Order Matters (?): Alternatives to Conventional Practices for Ordinal Categorical Response Variables

Bradford S. Jones
Department of Political Science
University of Arizona
Tucson, AZ 85721
bsjones@email.arizona.edu

Chad Westerland
Department of Political Science
University of Arizona
Tucson, AZ 85721
cwesterl@email.arizona.edu

April 26, 2006

1 Introduction

Response variables that can be construed of as ordered are commonplace in social science research. Likert-type items, scales for social spending, and ratings indicators are but a few of the myriad of ordinal items commonly used by social science researchers. Yet despite the prevalence of such items, standard applied methods often: 1) fail to exploit information found in ordinal items; 2) impose assumptions that frequently do not hold; 3) treat the ordinal item as if it were continuous; and/or 4) collapse categories on an ordinal scale to create a binary indicator. Each of these approaches are prevalent in applied social science research, particularly in political science and sociological applications.

In this paper, we discuss some of the implications of these approaches and then consider some alternative modeling strategies for ordinal categorical data. We note at the start that most of the models considered herein are straightforward applications of fairly standard cumulative link models. It is just the case that these models are only rarely considered and applied in the political science literature. Most of our discussion entails consideration of a variety of logit-type regression models. In this discussion, we note that common "casual" assumptions about modeling approaches with ordinal data are frequently wrong and are usually made out of convenience. Moreover, properties of seemingly appropriate statistical models for ordinal data, for example, the proportional odds model, are rarely acknowledged and are almost never tested to see if they hold. To illustrate some of the models, we present several substantive applications using national survey data on racial attitudes.

2 The Proportional Odds Model

Ordinal variables are commonplace in social science data sets. Likert-type scales, where survey respondents are asked to place themselves on a semantically balanced scale, are typically used by social scientists to measure attitudes or preferences over some issue. Similarly, scales recording intensity of opinion by asking respondents to place themselves on an n-point scale are also prevalent in frequently used surveys, for example the National Election Study (NES) or the General Social Survey (GSS). When these kinds of items are used as response variables in regression-type models, special issues emerge. And while these issues are generally well known—most good statistics texts on categorical data discuss them (c.f. Agresti 1996, 2002; Long 1997; Powers and Xie 2000)—they are often outright ignored in applied work (as we discuss below).

To motivate what follows, it is worth briefly discussing some of the implications of ordinal categorical response variables.[1] To fix ideas, imagine that Y is a four-point Likert item measuring attitudes toward government spending with scale scores "greatly decrease (spending)"=1, "decrease"=2, "increase"=3, and "greatly increase"=4. In a regression-type setting, a researcher may be interested in modeling Y as a function of some covariates $\mathbf{x} = (x_1, x_2, \ldots, x_k)$. How some variable x relates to the scale, for example, "moving up" or "moving down" the scale, is presumably of interest. In many social science settings, models for ordinal scales are motivated by random utility (McKelvey and Zavoina 1975).[2] Under this perspective, the observed scale scores on Y are assumed to be discretized measurements on an otherwise continuous, but latent, response variable Y^*. As such, researchers commonly interpret models with ordinal response variables probabilistically: regression parameters give information about how the probability (or odds) for some level of government spending increase or decrease with respect to changes in x. This kind of motivation, then, leads to consideration of

$$Y_i^* = \alpha + \mathbf{x}'\beta + \epsilon_i. \qquad (1)$$

Because the response variable in model (1) is unobserved, Y^* is connected to Y through a series of "cut points" such that

$$Y = 1 \quad \equiv \quad Y^* \leq \alpha_1$$

[1] Part of the following discussion is adapted from Jones and Sobel (2000).

[2] Though it is important to note that the utility perspective *need not* be adopted in order to motivate the models discussed herein (Jones and Sobel 2000; Powers and Xie 2000).

$$
\begin{aligned}
Y = 2 &\equiv \alpha_1 < Y^* \leq \alpha_2 \\
Y = 3 &\equiv \alpha_2 < Y^* \leq \alpha_3 \\
Y = 4 &\equiv Y^* > \alpha_3. \qquad (2)
\end{aligned}
$$

