Prevention Science
https://doi.org/10.1007/s11121-021-01302-y

Ordered Regression Models: a Tutorial

Andrew S. Fullerton[1] · **Kathryn Freeman Anderson**[2]

Accepted: 14 September 2021

Abstract

Ordinal outcomes are common in the social, behavioral, and health sciences, but there is no commonly accepted approach to analyzing them. Researchers make a number of different seemingly arbitrary recoding decisions implying different levels of measurement and theoretical assumptions. As a result, a wide array of models are used to analyze ordinal outcomes, including the linear regression model, binary response model, ordered models, and count models. In this tutorial, we present a diverse set of ordered models (most of which are under-utilized in applied research) and argue that researchers should approach the analysis of ordinal outcomes in a more systematic fashion by taking into consideration both theoretical and empirical concerns, and prioritizing ordered models given the flexibility they provide. Additionally, we consider the challenges that ordinal independent variables pose for analysts that often go unnoticed in the literature and offer simple ways to decide how to include ordinal independent variables in ordered regression models in ways that are easier to justify on conceptual and empirical grounds. We illustrate several ordered regression models with an empirical example, general self-rated health, and conclude with recommendations for building a sounder approach to ordinal data analysis.

Keywords Ordered regression models · Ordered logit · Ordinal variables

Ordinal outcomes are variables with categories sorted on a single characteristic with unknown spacing between categories. They are very common in the social, behavioral, and health sciences. However, approaching the analysis of ordinal outcomes in a systematic way is surprisingly uncommon in applied research. Researchers approach the (re)coding of ordinal outcomes and choice of a methodological technique in a myriad of ways. Prior to statistical advances in the 1980s, it was commonplace to assume that the variable was continuous in order to use the linear regression model, and this is still considered an acceptable approach by some today. Another popular approach is to dichotomize the outcome in order to use the binary regression model (e.g., logit or probit). Why is ordinality considered so problematic?

Ordinal outcomes lie at the intersection of continuous and categorical variables. They are categorical by definition, but they certainly have continuous elements as well. For example, the first ordered regression model, ordered probit, was conceptualized as a model for imperfectly measured continuous variables (McKelvey & Zavoina, 1975). Ordinal variables are theorized as truly continuous measures that are only measured at a few discrete points due to practical constraints. From the start, ordinal regression models were motivated by the assumption of a latent continuous variable generating the observed data (Agresti, 2010; Long, 1997). However, this assumption is not necessary for either model estimation or interpretation (McCullagh, 1980).

Ordered regression models are also inherently categorical due to the fact that they are extensions of the binary regression model (Fullerton, 2009). The main criticism of the binary approach to ordinal outcomes is that the choice of one dichotomization of Y over another is arbitrary. For an outcome with M categories, there are $M-1$ ways to create a binary outcome retaining the category order. For some variables, such as the measure we use as the example in this paper, general self-rated health, the choice of a cutoff seems relatively straightforward. For this measure, respondents are asked to indicate whether their health, in general, is excellent (1), good (2), fair (3), or poor (4). Most analysts combine categories 1 and 2 in order to create a binary measure of "good health" or categories 3 and 4 for a measure of "poor health." There are a number of other commonly used ordinal health measure in the literature, as well as some that are

✉ Andrew S. Fullerton
andrew.fullerton@okstate.edu

1 Oklahoma State University, Stillwater, USA

2 University of Houston, Houston, USA

being re-constructed as ordinal, such as the physical component summary (PCS) and the mental component summary (MCS), using confirmatory factor analysis of polychoric correlations (Tucker et al., 2013, 2014).

However, it would not be difficult to image scenarios in which researchers may recode the ordinal outcome in a different way. For example, analysts typically combine sparse outcome categories due to concerns over convergence problems. Thus, if there was a lot of clustering at one end of the ordinal scale or the other, then a researcher might decide to compare category 1 or 4 to the remaining categories combined. However, the situation is more difficult when there is a neutral middle category, such as the "neither agree nor disagree" category in Likert scales (Likert, 1932). The analyst must arbitrarily decide to combine the middle category with one end of the scale or treat the cases as missing values. Neither strategy is without limitations (Nadler et al., 2015). Additionally, one may have a theoretical reason for preferring one coding scheme over another, which is an important yet overlooked consideration (Fullerton & Xu, 2016).

Ordered regression models bypass the problematic choice of a cutoff point for dichotomizing y by considering all $M-1$ alternatives simultaneously. Each "cutpoint" equation (McCullagh, 1980) is estimated simultaneously with a common set of slopes and an equation-specific constant, which yields a relatively parsimonious model. This slope constraint, known as the "parallel regression assumption" (Long, 1997, p.140), ensures that the model is ordinal and increases the level of parsimony by reducing the number of required parameters per independent variable from $M-1$ to 1. However, it also can produce misleading results if the coefficients vary substantially across equations (Fullerton & Xu, 2016).

The binary approach to analyzing ordinal outcomes amounts to focusing on one cutpoint equation. This is problematic for several reasons. First, this implicitly assumes that the parallel regression assumption is reasonable (Long, 1997). However, this assumption is frequently violated in practice (Long & Freese, 2014). As a result, the arbitrary choice of one equation over another could have a non-trivial impact on the results. Ordinal regression models are preferred over binary models because they utilize information from every cutpoint equation.

Despite the references to a single "ordinal logit/probit" model, there are at least a dozen different ordinal regression models (Fullerton, 2009). In this study, we present the three classes of ordered regression models (cumulative, stage, and adjacent) with four variations within each class based on the application of the parallel regression assumption to all, some, or none of the independent variables.[1] The focus in this study is the traditional or "frequentist" approach to statistics, but recent developments in Bayesian data analysis have made this alternative approach very useful for the analysis of ordinal outcomes (see Agresti, 2010, Ch.11; Bürkner & Vuorre, 2019). For example, the Bayesian approach to ordered regression models provides additional means of testing the parallel regression assumption and more flexible cumulative models that avoid some convergence issues (for more details, see Xu et al., 2019). We begin with the most popular model: the cumulative ordered regression model.

Cumulative Models

Cumulative ordered regression models are an extension of the binary regression model that treat the original, ordinal outcome as a set of $M-1$ binary outcomes corresponding to different cutpoints in the cumulative distribution (McCullagh, 1980). The probability of interest is the cumulative probability, $\Pr(\mathrm{Y} \le m)$, which is the probability at or below category m. Let us consider a four-category outcome, such as general self-rated health, as an example. The cumulative ordered regression model connects the original, ordinal outcome, y, to the set of binary outcomes, y_1 to y_3, with the following measurement equations (Fullerton & Xu, 2016, p.5):

$$y_1 = \begin{cases} 1 & \text{if } y = 1 \\ 0 & \text{if } y = 2-4 \end{cases} \tag{1}$$

$$y_2 = \begin{cases} 1 & \text{if } y = 1-2 \\ 0 & \text{if } y = 3-4 \end{cases} \tag{2}$$

$$y_3 = \begin{cases} 1 & \text{if } y = 1-3 \\ 0 & \text{if } y = 4 \end{cases} \tag{3}$$

The cumulative ordered regression model simultaneously estimates these three binary regression models corresponding to three different cumulative probabilities. For self-rated health, the three equations focus on different points in the cumulative distribution: (1) excellent vs. good to poor health, (2) excellent or good vs. fair or poor health, and (3) excellent to fair vs. poEqor health. The traditional version of this model imposes the parallel regression assumption, which constrains the coefficients for a given independent variable, x_k, to remain constant across equations.

