1948. *Journal of Econometrics* 95: 391–413. [https://doi.org/10.1016/S0304-4076\(99\)00044-5](https://doi.org/10.1016/S0304-4076(99)00044-5).
- Neyman, J., and E. L. Scott. 1948. Consistent estimates based on partially consistent observations. *Econometrica* 16: 1–32. <https://doi.org/10.2307/1914288>.
- Santos Silva, J. M. C., and G. C. R. Kemp. 2016. Partial effects in fixed effects models. UK Stata Users Group meeting proceedings. [https://www.stata.com/meeting/uk16/slides/santos\\_uk16.pdf](https://www.stata.com/meeting/uk16/slides/santos_uk16.pdf).
- Stutzer, A., and B. S. Frey. 2006. Does marriage make people happy, or do happy people get married? *Journal of Socio-Economics* 35: 326–347. <https://doi.org/10.1016/j.socec.2005.11.043>.
- Wagner, G. G., J. R. Frick, and J. Schupp. 2007. The German Socio-Economic Panel Study (SOEP)—Scope, evolution and enhancements. *Schmollers Jahrbuch* 127: 139–169.
- Winkelmann, R., and S. Boes. 2010. *Analysis of Microdata*. 2nd ed. Verlag: Springer.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

### About the authors

Gregori Baetschmann is a lecturer in the Department of Sociology at the University of Bern, Switzerland.

Alexander Ballantyne is a PhD candidate in the Department of Economics at the University of Melbourne, Australia.

Kevin Staub is a senior lecturer in the Department of Economics at the University of Melbourne, Australia.

Rainer Winkelmann is a professor in the Department of Economics at the University of Zurich, Switzerland.