The cut points, given by α_j, partition the latent variable in terms of the J categories of Y. Hence, for a J-category response variable, $J-1$ cut points fully partition Y^*. A statistical model relating $\mathbf{x}$ to Y can be constructed from (1) if a distribution function is specified for ϵ_i. If the standard logistic function is applied, the proportional odds (ordinal logit) model is obtained and is given by

$$
\Pr(Y \leq y_j \mid \mathbf{x}) = \frac{\exp(\alpha_j - \mathbf{x}'\beta)}{1 + \exp(\alpha_j - \mathbf{x}'\beta)}, \qquad (3)
$$

where α_j correspond to $J-1$ intercepts. These parameters relate back to model (1) through the relationship given in (2); as such, these intercepts serve as the estimated cut points discussed above. The parameters in (3) are usually estimated as a linear model for the log-odds ratio using the logit link:

$$
\log\left[\frac{\Pr(Y \leq y_j \mid \mathbf{x})}{\Pr(Y > y_j \mid \mathbf{x})}\right] = \alpha_j - \mathbf{x}'\beta, \qquad j = 1, 2, \ldots j-1. \qquad (4)
$$

Ordinality in Y is preserved, subject to the constraint that $\alpha_1 \leq \alpha_2 \leq \ldots \alpha_{j-1}$ (Ananth and Kleinbaum 1997). The derivation of this model is usually credited to Walker and Duncan (1967), Williams and Grizzle (1972), Simon (1974) and especially McCullagh (1980), who referred to this model as the "proportional odds" model.[3] The proportional odds model in (4) has some attractive features. Because the regression parameters β are invariant to the cut points (note they are not indexed by j), the odds ratios are the same over the $j-1$ cumulative probabilities (Liu and Agresti 2005). Thus, this model has the important property that

$$
\frac{\Pr(Y \leq y_j \mid \mathbf{x} = x_1) / \Pr(Y > y_j \mid \mathbf{x} = x_1)}{\Pr(Y \leq y_j \mid \mathbf{x} = x_2) / \Pr(Y > y_j \mid \mathbf{x} = x_2)} =
$$

[3] Most of our discussion in the remainder of this paper will concentrate on the logit link, though as Liu and Agresti (2005) note, any cumulative link function can be used for ordinal response variables (for example the standard normal link or the complementary log-log link).

$$\frac{\exp(x_1\beta)}{\exp(x_2\beta)} = \exp\{(x_1 - x_2)'\beta\}, \tag{5}$$

where x_1 and x_2 are two different values taken by $\mathbf{x}$. The result in (5) shows that odds ratios derived from (4) are proportional to the distance between values of $\mathbf{x}$ (Powers and Xie 2000). Because of this proportionality, the model gets its name "proportional odds." If this property holds, then the proportional odds model gives rise to stochastic ordering in the cumulative probabilities (McCullagh 1980). That is, if $\exp\{(x_1 - x_2)'\beta\} > 0$, then $\Pr(Y \leq j \mid \mathbf{x} = x_1 < \Pr(Y \leq j \mid \mathbf{x} = x_2)$ for $j = 1, 2, \ldots, j-1$, implying that the response distribution is stochastically higher in covariate class x_2 than x_1 (Jones and Sobel 2000). Proportional odds is therefore a convenient property because knowing β is sufficient to describe how responses on Y "move" with respect to changes in $\mathbf{x}$; the cut points simply move the cumulative probabilities to the left or right, but the shape of the function is determined by β, which is invariant to the cut points. Long (1997) describes this as "parallel regression."

Clearly, the proportional odds model seems naturally suited to ordinal survey items. The model is easy to interpret and generally poses no particular problems to estimate. Nevertheless, at least three issues are worth raising with respect to the proportional odds model. First, it is not always clear that an ordinal response variable, given some set of covariates $\mathbf{x}$ is truly ordinal; second, even if it is ordinal, it is not always obvious that the contrasts given by the proportional odds model are the contrasts the researcher will be interested in; and third (which is closely related to the previous point), it is very likely that in any particular application, the proportional odds assumption *does not hold*. It should be clear from (5) that proportionality in the odds ratios is a property of the model, and not (necessarily) a property of the "real world." Minimally, the proportional odds assumption should, as a matter of course, be tested for in any applications using the proportional odds model. Unfortunately, most social science work, especially political science analyses, using ordinal categorical response variables usually *never* evaluate this assumption (Jones and Sobel 2000).