Cumulative Ordered Regression Models with Proportional Odds

As a latent variable model, the observed ordinal outcome, y, is connected to the underlying continuous measure, y^*, through the following equation:

[1] There are constrained and unconstrained partial models, which yields four versions of each approach.

$$y = m \text{ if } \tau_{m-1} \le y* < \tau_m (1 \le m \le M) \tag{4}$$

The observed outcome categories correspond to different segments of the latent continuous variable, and the "cutpoints" (τ_m) are estimated along with the slopes using maximum likelihood estimation. The parallel regression assumption constrains the slopes across cutpoint equations. If $M = 4$, then the model imposes the following restriction for each independent variable: $\beta_{k1} = \beta_{k2} = \beta_{k3}$. The only coefficients that vary across cutpoint equations are the cutpoints. This yields a relatively parsimonious equation for the cumulative probability (Long, 1997, p.121):

$$\Pr(y \le m|\mathbf{x}) = \Pr\left(y^* < \tau_m|\mathbf{x}\right) = F(\tau_m - \mathbf{x\beta}) \tag{5}$$

The parallel regression assumption is a defining feature of the traditional ordered regression model. For cumulative models, it ensures that the model is ordinal and is motivated by a latent continuous variable generating the data. It also guarantees that the predicted probabilities are non-negative (Greene & Hensher, 2010). Relaxing the assumption creates the potential for negative predicted probabilities due to the fact that non-parallel cumulative probability curves eventually intersect (McCullagh & Nelder, 1989, p.155). However, negative predicted probabilities are relatively rare, and one can resolve this problem by recoding continuous variables into binary variables (see Hedeker et al., 1999, p.63–64).

The parallel regression assumption also increases the level of parsimony in the model. A common slope across cutpoint equations for each independent variable results in a savings of $M - 2$ parameters per variable. However, there is a cost to this parsimony. The model restriction may produce misleading results if there is a substantial amount of coefficient variation across cutpoint equations. In extreme cases, the restriction can produce non-significant coefficients that average together significant coefficients with opposite signs. For example, there may be a significant, negative association in the first cutpoint equation and a significant, positive association in the second cutpoint equation. The application of the parallel regression assumption leads to an averaging of the coefficients that suggests the lack of an association between x and y. However, in this case, there is an association, but the relationship is asymmetrical (see Lieberson, 1985; Fullerton & Dixon, 2009).

The parallel regression assumption is misunderstood in the applied social science literature in two key respects. First, a significant test statistic, such as the Brant test (Brant, 1990; Long & Freese, 2014), does not necessarily mean that the assumption is "violated" to such an extent that the model is no longer appropriate or useful. Conceptualized as a test of competing models (with and without cutpoint-specific slopes), a significant test statistic indicates that the less parsimonious non-parallel model fits the data "better." However, the significant coefficient variation across equations may not lead to substantively different conclusions. Thus, the simpler model requiring fewer parameters may still be the preferable model (Fullerton & Xu, 2016).

Second, if the parallel regression assumption is reasonable, then the analyst would be justified in simply collapsing the categories into a single binary outcome given that the coefficients are constant across equations (Long, 1997). An inspection of journals in the social sciences reveals an inverted use of this logic: researchers decide to dichotomize the outcome and employ a binary regression model *because* the parallel regression assumption is violated. A meaningful violation of the assumption suggests that the coefficients are equation-dependent. A model that allows for this dependence, such as partial proportional odds, would be preferable (Peterson & Harrell, 1990). We will introduce this model in the next section.

Regardless of the results from tests of the parallel regression assumption, the decision to treat the outcome as ordinal, binary, or even continuous is never completely empirical. Every methodological decision is also at least implicitly theoretical (Bourdieu et al., 1991). The recoding choices and selection of a methodological technique for analyzing ordinal outcomes would be easier to justify if researchers explicitly incorporated theoretical concerns into this process. For example, does the theory employed suggest the outcome is a quantitative resource or a type of group membership? The advantage of using ordered regression models is that they can accommodate either theoretical position. Once a researcher has dichotomized the outcome, the theoretical assumption of either group membership or threshold effect is already established.

Cumulative Ordered Regression Models that Relax Proportional Odds

If there are theoretical and/or empirical reasons for relaxing the proportional odds or parallel regression assumption, then researchers may rely on a partial or non-parallel cumulative model (see Fullerton, 2009; Fullerton & Xu, 2016, Ch. 3–4). Partial models relax the parallel regression assumption for a subset of independent variables (Fullerton & Xu, 2016, p.57):

$$\Pr(y \le m|\mathbf{x}) = \Pr\left(y^* < \tau_m|\mathbf{x}\right) = F\left(\tau_m - \mathbf{x\beta} - \boldsymbol{\omega\eta}_m\right) \tag{6}$$

where ω is a vector of independent variables forming this subset and η is a vector of coefficients that vary across cutpoint equations. There is also a constrained version of the partial model, but we only consider the unconstrained version in this study (see Fullerton & Xu, 2012).

Partial models retain the parallel assumption for one subset of variables for which it is theoretically or empirically meaningful and relax it for the remaining variables.

By contrast, non-parallel ordered models relax the parallel regression assumption for every independent variable in the model. The omnibus test of the parallel regression assumption reported in some programs is essentially a comparison of the model fit in the non-parallel model relative to the parallel model. The equation for the non-parallel model is (Fullerton & Xu, 2016, p.86):

$$\Pr(y \le m|\mathbf{x}) = \Pr\left(y^* < \tau_m \middle| \mathbf{x}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta}_m) \tag{7}$$

The subscript, m, indicates that the coefficients for each independent variable are allowed to vary across cutpoint equations. This model is also known as the "generalized ordered" logit or probit model (Williams, 2006, 2016). It is less efficient than the parallel and partial models because it has cutpoint equation-specific slopes for every independent variable, which dramatically increases the number of estimated parameters as the number of outcome categories and independent variables increase. However, if the sample size is large, then the inefficiency due to the additional parameters may not be costly, and the added flexibility allows the researcher to explore new theoretical and empirical possibilities in the model.

For example, the partial and non-parallel cumulative models allow one to consider the possibility of asymmetrical effects (Lieberson, 1985). Opposition to and support for a particular policy or idea may not have effects that are simply mirror images of one another. Factors may shape opposition and support in different ways relative to a neutral middle position (e.g., Fullerton & Dixon, 2009). The imposition of the parallel regression assumption prevents researchers from observing these asymmetrical relationships, which highlights the importance of partial and non-parallel models for detecting more nuanced relationships.