Ironically, the kinds of problems discussed above are often *never* encountered in many analyses. Indeed, as Liu and Agresti (2005) note, researchers still commonly proceed by applying normal-theory (Gaussian) models to ordered categorical response variables, for example, ordinary least squares. Under this approach, one *explicitly assumes* that Y is equal-interval scored. The problem with this is we almost never know the "true" interval between the observed scale scores. Indeed, this is one motivating factor giving rise to the random utility perspective discussed above. On the face of it, this seems an obviously inappropriate strategy. Nevertheless, it is a commonly employed strategy by political scientists and sociologists. In a "non-scientific" survey of research using ordinal response variables between 1995–2005, we found 57 applications.[4] Of these 57 papers, 23 (40 percent) used simple OLS regression; 28 (49 percent) used cumulative link models (like the proportional odds model), and 6 (11 percent) used alternatives to the standard cumulative link models. The use of OLS is usually justified on the grounds that OLS results are "similar to" results from cumulative link models and therefore OLS results are given because of their "ease of interpretation" (statements like these can be found in a variety of papers, for example Hill 2002, Loftus 2001, Zuckerman and Jost 2001). Still, other researchers provide no justification for assuming equal-interval scoring (c.f. Oliver and Mendelberg 2000). In the next section, we will discuss some of the problems in applying OLS to categorical response variables as well as turn attention to the three issues raised above regarding the proportional odds model.

2.1 The Proportional Odds Assumption

By proceeding with the model given by (4), the researcher is making the relatively strong assumption that the covariate effects are invariant to the cut points, thus implying proportionality in the odds ratios. Of course, if one forgoes the proportional odds model (or something like it) and chooses to apply Gaussian methods, one is making an even stronger

[4]We thank Jennifer Byrne for her research assistance in doing this analysis. The search engine J-Stor was used in this survey.

statement: the scale-score intervals are known and so the regression coefficient informs the researcher explicitly on how $E(Y)$ is changing with respect to changes in $\mathbf{x}$. In either case, one implicitly assumes that once estimated, β describes "movement" over the scale. In the case of OLS, the interpretation is tied directly to the expected value of Y; in the case of the proportional odds model, the interpretation is connected to the conditional or cumulative probabilities (or log-odds or odds ratios). To illustrate some issues, we estimate a model of white attitudes on affirmative action using data from the 1991 National Race and Politics Study (Sniderman, Tetlock, and Piazza 1991).

The dependent variable is a four point item asking about opposition to giving African-Americans preferential treatment in university admissions.[5] We include in our model measures of both symbolic racism and of traditional racial prejudice. Following Oliver and Mendleberg (2000), we construct the symbolic racism measure from a three point likert scale item concerning the amount of attention given to minorities by the government and two 11 point items rating the amount of anger felt by the respondent toward special advantages given to blacks in employment and toward minority leaders who are always complaining about discrimination. The measure of traditional racism consists of the difference between the average scores of five responses each to 11 point positive and negative racial stereotype items. Both symbolic racism and traditional prejudice are rescaled from 0 to 1. We also include education, which is a six point scale rescaled from 0 to 1, ideology, which is scored as -1 if liberal, 0 if moderate, and 1 if conservative, and gender (1 = female). We include in the analysis only white, non-Latino respondents.

In Table 1, we give the OLS and proportional odds estimates for the model of affirmative action attitudes. The way we have presented the parameter estimates, a positively signed coefficient implies an increase in x_k is related to higher scores on Y (or in the cumulative probabilities), and thus, greater opposition to affirmative action. In "eyeballing" the two

[5] More precisely, the measure is a combination of two questions. Respondents were asked if they support or oppose preferential treatment and then asked if they strongly support (oppose) or somewhat support (oppose) the policy. Higher scores represent greater opposition to affirmative action.

sets of coefficients, one might conclude (as many do at this point) that the OLS results are "similar to" the proportional odds estimates. Indeed, at conventional levels of statistical significance, what is statistically significant in one model is significant in the other. For example, the confidence interval around, say, the symbolic racism scale is not wide in either of the models. The major presentational difference between the two models are the $J-1$ cut point parameters given by the proportional odds model. In the OLS model, only a single intercept is needed, owing to the luxury afforded the equal-interval scoring assumption. Because the proportional odds model is parameterized as a linear model for the log-odds ratios *and* because of the presence of two additional intercept parameters, researchers sometimes conclude OLS results are "more interpretable" because they connect directly back to Y, and not to cumulative or conditional probabilities. In order to get quantities of more natural interest, additional work needs to be done to the proportional odds model.