Stage Models

Continuation Ratio Models with Proportional Odds

The popularity of cumulative models is evidenced by the fact that practitioners refer to them simply as ordinal regression. In fact, there are several types of ordinal regression models. The second class of ordered models, stage models (or "continuation ratio" [Fienberg, 1980, p.110] models), conceptualize the cutpoint equations as an irreversible sequence of stages. In order to reach later stages, respondents must first "survive" the transitions from each of the previous stages. For example, in order to earn a Ph.D., one must first successfully pass through the several educational stages (e.g., high school and college). The sample size is progressively smaller in later cutpoint equations or stages because only those respondents that "survived" the earlier transitions are capable of experiencing the event and therefore included in the sample. The stage or "continuation ratio" model with the parallel assumption (and the cloglog link function) is equivalent to the proportional hazards model (Cox, 1972) used in event history analysis (Fullerton, 2009, p.317). For a four-category outcome, the continuation ratio ordered regression model connects the original, ordinal outcome, y, to the set of binary outcomes, y_1 to y_3, with the following measurement equations (Fullerton & Xu, 2016, p.6–7):

$$y_1 = \begin{cases} 1 & \text{if } y = 1 \\ 0 & \text{if } y = 2 - 4 \end{cases} \tag{8}$$

$$y_2 = \begin{cases} 1 & \text{if y} = 2 & | \quad y \ge 2 \\ 0 & \text{if y} = 3 - 4 & | \quad y \ge 2 \end{cases} \tag{9}$$

$$y_3 = \begin{cases} 1 & \text{if y} = 3 & | \quad y \ge 3 \\ 0 & \text{if y} = 4 & | \quad y \ge 3 \end{cases} \tag{10}$$

The probability of interest is the conditional probability and the equation for the continuation ratio model with proportional odds is (Fullerton & Xu, 2016, p.32):

$$\Pr(y = m|y \ge m,\ \mathbf{x}) = \Pr\left(y^* < \tau_m \middle| y \ge m, \mathbf{x}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta}) \tag{11}$$

In other words, the focus is on the probability of a specific category given that one has progressed to that stage of the sequential process. Continuation ratio models are useful for analyzing ordinal outcomes that are the result of a sequential process, but most ordinal outcomes are not sequential. For example, Likert scale questions do not have a natural starting point and one does not have to proceed through each level of support or opposition to reach the position at the other extreme. However, stage models will often produce similar results to other ordinal models that apply the proportional odds assumption in the same manner (Fullerton, 2009). Analysts could simply note whether the results change in a meaningful way if one chooses the other extreme (e.g., category 4 rather than category 1) as the starting point. In practice, it is the application of the proportional odds assumption rather than the class of model (cumulative, stage, or adjacent) that has the most important influence on the results.

Continuation Ratio Models without Proportional Odds

We may relax the proportional odds assumption for a subset of independent variables with the partial model and for the entire model with a non-parallel model. The equation for the conditional probability in the unconstrained partial model is (Fullerton & Xu, 2016, p.64)[2]

$$\begin{aligned}\Pr(y = m|y \geq m, \mathbf{x}, \boldsymbol{\omega}) &= \Pr\left(y^* < \tau_m|y \geq m, \mathbf{x}, \boldsymbol{\omega}\right) \\ &= F\left(\tau_m - \mathbf{x}\boldsymbol{\beta} - \boldsymbol{\omega}\boldsymbol{\eta}_m\right)\end{aligned} \tag{12}$$

Relaxing the proportional odds assumption for the entire model results in the non-parallel continuation ratio model, which is also known as the sequential logit or probit model (Amemiya, 1981; Tutz, 1991). The equation for the conditional probability is (Fullerton & Xu, 2016, p.93)

$$\Pr(y = m|y \geq m,\ \mathbf{x}) = \Pr\left(y^* < \tau_m|y \geq m, \mathbf{x}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta}_m) \tag{13}$$

Stage models require one additional assumption that is not required in cumulative models. We must assume that the error terms are uncorrelated across cutpoint equations (Maddala, 1983, p.51). In other words, we assume that there is no selection bias in later stages, which in practice is often an unrealistic assumption. The later-stage samples are typically more homogenous, which can introduce additional bias into the estimates (Buis, 2011; Cameron & Heckman, 1998; Mare, 2011; Xie, 2011). Additionally, one may interpret changing coefficients across stages as due to changes in unobserved heterogeneity rather than true "effects." This may be seen as a limitation of the model, but it also opens up additional solutions to violations of the proportional odds assumption (e.g., see Williams, 2009).

Adjacent Models

Adjacent Category Models with Proportional Odds

If the outcome categories are of substantive interest to the researcher and the variable is not the result of a sequential process, then the adjacent class of models is a useful, yet underutilized, alternative to cumulative models. The probability of interest in the adjacent model is a conditional probability, as it is in stage models, but the focus is limited to two adjacent categories at a time. Adjacent models grew out of the log-linear or association model tradition (Goodman, 1983). It is important to recognize that ordered regression models are equivalent to log-linear models in certain cases. For example, a parsimonious adjacent category logit model with one ordinal independent variable that imposes the parallel regression assumption is equivalent to the uniform association model for analyzing contingency tables (Clogg & Shihadeh, 1994, p.148). In the case of a model with a nominal-level independent variable and the parallel regression assumption, the model is equivalent to the column effects association model (Clogg & Shihadeh, 1994, p.148). Agresti (2010) provides additional examples of the connections between adjacent category models and log-linear models and is an excellent resource for those interested in understanding a broad range of ordinal data analysis techniques.

The adjacent model does not suffer from the internal contradiction in cumulative models that makes negative predicted probabilities possible, and it does not require a sequential data-generating process. Additionally, it is the only class of ordered models that allows one to focus on comparisons of individual categories, which may be of theoretical interest to the analyst. Therefore, it is surprising that it is a relatively unknown method for analyzing ordinal outcomes in the applied literature. For a variable such as self-rated health, the categories may not be sufficiently meaningful to warrant using an adjacent category model, but sequential outcomes (such as educational attainment) would be good candidates for these models given that each category reflects a meaningful stage in a certain process and the focus is on comparisons of adjacent stages.

For a four-category outcome, the adjacent category ordered regression model connects the original, ordinal outcome, y, to the set of binary outcomes, y_1 to y_3, with the following measurement equations (Fullerton & Xu, 2016, p.7–8):

$$y_1 = \begin{cases} 1 & \text{if y} = 1 \quad | \quad y = 1 \text{ or } 2 \\ 0 & \text{if y} = 2 \quad | \quad y = 1 \text{ or } 2 \end{cases} \tag{14}$$

$$y_2 = \begin{cases} 1 & \text{if y} = 2 \quad | \quad y = 2 \text{ or } 3 \\ 0 & \text{if y} = 3 \quad | \quad y = 1 \text{ or } 3 \end{cases} \tag{15}$$

$$y_3 = \begin{cases} 1 & \text{if y} = 3 \quad | \quad y = 3 \text{ or } 4 \\ 0 & \text{if y} = 4 \quad | \quad y = 3 \text{ or } 4 \end{cases} \tag{16}$$

The equation for the conditional probability in the adjacent category ordered regression model with proportional odds is (Fullerton & Xu, 2016, p.40)

$$\Pr(y = m|y = m \text{ or } y = m + 1, \mathbf{x}) =$$

$$\Pr\left(y^* < \tau_m|y = m \text{ or } y = m + 1, \mathbf{x}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta}) \tag{17}$$

Adjacent Category Models Without Proportional Odds

If proportional odds is not a reasonable assumption for some variables, we may relax it for a subset with the partial

[2] For a discussion of the constrained partial model, see Fullerton and Xu (2016, p.65).

adjacent category model or for the entire model with the non-parallel adjacent category model. The equation for the conditional probability in the unconstrained partial adjacent category ordered regression model is (Fullerton & Xu, 2016, p.71)

$$\Pr(y = m|y = m \,\text{or}\, y = m + 1, \mathbf{x}, \boldsymbol{\omega}) =$$

$$\Pr\left(y^* < \tau_m \middle| y = m \,\text{or}\, y = m + 1, \mathbf{x}, \boldsymbol{\omega}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta} - \boldsymbol{\omega}\boldsymbol{\eta}_m) \tag{18}$$

The non-parallel adjacent category logit model is a re-parameterized version of the more familiar multinomial logit model for nominal outcomes. The cutpoint equations are expressed in terms of adjacent comparisons, but using the properties of natural logs one can easily express them using a common baseline category. The equation for the conditional probability in the non-parallel adjacent category ordered logit model is (Fullerton & Xu, 2016, p.98)

$$\Pr(y = m|y = m \,\text{or}\, y = m + 1, \mathbf{x}) =$$

$$\Pr\left(y^* < \tau_m \middle| y = m \,\text{or}\, y = m + 1, \mathbf{x}\right) = F(\tau_m - \mathbf{x}\boldsymbol{\beta}_m) \tag{19}$$

One additional point to take into consideration with the adjacent class of ordered regression models is the independence of irrelevant alternatives (IIA) assumption (see Long, 1997). As in multinomial logit, we must assume that the introduction of additional categories would not affect the ratio for the probabilities of two existing categories. Unfortunately, existing tests of the IIA assumption suffer from problems associated with size properties of the formal tests (see Cheng & Long, 2007). Therefore, one should only use the adjacent category ordered models if a reasonable case can be made on conceptual grounds that the outcome categories are distinct and independently weighed by respondents (McFadden, 1973).

Ordinal Independent Variables

Ordinal independent variables are handled in a variety of ways in the applied literature often with little to no justification given for decisions made in the recoding process as with ordinal dependent variables. The three most common approaches are to retain the original variables (i.e., assume an interval-ratio level of measurement), represent it with a set of binary variables for each category (leaving out one category to serve as the reference), or collapse categories to create a single, binary variable. The only way to retain the ordinal nature of the variable is to utilize the second option, a set of binary variables. As a set, the binary variables actually assume a nominal level of measurement for the variable because there are no restrictions imposed on the coefficients.[3] However, what analysts may be unaware of is that the set of binary variables effectively acts as a "potentially" ordinal variable. It allows for an ordinal pattern to emerge just as a non-parallel adjacent category model is a nominal method that allows for ordinal patterns to emerge, which is suitable for nominal outcomes given that it is equivalent to a multinomial logit model. Collapsing the M categories into two assumes that if the variable is ordinal, there is only one meaningful threshold. In most cases, there is no theoretical rationale or empirical justification for making that assumption.

In the example we present in the next section, we will illustrate how to empirically distinguish between different ways of handling ordinal independent variables and how the continuous, binary set, and single binary options are interrelated and easily compared using familiar tests for nested models (Wald and LR). The connection between a set of binary variables and a single binary variable is the most obvious. If there are three binary coefficients corresponding to a four-category ordinal independent variable, then the model with a single binary variable (e.g., 1–2 vs. 3–4) may be expressed as a constrained form of the model with a set of binary variables for categories 1 through 3 (4 = reference) with the following constraints:

$$\text{Constraint \#1}: \quad \beta_1 = \beta_2 \tag{20}$$

$$\text{Constraint \#2}: \quad \beta_3 = 0 \tag{21}$$

However, we can also express the model with the original variable (ordinal but treated as continuous) as a constrained form of the binary set model with the following constraints:

$$\text{Constraint \#3}: \quad \beta_1 = \beta_2 = \beta_2 - \beta_3 \tag{22}$$

$$\text{Constraint \#4}: \quad \beta_2 - \beta_3 = \beta_3 \tag{23}$$

A statistically significant Wald or LR test of either set of constraints suggests that the binary set model is preferable over the simpler continuous or single binary versions of the model. One can also compare the three models using information criteria measures such as the AIC and BIC. These empirical tests are relatively simple yet under-utilized tools for making recoding decisions in a more systematic manner for ordinal independent variables. Ideally, theoretical concerns would guide recoding decisions, but in their absence (or perhaps in tandem), these empirical measures should be useful for health researchers.

[3] Espinosa and Hennig (2019) developed an ordinal model that allows one to impose a monotonicity constraint on ordinal independent variables via constrained MLE (p.872). The models we present in this study do not impose this inequality constraint. Ordinal patterns may emerge for groups of binary independent variables, but they are not imposed in the models.

Example

Data and Methods

To illustrate the cumulative method using an empirical example, we use data from the 2018 General Social Survey (GSS). The GSS is a telephone survey of American adults that has been conducted biannually by NORC at the University of Chicago since 1972. The GSS generally focuses on attitudinal and behavioral outcomes, and in recent years, the GSS has included questions on both job insecurity and health.

First, the dependent variable across all models is general *self-rated health*. The item in this data source is specifically worded as, "Would you say that in general your health is excellent, very good, good, fair, or poor?" with the five response options embedded in the question. We leave this variable as ordinal, and only re-code the variable into four categories by combining the "fair" and "poor" categories due to the low number of responses for the "poor" category.[4] Note that for this variable, higher values indicate poorer health. General self-rated health is an outcome that is well-suited for the cumulative ordered regression model because it is plausible that there is an underlying continuous measure of perceived health, it is not the result of a sequential decision-making process, and the individual categories are not of particular substantive interest.[5]

Our main independent variable of interest is another ordinal measure for *perceived job insecurity*, which was only asked of respondents who are employed. This item is specifically worded as, "Thinking about the next 12 months, how likely do you think it is that you will lose your job or be laid off–very likely, fairly likely, not too likely, or not at all likely?" with the four responses options embedded in the question. We code this variable in three different ways in order to illustrate differences for the purpose of illustration across the various coding schemes for ordinal independent variables as well. First, we include a measure of the variable in its original, ordinal coding and treat it as continuous where higher values indicate a higher degree of job insecurity. However, we combine two of the categories, "very likely" and "fairly likely" as the "very likely" category had a small number of responses.[6] Thus, the coding scheme is 1 = not at all likely, 2 = not too likely, and 3 = very/fairly likely. Next, we take these same three categories and re-code them into a set of binary variables, with "not at all likely" as the reference category. Finally, we dichotomize the variable into a single binary variable for job insecurity, where 1 = very/fairly likely and 0 = not at all likely/not too likely. The vast majority (71%) of respondents reports the lowest level of job insecurity, 22% have a medium level of job insecurity, and the remaining 7% have a high level of job insecurity. We also include a number of socio-demographic control variables (see Table 1).