Our interest centers on the proportional odds assumption, however. Consider the symbolic racism scale. In Figure 1, we plot the estimated odds ratio associated with this scale, given by $\exp(\hat{\beta} * x_{SR})$. Clearly, the odds of responding in a higher category (less supportive of affirmative action policies) substantially increase as scores on the scale increase. This result is consistent with theoretical expectations: individuals harboring racially intolerant attitudes tend to also be in opposition to affirmative action policies. Figure 1 nicely illustrates a desirable feature of the proportional odds model: the relationship is invariant to the cut points and so it is unnecessary to consider them in calculating odds ratios.[6] Further, it must be true by definition of the model that the odds ratios are proportional to changes in scale scores. The only impact the cut points have is to move the response distribution to the left or right; the relationship shown in Figure 1 is therefore unaffected by the α_j.

But suppose this assumption were wrong? In this setting, one or more of the log-odds ratios would vary with respect to the cut points. Consequently, the associated odds ratios

[6] It is worth noting that because attention often centers on odds ratios, many statistical treatments of the proportional odds model refer to the cut points as nuisance parameters. As Liu and Agresti (2005) write, the cut points "are usually nuisance parameters of little interest" (p. 3).

would differ, depending on the scale location. In this sense, the unsubscripted β from equation (4) would need to be indexed by j. Fortunately, the proportional odds assumption is, in fact, a testable assumption; unfortunately, in applied work, it is almost never done. Two standard tests have been proposed for evaluating the proportional odds assumption. The first is a global Lagrange multiplier test. The proportional odds model can be considered as a series of J-1 binary logits where the $\beta's$ are constrained across the models such that:

$$\beta_1 = \beta_2 = \ldots = \beta_{j-1} = \beta \tag{6}$$

The null hypothesis of the Lagrange multiplier test is that the log likelihood of the constrained model is no different than that of the unconstrained model. The resulting test statistic is distributed as chi-square with $K(J\text{-}2)$ degrees of freedom, where K is the number of covariates in the model.

While the global test is useful, it does not diagnose violations of the proportional odds assumption for individual covariates. Brant (1990) suggests a Wald test for the overall model and for each variable in the model.[7] The test is conducted by first running J-1 cumulative logits. The global null hypothesis in (6) can be tested with a Wald statistic, which is also distributed chi-square with $K(J\text{-}2)$ degrees of freedom. If the proportional odds assumption holds for covariate K, then the coefficients for that covariate should be equal across the cumulative logits. Thus, a Wald statistic can be constructed for each covariate to test the null hypothesis:

$$\beta_{K1} = \beta_{K2} = \ldots = \beta_{K,j-1} = \beta \tag{7}$$

The test statistic is distributed chi-square with J-2 degrees of freedom.

In applying the Brant test to the proportional odds model from Table 1, we obtain the results presented in Table 2.[8] Globally, we can reject the null that proportional odds hold for the model. Using the Brant covariate specific test, we see that the symbolic racism scale,

[7] See also Peterson and Harrell (1990).

[8] This test can be implemented in Stata by the `spost` routine created by Long and Freese. See also Long (1997) for an explanation of how to conduct the test for statistical software with a matrix language.

the racial prejudice scale, and the indicator for female respondents indicate nonproportionality in the odds. Substantively, this suggests the relationship between these covariates and affirmative action attitudes might vary over the scale scores, a relationship that is assumed *not* to exist in the standard proportional odds model. In short, if the central property of the proportional odds model does not hold, then alternative modeling strategies should be considered. But before considering some alternatives, consider the implications for applying OLS to these data, as is done in Table 1. The alleged similarity in coefficients between the OLS and proportional odds models is illusory. If the proportional odds assumption does not hold, it makes no difference how "similar" linear regression results are to the logit coefficients because the covariate effect is changing with respect to scale location. In using the proportional odds model and testing for this property, one can detect this; in applying OLS, one cannot. Therefore, normal-theory regression models estimated out of convenience will often be misleading. Further, as Lall, Walters, and Morgan (2002), Peterson and Harrell (1990), and Long and Freese (2003) note, the proportional odds assumption commonly will not hold. Indeed, as the number of K parameters to be estimated increases, the chances of finding nonproportionality in the odds ratios will increase. Thus, if one defines a "well specified" model as one having many parameters (usually not a good definition!), then the likelihood the proportional odds assumption is false increases with "improved" specification. Clearly, alternatives should be considered.