We present the results from a series of ordinal methods employing a variety of model specifications in order to compare across modeling differences in the dependent variable, as well as the three codings of the ordinal independent variable. First, we present a series of parallel cumulative ordered logistic regression models with the parallel regression assumption (see Table 2). Further, using the Brant test (Brant, 1990), we tested whether or not the parallel regression assumption is reasonable across all variables. We found that the assumption was violated for education. Thus, in a second set of models, we present a series of partial cumulative ordered logistic regression models with the parallel regression assumption relaxed for education (see Table 3). Finally, for comparison purposes, we present a set of non-parallel cumulative ordered logistic regression models with the assumption relaxed for every variable (see Table 4).[7] We present average discrete change (ADC) coefficients, which reflect the average change in the predicted probability given a change from 0 to 1 in the binary variables and a standard deviation change from the mean in the continuous variables (see Long & Freese, 2014, p.343).

Results

In Table 2, we present results from the parallel version of the cumulative ordered regression model of self-rated health, which imposes the parallel regression (or "proportional odds") assumption for every variable. We will focus the

[4] The "poor" category only had 1.53% of the overall sample. The LR test for combining the "fair" and "poor" categories was not significant ($X^2 = 9.798$, $p = 0.367$), providing an empirical justification for this approach.

[5] For illustrations of stage and adjacent models, see Fullerton and Xu (2016, 2018) and Bauldry et al. (2018). We also present examples for these approaches in the online appendix.

[6] The "fairly likely" category only had 4.36% of the sample, and the "very likely" category only had 2.59% of the sample. The LR tests for combining these categories was not significant ($X^2 = 6.688$, $p = 0.571$).

[7] We used the gologit2 (Williams 2006) command in Stata to estimate the cumulative models. Although the focus in this example is on cumulative models, we also estimated stage and adjacent models for comparison purposes. The results are substantively similar to the cumulative results. We used the gencrm command (Bauldry et al., 2018) and the regular mlogit command in Stata to estimate the stage and adjacent models, respectively. See the online appendix for the results and details regarding estimation using Stata.

Table 1 Descriptive statistics for variables in cumulative models of self-rated health

Variable Name	Mean/%	St. Dev	Min	Max	Description
Dependent variables					
Self-rated health	20.73	--	1	4	General self-rated health (1 = Excellent, 4 = Fair/Poor)
Excellent	32.74	--			
Very Good	31.92	--			
Good	14.61	--			
Fair/Poor					
Independent variables					
Job insecurity:					
Job insecurity (Cont.)	1.36	0.61	1	3	How likely to lose one's job (1 = Not at all likely, 2 = Not too likely, 3 = Fairly/very likely)
Job insecurity (set binary)					
Not at all likely	71.26	--	0	1	1 = Not at all likely, 0 = else
Not too likely	21.79	--	0	1	1 = Not too likely, 0 = else
Fairly/very likely	6.95	--	0	1	1 = Fairly/very likely, 0 = else
Job insecure (binary)	6.95	--	0	1	1 = Fairly/very likely, 0 = else
Age	44.91	14.22	18	86	Age in years
Female	53.00	--	0	1	1 = Female, 0 = else
Race					
White (ref.)	62.54	--	0	1	1 = White, 0 = else
Black	15.78	--	0	1	1 = Black, 0 = else
Latino	16.73	--	0	1	1 = Latino, 0 = else
Other	4.95	--	0	1	1 = Other race, 0 = else
Education	14.14	2.85	0	20	Years of schooling
Log Income	10.22	0.91	5.43	11.69	Log of household income
Part Time	16.73	--	0	1	1 = Part time, 0 = else

N = 849. Data come from the 2018 General Social Survey (GSS)

discussion on how the relationship between job insecurity and health varies by the measurement of job insecurity as continuous, a set of binary variables for each category, or a single binary variable (high vs. low to medium insecurity). The ADC coefficients for the continuous measure of job insecurity in Model 1 reveal that job insecurity, conceptualized as a continuous, negative resource (or stressor), is significantly associated with poorer health. On average, a standard deviation increase in job insecurity is associated with a decrease of 0.032 in the predicted probability of reporting excellent health and an increase of 0.027 in the predicted probability of reporting fair or poor health. If we conceptualize job insecurity as group membership (low, medium, or high insecurity groups), then the results in Models 2 and 3 allow us to model the relationship between insecurity and health accordingly and reveal a similar pattern. The groups with the highest levels of insecurity tend to have the poorest self-rated health. The likelihood ratio (LR) tests indicate that the model with a set of binary variables for job insecurity provides a better fit than a single binary measure, but the model with job insecurity as a continuous measure provides an even better fit. The AIC and BIC results confirm this. If there is a theoretical rationale for including job insecurity as a continuous measure, then these results suggest that this is a reasonable coding decision even though technically job insecurity is an ordinal measure. The less constrained results in Model 2 follow a monotonic pattern that is more efficiently approximated with a continuous term. Therefore, these results provide an empirical justification for treating insecurity as continuous.

The other focal variable in this example is education, which is positively related to self-rated health. On average, a standard deviation increase in years of schooling is associated with an increase of 0.052 in the predicted probability of reporting excellent health and a decrease of 0.033 in the predicted probability of reporting fair or poor health. However, the Brant test (results available upon request) indicates that the raw coefficients for education vary in a statistically significant way across cutpoint equations. This does not guarantee that the violation of the assumption is substantively important (Fullerton & Xu, 2016), but it does signal that we should investigate this possibility by estimating partial and/or non-parallel cumulative models. The Brant test (Brant, 1990) was non-significant for the remaining variables, which suggests it is a reasonable assumption on empirical grounds.

Table 2 Average discrete change (ADC) coefficients and (Z-ratios) from cumulative ordered logit models of poor self-rated health with the parallel regression assumption (1 = Excellent, 2 = Very Good, 3 = Good, 4 = Fair/Poor)

Variable name	(1) Continuous				(2) Set of binary				(3) Binary			
	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor
Job insecurity (Cont.)	−0.032***	−0.018**	0.022***	0.027**								
	(−3.546)	(−3.034)	(3.598)	(3.155)								
Job insecurity (Ref. = Not at all likely)												
Not too likely					−0.050*	−0.031*	0.034**	0.046*				
					(−2.559)	(−2.010)	(2.704)	(2.090)				
Fairly/Very likely					−0.091***	−0.071*	0.057***	0.105*				
					(−3.509)	(−2.382)	(4.620)	(2.365)				
Job insecurity (binary)									−0.081**	−0.060*	0.052***	0.089*
									(−2.979)	(−2.088)	(3.635)	(2.108)
Age	−0.021*	−0.011*	0.015*	0.017*	−0.021*	−0.011*	0.015*	0.017*	−0.021*	−0.011*	0.015*	0.017*
	(−2.206)	(−1.968)	(2.225)	(2.037)	(−2.205)	(−1.968)	(2.224)	(2.036)	(−2.193)	(−1.960)	(2.212)	(2.026)
Female	−0.001	0.000	0.001	0.001	−0.001	0.000	0.001	0.001	−0.001	0.000	0.000	0.000
	(−0.039)	(0.039)	(0.039)	(0.039)	(−0.039)	(−0.039)	(0.039)	(0.039)	(−0.025)	(−0.025)	(0.025)	(0.025)
Race												
Black	−0.015	−0.008	0.011	0.012	−0.015	−0.008	0.011	0.012	−0.017	−0.009	0.012	0.014
	(−0.572)	(−0.524)	(0.575)	(0.539)	(−0.571)	(−0.524)	(0.574)	(0.538)	(−0.642)	(−0.584)	(0.646)	(0.601)
Latino	−0.062**	−0.041*	0.042**	0.061*	−0.062**	−0.041*	0.042**	0.061*	−0.065**	−0.044*	0.043***	0.065*
	(−2.863)	(−2.155)	(3.140)	(2.221)	(−2.861)	(−2.154)	(3.137)	(2.220)	(−3.040)	(−2.273)	(3.376)	(2.329)
Other	0.029	0.011	−0.020	−0.020	0.029	0.011	−0.021	−0.020	0.030	0.012	−0.021	−0.021
	(0.592)	(0.725)	(−0.596)	(−0.656)	(0.593)	(0.727)	(−0.597)	(−0.657)	(0.610)	(0.753)	(−0.615)	(−0.678)
Education	0.052***	0.017***	−0.036***	−0.033***	0.052***	0.017***	−0.036***	−0.033***	0.051***	0.017***	−0.036***	−0.033***
	(4.050)	(4.793)	(−4.153)	(−4.627)	(4.038)	(4.776)	(−4.140)	(−4.610)	(3.993)	(4.727)	(−4.093)	(−4.553)
Log income	0.065***	0.020***	−0.045***	−0.040***	0.065***	0.020***	−0.045***	−0.040***	0.068***	0.020***	−0.047***	−0.041***
	(4.624)	(5.376)	(−4.840)	(−5.390)	(4.624)	(5.374)	(−4.841)	(−5.390)	(4.793)	(5.478)	(−5.040)	(−5.610)
Part time	0.032	0.012	−0.023	−0.022	0.032	0.012	−0.023	−0.022	0.029	0.011	−0.021	−0.020
	(1.074)	(1.342)	(−1.083)	(−1.200)	(1.074)	(1.342)	(−1.084)	(−1.201)	(0.987)	(1.205)	(−0.994)	(−1.093)
AIC	2185.687				2187.685				2191.042			
BIC	2242.616				2249.358				2247.970			
LR test					0.000[a]				5.360*[b]			

N = 849. Z-ratios in parentheses. The average discrete change coefficients reflect the average change in the predicted probability associated with a change from 0 to 1 in the variable for binary variables and a standard deviation change from the mean for continuous variables. a. LR test of Model 2 compared to Model 1. b. LR test of Model 3 compared to Model 2

*p < .05, **p < .01, ***p < .001

Table 3 Average discrete change (ADC) coefficients and (Z-ratios) from cumulative ordered logit models of poor self-rated health with the parallel regression assumption relaxed for education (1 = Excellent, 2 = Very good, 3 = Good, 4 = Fair/Poor)

Variable name	(1) Continuous				(2) Set of binary				(3) Binary			
	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor
Job insecurity (Cont.)	−0.032***	−0.017**	0.023***	0.026**								
	(−3.498)	(−2.942)	(3.511)	(3.114)								
Job insecurity (Ref. = Not at all likely)												
Not too likely					−0.051**	−0.030*	0.037**	0.045*				
					(−2.596)	(−2.003)	(2.670)	(2.116)				
Fairly/Very likely					−0.090***	−0.067*	0.062***	0.096*				
					(−3.394)	(−2.271)	(4.009)	(2.307)				
Job insecurity (binary)									−0.080**	−0.057*	0.056**	0.081*
									(−2.861)	(−1.980)	(3.203)	(2.043)
Age	−0.021*	−0.010	0.015*	0.016*	−0.021*	−0.010	0.015*	0.016*	−0.021*	−0.010	0.015*	0.016*
	(−2.164)	(−1.910)	(2.166)	(1.999)	(−2.159)	(−1.906)	(2.161)	(1.995)	(−2.144)	(−1.896)	(2.146)	(1.982)
Female	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(−0.012)	(−0.012)	(0.012)	(0.012)	(−0.012)	(−0.012)	(0.012)	(0.012)	(−0.006)	(−0.006)	(0.006)	(0.006)
Race												
Black	−0.015	−0.007	0.011	0.012	−0.015	−0.007	0.011	0.012	−0.017	−0.008	0.013	0.013
	(−0.563)	(−0.513)	(0.562)	(0.531)	(−0.563)	(−0.514)	(0.562)	(0.531)	(−0.636)	(−0.575)	(0.635)	(0.596)
Latino	−0.059**	−0.037*	0.042**	0.054*	−0.059**	−0.037*	0.042**	0.054*	−0.062**	−0.040*	0.045**	0.057*
	(−2.651)	(−1.981)	(2.767)	(2.086)	(−2.643)	(−1.976)	(2.758)	(2.081)	(−2.835)	(−2.101)	(2.987)	(2.201)
Other	0.031	0.011	−0.022	−0.020	0.031	0.011	−0.022	−0.020	0.032	0.011	−0.023	−0.020
	(0.601)	(0.755)	(−0.611)	(−0.666)	(0.597)	(0.750)	(−0.607)	(−0.662)	(0.613)	(0.776)	(−0.623)	(−0.682)
Education	0.026	0.039*	−0.015	−0.050***	0.026	0.039*	−0.015	−0.050***	0.026	0.039*	−0.014	−0.050***
	(1.559)	(2.297)	(−1.065)	(−6.276)	(1.560)	(2.297)	(−1.066)	(−6.263)	(1.542)	(2.277)	(−1.039)	(−6.196)
Log income	0.067***	0.018***	−0.047***	−0.038***	0.067***	0.018***	−0.047***	−0.038***	0.070***	0.019***	−0.049***	−0.040***
	(4.636)	(4.870)	(−4.945)	(−5.385)	(4.636)	(4.870)	(−4.944)	(−5.385)	(4.802)	(4.905)	(−5.147)	(−5.600)
Part time	0.030	0.011	−0.022	−0.020	0.030	0.011	−0.022	−0.020	0.028	0.010	−0.020	−0.018
	(0.998)	(1.242)	(−1.014)	(−1.104)	(0.997)	(1.241)	(−1.013)	(−1.104)	(0.911)	(1.107)	(−0.923)	(−0.999)
AIC	2180.432				2182.430				2185.906			
BIC	2246.849				2253.591				2252.323			
LR test					0.000[a]				5.480*[b]			