3 Multinomial Models for Ordinal Response Variables

A variety of models have been proposed, apart from the proportional odds model, for ordinal categorical response variables. In this section, we consider several we think have applicability to the kinds of survey items social scientists regularly work with. Some of the models discussed are designed to explicitly deal with nonproportional odds; other models are attractive because of the different kinds of contrasts that stem from them. Most of the models discussed below can be thought of as "multinomial" models insofar as estimation of them usually results in multiple parameter vectors. As such, they seemingly increase in complexity as the number of logit functions increase (usually resulting in $J-1$ regression functions). To continue the discussion from the previous section and to motivate what follows, we consider first, models for nonproportional odds.

3.1 Nonproportional Odds

Nonproportional odds imply the relationship between some covariate $\mathbf{x_k}$ and Y changes over the scale. Notationally, we can rewrite model (4) as

$$\log\left[\frac{\Pr(Y \leq y_j \mid \mathbf{x})}{\Pr(Y > y_j \mid \mathbf{x})}\right] = \alpha_j - \mathbf{x}'\beta_{\mathbf{j}}, \qquad j = 1, 2, \ldots j-1, \tag{8}$$

to reflect nonproportional odds. Here, the β are indexed by J, indicating the relationship is assumed to hold differently over the scale scores. Estimating something like this would require multiple parameters on β to be estimated, thus giving rise to a multinomial type model. Several modeling strategies have been proposed for (8), though, as noted earlier, applied political science work rarely utilizes these models. We consider first a model proposed by Peterson and Harrell (1990), which they called an "unconstrained partial proportional odds model."

Peterson and Harrell's "Unconstrained" Model

To motivate the unconstrained partial proportional odds (hereafter UPP) model, consider a fully generalized cumulative logit model of the form

$$\Pr(Y \leq y_j \mid \mathbf{x}) = \frac{\exp(-\alpha_j - \mathbf{x}'\beta_{\mathbf{j}})}{1 + \exp(-\alpha_j - \mathbf{x}'\beta_{\mathbf{j}})}, \tag{9}$$

with log-odds

$$\log\left[\frac{\Pr(Y \leq y_j \mid \mathbf{x})}{\Pr(Y > y_j \mid \mathbf{x})}\right] = -\alpha_j - \mathbf{x}'\beta_{\mathbf{j}}, \qquad j = 1, 2, \ldots j-1. \tag{10}$$

In (10) the notational convention is adapted from Peterson and Harrell (1990) (see also Ananth and Kleinbaum 1997).[9] This is a fully generalized cumulative logit model because each of the K regression parameters are estimated separately for each scale score. Because of this, $J-1$ equations need to be estimated in order to derive the full set of cumulative probabilities. So for a four-category ordinal variable with α_j and β_j, $j = 1, 2, 3$, the cumulative probabilities would be given by

$$\begin{aligned} \Pr(2+ \ \ vs.\ 1) &= \frac{1}{1+\exp(-\alpha_1 - \mathbf{x}'\beta_{\mathbf{1}})} \\ \Pr(3+ \ \ vs.\ 2,1) &= \frac{1}{1+\exp(-\alpha_2 - \mathbf{x}'\beta_{\mathbf{2}})} \\ \Pr(4+ \ \ vs.\ 3,2,1) &= \frac{1}{1+\exp(-\alpha_3 - \mathbf{x}'\beta_{\mathbf{3}})}. \end{aligned} \tag{11}$$

This statement, it should be clear, is greatly simplified in the proportional odds model of (4) because a single parameter vector is estimated. This gives rise to the proportional odds model's property of stochastic ordering. Here, the β_k are estimated uniquely for each contrast over the scale. An application of model (10) in political science can be found in Branton and Jones (2005). The desirable feature of (10) is the proportional odds assumption is relaxed; the undesirable feature of the model is that many parameters must be estimated, some of which may not be of substantive interest. Thus, if Y has four categories and there are three covariates, there are $K + 1(J-1) = 12$ parameters to estimate. Because the number of estimated parameters proliferate with J and $\mathbf{x}$, Peterson and Harrell (1990) proposed the UPP model. Under their formulation, two sets of coefficients are estimated: one set having proportional odds, the second set having non-proportional odds. This model has the form:

$$\log\left[\frac{\Pr(Y \leq y_j \mid \mathbf{x})}{\Pr(Y > y_j \mid \mathbf{x})}\right] = -\alpha_j - \mathbf{x}'\beta - \mathbf{t}'\gamma_{\mathbf{j}}, \qquad j = 1, 2, \ldots j-1. \tag{12}$$

[9] This notation is not necessary and it is important to note that different software code will parameterize this (as well as (4) differently). For example, using `R` and Yee's (2003) `VGAM` package, the log-odds are given by (9). Using the Stata `ado` module, `glogit2` (Williams 2006), the odds ratios are given by $\alpha_j + \mathbf{x}'\beta$ (that is, the signs are different). The quantities of interest from either parameterization will be identical, of course.

Under (12), $\mathbf{x}$ is a $p \times 1$ vector of covariates that maintain the proportional odds property and $\mathbf{t}$ is a $q \times 1$ vector of covariates that are nonproportional. If each covariate exhibits nonproportionality, then model (12) can be shown to be equivalent to the fully generalized model (10); if $q < p$, then only a subset of parameters are nonproportional, thus giving rise to the name "partial" proportional odds. The $\gamma_{\mathbf{j}}$ are the parameters associated with the *incremental* change in the log-odds associated with the nonproportional covariate (Peterson and Harrell 1990; Ananth and Kleinbaum 1997). If the $\gamma_j = 0$, then proportionality holds and the odds ratios are derived as in model (3).

To illustrate the UPP model, we return to the affirmative action data. Recall from Table 2, we found evidence that several of the covariates from the proportional odds model displayed nonproportionality. We refit the model from Table 1 using model (12). The results are given in Table 3. Of the five covariates, three are estimated as having nonproportional odds: symbolic racism, racial prejudice, and the indicator variable for female. Consider the symbolic racism scale. The estimate for the log-odds of a respondent answering in categories 2+ versus 1 is 1.79 with an odds ratio of about 6. However, since this covariate exhibits nonproportionality, the log-odds ratios vary over the scale scores. For the log-odds of answering in categories 3+ versus 2 or 1, the estimate is given by $\beta + \gamma_2 = 1.79 + .80 \approx 2.59$. The odds ratio is thus exp(2.59) or approximately 13.33. Finally, the log-odds for answering in category 4 (being least supportive of affirmative action) are $\beta + \gamma_3 \approx 3.13$ with an odds ratio of about 23. Nonproportional odds informs us that the odds of scoring in a higher category as versus a lower category on the response variable are increasing nonproportionally over the scale. This is in contrast to Table 1, which suggested the odds ratio for the symbolic racism scale was about 12 and was invariant to the cut points. Here, we find that for increases on the symbolic racism scale (i.e. respondents who are more racially intolerant), the odds of answering in the highest categories (3 or 4) are much higher than answering in the lower categories. In this sense, the strongest effects are observed above the implied midpoint of the scale. For our purposes, note the effect of the γ_j. These parameters serve to indicated by how much the odds are increasing. Thus, as Peterson and Harrell (1990) show and as noted above, the $\gamma_{\mathbf{j}}$ give us an estimate of the incremental change in the odds ratios—precisely what we would like to estimate for nonproportional odds. Here, the odds are clearly increasing in Y. This kind of inference is impossible to obtain from the garden variety proportional odds model (as well as the OLS model).

Generalized Ordinal Logit Models

Fu (1998) and more recently Williams (2005, 2006) have discussed the "generalized ordinal logit" model. In the case of Fu (1998), his model (which can be implemented as an `ado` in Stata as `gologit`) is equivalent to the generalized model shown in (10). Williams (2005, 2006) proposes a restricted generalized ordinal logit (which can be implemented in Stata as a `ado` module called `gologit2`). Williams' (2005, 2006) treatment of nonproportional odds more closely follows the approach derived by Peterson and Harrell (1990). Under the generalized ordinal logit, the model has the following form:

$$\log\left[\frac{\Pr(Y \leq y_j \mid \mathbf{x})}{\Pr(Y > y_j \mid \mathbf{x})}\right] = \alpha_j + \mathbf{x}'\beta + \mathbf{z}'\zeta_{\mathbf{j}}, \qquad j = 1, 2, \ldots j - 1. \tag{13}$$