$N = 849$. Z-ratios in parentheses. The average discrete change coefficients reflect the average change in the predicted probability associated with a change from 0 to 1 in the variable for binary variables and a standard deviation change from the mean for continuous variables. a. LR test of Model 2 compared to Model 1. b. LR test of Model 3 compared to Model 2

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 4 Average discrete change (ADC) coefficients and (Z-ratios) from cumulative ordered logit models of poor self-rated health with the parallel regression assumption relaxed for all variables (1 = Excellent, 2 = Very good, 3 = Good, 4 = Fair/Poor)

Variable name	(1) Continuous				(2) Set of binary				(3) Binary			
	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor	Excellent	Very good	Good	Fair/Poor
Job insecurity (Cont.)	−0.045**	−0.008	0.037*	0.016								
	(−3.208)	(−0.494)	(2.297)	(1.428)								
Job insecurity (Ref. = Not at all likely)												
Not too likely					−0.073**	−0.033	0.106**	−0.001				
					(−2.693)	(−0.917)	(2.744)	(−0.021)				
Fairly/Very likely					−0.116**	−0.027	0.060	0.083				
					(−2.928)	(−0.437)	(0.970)	(1.626)				
Job insecurity (binary)									−0.106*	−0.011	0.035	0.082
									(−2.450)	(−0.175)	(0.558)	(1.650)
Age	−0.030*	−0.007	0.037*	0.000	−0.030*	−0.006	0.036*	0.000	−0.031*	−0.005	0.035*	0.000
	(−2.312)	(−0.415)	(2.255)	(−0.014)	(−2.299)	(−0.358)	(2.188)	(−0.019)	(−2.312)	(−0.315)	(2.118)	(0.027)
Female	−0.014	−0.007	0.056	−0.035	−0.014	−0.007	0.057	−0.035	−0.013	−0.007	0.056	−0.036
	(−0.523)	(−0.223)	(1.766)	(−1.801)	(−0.540)	(−0.226)	(1.774)	(−1.787)	(−0.470)	(−0.224)	(1.739)	(−1.847)
Race												
Black	0.021	−0.038	−0.042	0.059	0.021	−0.038	−0.038	0.054	0.014	−0.033	−0.039	0.058
	(0.519)	(−0.861)	(−0.916)	(1.521)	(0.521)	(−0.854)	(−0.838)	(1.431)	(0.351)	(−0.744)	(−0.859)	(1.513)
Latino	−0.065	−0.017	0.016	0.066	−0.066	−0.017	0.015	0.067	−0.072*	−0.015	0.017	0.070
	(−1.908)	(−0.398)	(0.364)	(1.784)	(−1.921)	(−0.384)	(0.345)	(1.801)	(−2.203)	(−0.350)	(0.373)	(1.868)
Other	0.010	0.083	−0.122	0.030	0.008	0.079	−0.105	0.018	0.009	0.080	−0.110	0.021
	(0.155)	(1.072)	(−1.696)	(0.449)	(0.133)	(1.018)	(−1.483)	(0.290)	(0.150)	(1.028)	(−1.527)	(0.327)
Education	0.025	0.043*	−0.019	−0.049***	0.025	0.044*	−0.021	−0.048***	0.024	0.044*	−0.020	−0.048***
	(1.429)	(2.309)	(−1.343)	(−5.879)	(1.442)	(2.349)	(−1.485)	(−5.657)	(1.365)	(2.340)	(−1.342)	(−5.608)
Log income	0.067***	0.013	−0.039*	−0.041***	0.067***	0.013	−0.039*	−0.041***	0.072***	0.013	−0.043**	−0.042***
	(3.327)	(0.661)	(−2.447)	(−4.180)	(3.323)	(0.634)	(−2.429)	(−4.184)	(3.518)	(0.637)	(−2.684)	(−4.214)
Part time	0.015	0.052	−0.061	−0.006	0.014	0.051	−0.058	−0.008	0.008	0.053	−0.054	−0.007
	(0.349)	(1.122)	(−1.458)	(−0.187)	(0.330)	(1.107)	(−1.364)	(−0.261)	(0.184)	(1.153)	(−1.258)	(−0.246)
AIC	2196.315				2199.375				2203.512			
BIC	2338.637				2355.929				2345.834			
LR test					2.940[a]				10.140*[b]			

N = 849. Z-ratios in parentheses. The average discrete change coefficients reflect the average change in the predicted probability associated with a change from 0 to 1 in the variable for binary variables and a standard deviation change from the mean for continuous variables. a. LR test of Model 2 compared to Model 1. b. LR test of Model 3 compared to Model 2

$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

We also present the results from a partial cumulative model of self-rated health that relaxes the parallel assumption for education but retains it for the remaining variables in Table 3. Relaxing the parallel regression assumption for education has a substantial impact on the results, which is not always the case. The logit coefficients become larger in the second and third cutpoint equations. As a result, relaxing the parallel assumption results in a weaker relationship with the best level of self-rated health ("excellent") and a stronger relationship with the poorest level of self-rated health ("fair" or "poor"). This adds some nuance to the relationship between education and health. More formal schooling is associated with a lower probability of reporting poor health, but this does not automatically translate into a higher probability of being in excellent health. Thus, how important education is as a predictor of health depends on where one is focused in the distribution of health. Thus, relaxing the parallel assumption for education in this case provides more nuance to our understanding of these relationships as compared to the results presented in Table 2. As such, there is an empirical and theoretical reason to use this model over the set of models with the parallel regression assumption for all variables.

We present the results from a non-parallel cumulative model of self-rated health that relaxes the parallel assumption for every variable in Table 4. We see essentially the same pattern for education as in the partial model. However, now the parallel regression assumption is also relaxed for the remaining variables. For job insecurity, its association with self-rated health is stronger for excellent health and weaker for poor/fair health without the parallel regression assumption. However, the added coefficients (i.e., cutpoint equation-specific slopes) do not improve the model fit according to the AIC and BIC. The partial model is preferred over the non-parallel model regardless of how job insecurity is measured. On balance, the results in Tables 2, 3, and 4 suggest that a partial cumulative ordered regression model relaxing the parallel assumption for education would be an effective model to use. It is more efficient than the non-parallel model and allows us to uncover important asymmetries in the relationship between education and health that are suppressed in the parallel model.