We have written the Williams model in terms of two sets of parameter estimates, β and $\zeta_{\mathbf{j}}$. The β are the coefficient estimates for the covariates maintaining the proportional odds assumption and the ζ_j are the estimates for the $\mathbf{z}$ covariates having nonproportional odds. The only difference between the model in (13) and the UPP model is that in the generalized ordinal logit parameterization, the actual log-odds coefficients are directly reported. In the UPP model, only the increments (given by γ_j) are reported. It should be clear that in Williams' parameterization, $\zeta_2 - \zeta_1 = \gamma_2$. To illustrate the Williams model, consider Table 3a. In this parameterization, three logit equations are jointly estimated. The first column (corresponding to "C_1") gives the log-odds of responding above versus below $Y = 1$; the second column gives the log-odds of responding in categories 3 or 4 versus 1 or 2; finally, the third column gives the log-odds of responding in category 4 versus all others. In Table 3a, the covariates showing nonproportionality have estimates corresponding to $\zeta_\mathbf{j}$. As such, over the $J-1$ logits, the parameter estimates will vary. Again, looking at the symbolic racism scale, the log-odds estimates over the logits are: 1.79, 2.59, and 3.13. The correspondence between Williams' generalized ordinal logit model and the UPP model is straightforward. To obtain the odds increments, subtract the log-odds ratios from each adjacent logit. Thus the log-odds estimate of 2.59-1.79 gives .80, which is γ_2 from Table 3. Similarly, to obtain ζ_2 from the UPP model, simply add $\gamma_2 + \beta$. The two models are identical (and hence, fit statistics like the likelihood ratio, will be the same) as both estimate 14 unique parameters: 3 each for the nonproportional odds; 1 each for the proportional odds; and 3 intercepts. In Table 3a, note that the education and ideology covariate estimates are constrained to be equal over the $J-1$ logits. This is equivalent to saying the odds are proportional. Note also how ordinality in the scale is preserved. In the UPP model, the $\gamma_\mathbf{j}$ provide information on incremental shifts in the odds over the scale; in the generalized ordinal logit, the separate ζ_j give this information. Because the models may be estimated with standard likelihood methods (using a logit link for a cumulative response variable), usual goodness-of-fit statistics can be applied. In comparing the fit of the partial proportional odds model to the proportional odds model (from Table 1), standard fit indices suggest the partial proportional odds model is preferable. Note that since the proportional odds model is obtained when $\gamma_j = 0$, this model is nested under the UPP (or generalized ordinal). The likelihood ratio test is 30.55. On six degrees of freedom, the p-value is .00003. Moreover, the AIC for the UPP model is 4576 and for the proportional odds model, 4595, indicating the UPP model is preferable on these grounds.

The two (complementary) approaches outlined above would seem preferable to the garden variety proportional odds model when the parallel regressions assumption does not hold. The problem, however, is that many users simply fail to evaluate this assumption and proceed as in Table 1. Worse, if users proceed with normal-theory methods, not only is the model putatively inappropriate (on many fronts), possibly useful information on nonproportional odds is lost. Of course, all of the preceding worked on the premise that: 1) ordinality was present in the data and/or 2) the contrasts afforded the proportional odds model are of primary interest. Nevertheless, it is conceivable that *non*-ordinal models might actually be *preferred* in some instances to ordinal models. Further, alternative contrasts might reveal information not easily obtained from the proportional odds model (or the UPP or generalized ordinal logit model). We turn to these issues/models next.

3.2 Alternative Modeling Strategies

In the previous section, we considered a class of multinomial models that preserved the information on ordinality in Y while relaxing the proportional odds assumption. Nevertheless, there are alternative strategies apart from those just discussed that can be applied to putatively ordinal response variables from surveys. To fix some ideas, we first consider the baseline category logit model.

Baseline Category Logit

Possibly the most general model is a multinomial model where ordinality is not assumed. The baseline category logit model, or "multinomial logit," model proceeds by treating a polytomous response variable has having an arbitrarily defined baseline to which the log-odds ratios are referenced. This model is usually applied for truly nominal and non-ordered data. In the context of an ordinal response variable, the log-odds ratios are

$$\log\left[\frac{\Pr(Y = y_j \mid \mathbf{x})}{\Pr(Y = 1 \mid \mathbf{x})}\right] = \alpha_j + \mathbf{x}'\beta_{\mathbf{j}}, \quad j = 1, 2, \ldots j - 1, \tag{14}$$

where the denominator in the log-odds ratio corresponds to the baseline category. The choice of the baseline category is irrelevant. Because the parameters are subscripted by J, the resulting model yields $J-1$ nonredundant logits. If the response variable is four-category, the contrasts from the model are category $J = 2$ versus 1, $J = 3$ versus 1, and $J = 4$ versus 1. The model in (14) is actually quite similar to the fully generalized ordinal model shown in (9); however, ordinality in the response variable is not assumed (or is not preserved) in this particular setup. In principle, there is nothing particularly wrong in applying this model to ordinal data. Indeed, even if a response variable is ostensibly ordered, ordinality may not actually be found conditional on the covariates, (i.e. the response distribution may not monotonically increase or decrease in $\mathbf{x}$.) In such a case, applying the baseline category logit model may be a reasonable strategy. Nevertheless, on the face of it, with an ordinal response variable, there is usually some compelling reason to derive ordinal (or "ranked") interpretations from the model. To see, however, the connection between baseline category logit model and ordinal data, we consider a reparameterized baseline category logit.