Alternative approaches

In order to illustrate the utility of ordered regression models, we also compared the results from Tables 2, 3, and 4 with results from two common alternative approaches to analyzing ordinal outcomes: (1) a linear regression model (LRM) treating the dependent variable as continuous and (2) a binary logistic regression model treating the dependent variable as dichotomous (1 = poor/fair health, 0 = else). For the LRM, our substantive conclusions are similar but there is one important difference. The LRM over-estimates the strength of the association between education and health in some parts of the outcome distribution. More specifically, the ordered regression models relaxing the parallel assumption for education reveal that education has a strong, negative association with the odds of fair or poor health but no association with the odds of excellent health (compared to the remaining categories). Thus, the positive relationship between education and health is primarily a buffering effect against poor health. Highly educated respondents have a lower odds of experiencing poor health but factors other than education, such as job insecurity and income, help explain differences at the other end of the health spectrum.

For the binary approach, the differences are even more substantial. In the logit model, job insecurity is *not* significantly associated with poor/fair self-rated health. Thus, this modeling strategy would change our conclusions. The decision to collapse outcome categories in order to utilize a binary regression model would lead one to draw different theoretical implications. Researchers routinely make this coding choice because binary regression models are viewed as simpler than ordered models, but these results remind us that simpler models sometimes lead us to the wrong conclusions. Additionally, coding decisions such as this one have important theoretical implications (i.e., viewing self-rated health as a binary measure changes its conceptual meaning), which researchers should address. The full results from these alternative models are available in the online appendix.

Conclusion

Although references to the "ordinal logit/probit model" are replete in the literature, there is actually a diverse array of different models that vary depending on the probability of interest and the application of the parallel regression assumption (Fullerton, 2009). Ordered regression models are an attractive alternative for ordinal outcomes to the linear regression model, which imposes the unrealistic assumption of equal spacing between categories, and the binary regression model, which requires arbitrarily collapsing categories and potentially losing valuable information. We also emphasize throughout that an integration of theoretical and empirical concerns will help guide researchers to the appropriate ordered regression model (cumulative/stage/adjacent and parallel/partial/non-parallel) and treatment of ordinal independent variables. The empirical example in this study, general self-rated health, is ideally suited for a cumulative model because it does not result from a sequential process and the individual categories are not of interest. However, there are other health outcomes that would be better suited to alternative ordered models. For example, stage models are the preferred model for analyzing the stages of a disease, and the adjacent model may be ideal for a measure such as body mass index (BMI) that is grouped into meaningful intervals. A more systematic approach to ordinal data analysis is critical to improving the practice of quantitative research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11121-021-01302-y.

Compliance with Ethical Standards

Ethical Approval The OSU Institutional Review Board granted the study exempt status.

Informed Consent Not applicable.

Conflict of Interest The authors declare no competing interests.

References

Agresti, A. (2010). *Analysis of ordinal categorical data* (2nd ed.). Wiley

Amemiya, T. (1981). Qualitative response models: a survey. *Journal of Economic Literature,* 19, 1483–1536

Bauldry, S., Xu, J., & Fullerton, A. S. (2018). Gencrm: a new command for generalized continuation-ratio models. *Stata Journal,* 18, 924–936

Bourdieu, P., Chambordeon, J., & Passeron, J. (1991). *The craft of sociology: Epistemological preliminaries*. Walter de Gruyter. Berlin, Germany.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics,* 46, 1171–1178

Buis, M. L. (2011). The consequences of unobserved heterogeneity in a sequential logit model. *Research in Social Stratification and Mobility,* 29, 247–262

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: a tutorial. *Advances in Methods and Practices in Psychological Science,* 2, 77–101

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

Cheng, S., & Long, J. S. (2007). Testing for IIA in the multinomial logit model. *Sociological Methods and Research,* 35, 583–600

Clogg, C. C., & Shihadeh, E. S. (1994). *Statistical models for ordinal variables*. Sage

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B,* 34, 187–220

Espinosa, J., & Hennig, C. (2019). A constrained regression model for an ordinal response with ordinal predictors. *Statistics and Computing,* 29, 869–890

Fienberg, S. E. (1980). *The analysis of cross-classified categorical data* (2nd ed.). MIT Press

Fullerton, A. S. (2009). A conceptual framework for ordered logistic regression models. *Sociological Methods and Research,* 38, 306–347

Fullerton, A. S., & Dixon, J. C. (2009). Racialization, asymmetry, and the context of welfare attitudes in the American states. *Journal of Political and Military Sociology,* 37, 95–120

Fullerton, A. S., & Xu, J. (2012). The proportional odds with partial proportionality constraints model for ordinal response variables. *Social Science Research,* 41, 182–198

Fullerton, A. S., & Xu, J. (2016). *Ordered regression models: Parallel, partial, and non-parallel alternatives*. Chapman & Hall/CRC Press

Fullerton, A. S., & Xu, J. (2018). Constrained and unconstrained partial adjacent category logit models for ordinal response variables. *Sociological Methods and Research,* 47, 169–206

Goodman, L. A. (1983). The analysis of dependence in cross-classifications having ordered categories, using log-linear models for frequencies and log-linear models for odds. *Biometrics,* 39, 149–160

Greene, W. H., & Hensher, D. A. (2010). Modeling ordered choices: A primer. Cambridge, UK: Cambridge University Press.

Hedeker, D. R., Mermelstein, R. J., & Weeks, K. A. (1999). The thresholds of change model: An application to analyzing stages of change data. *Annals of Behavioral Medicine,* 21, 61–70

Lieberson, S. (1985). *Making it count*. University of California Press

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology,* 140, 1–55

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage

Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using Stata* (3rd ed.). College Station, TX: Stata Press

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press

Mare, R. D. (2011). Introduction to symposium on unmeasured heterogeneity in school transition models. *Research in Social Stratification and Mobility,* 29, 239–245

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society Series B,* 42, 109–142

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology,* 4, 103–120

Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. *Journal of General Psychology,* 142, 71–89

Peterson, B., & Harrell, F. E., Jr. (1990). Partial proportional odds models for ordinal response variables. *Applied Statistics,* 39, 205–217

Tucker, G., Adams, R., & Wilson, D. (2013). Observed agreement problems between sub-scales and summary components of the SF-36 Version 2 — An alternative scoring method can correct the problem. *PLoS ONE* , 8, e61191

Tucker, G., Adams, R., & Wilson, D. (2014). Results from several population studies show that recommended scoring methods of the SF-36 and the SF-12 may lead to incorrect conclusions and subsequent health decisions. *Quality of Life Research,* 23, 2195–2203

Tutz, G. (1991). Sequential models in categorical regression. *Computational Statistics & Data Analysis,* 11, 275–295

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal,* 6, 58–82

Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *Sociological Methods and Research,* 37, 531–559

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *Journal of Mathematical Sociology,* 40, 7–20

Xie, Y. (2011). Values and limitations of statistical models. *Research in Social Stratification and Mobility,* 29, 343–349

Xu, J., Bauldry, S., & Fullerton, A. S. (2019). Bayesian approaches to assessing the parallel lines assumption in cumulative ordered logit models. *Sociological Methods and Research (In Press)*. https://doi.org/10.1177/0049124119882461