Adjacent Category Logit

The adjacent category logit model proceeds by forming $J-1$ logits for all pairs of adjacent categories (Agresti 1996). The model has not been widely applied in political analysis (though see Cameron, Epstein and O'Halloran 1996 or Jones and Sobel 2000) though it would seem to have some appeal given the prevalence of multicategorical response variables in political science research. This model has been much more widely applied in sociological applications. This is probably due to its close connection to log-linear models (Agresti 1996, Sobel, Becker, and Minick 1998, Jones and Sobel 2000, Powers and Xie 2000). The model gets its name because of how the log-odds ratios are parameterized. Note that in the baseline category model of (14), the log-odds are in reference to a baseline category (which is arbitrarily defined). In the adjacent category logit model, the log-odds of score $J+1$ are in reference to score J–that is, the "adjacent" category. The model can be written as

$$\log\left[\frac{\Pr(Y = y_{j+1} \mid \mathbf{x})}{\Pr(Y = y_j \mid \mathbf{x})}\right] = \alpha_j + \mathbf{x}'\beta_{\mathbf{j}}, \qquad j = 1, 2, \ldots j-1. \tag{15}$$

The resulting odds ratios from (15), then, give the odds of responding in a higher category versus the lower adjacent category. The model itself is easily estimated because the parameters and standard errors can be directly obtained from the baseline category logit model. To illustrate, we apply the adjacent category logit model to data on immigration attitudes.

Specifically, we use the 2004 National Election Studies (NES) to examine white, non-Latino evaluations of the effect of immigration on job availability. Our dependent variable is constructed from a question about whether it is extremely likely, very likely, somewhat likely, or not likely at all that current immigration levels will take jobs away from people already in the United States.[10] We include five independent variables in the model. First, we construct a measure of evaluations of Hispanic stereotypes from three seven-point questions, with high values meaning negative stereotypes. A score measuring the relative group evaluations is created from subtracting a Hispanic feeling thermometer from a white thermometer. Four items are scaled to create a measure of moral traditionalism where higher values equate to higher levels of traditionalism. Personal retrospective economic evaluations are also included in the model.[11] Finally, ideology is measured with a seven point self-placement question (strong liberals = 0). All five measures are recoded from 0 to 1. We hypothesize that larger values of the independent variables should increase the likelihood that a respondent believes that immigration will reduce the number of available jobs.[12] The results for an adjacent category logit model are given in Table 4.

The interpretation of the coefficients are akin to any logit model. A positively signed coefficient implies the log-odds are increasing in $\mathbf{x}$. What sets this model apart from others considered to this point are the contrasts it gives. The first column of estimates in Table 4, for example, gives the log-odds of responding in category 2 ("very likely") versus category 1 ("extremely likely"); column 2 gives the log-odds of responding in category 3 ("somewhat likely") versus category 2; and column 3 gives the log-odds for responding in category 4 ("not likely at all") versus category 3. To illustrate, consider the moral traditionalism scale. In the first column, the odds of 2 vs. 1 are $\exp(-.81) = .44$; the odds of 3 vs. 2 are $\exp(-1.52) = .22$; and the odds of 4 vs. 3 are $\exp(-1.75) = .17$. Interestingly, the odds

[10]We code extremely likely = 1, not very likely = 4

[11]This is a five-point question asking how much better or worse the respondent's family is economically than one year ago.

[12]See Branton, Jones, and Westerland (2006) for a more complete discussion.

ratios are monotonically decreasing over the range of the scale. Further, the confidence interval around the contrast 2 vs. 1 is much wider than the other contrasts suggesting the substantive impact of moral traditionalism is strongest for the higher categories of the scale but not the lower categories.

The appealing feature of the adjacent category logit model is that the contrasts forthcoming from this model would seem to be of natural interest in many political science applications (Jones and Sobel 2000). The model is general in the sense that the odds ratios can (obviously) vary over the adjacent categories. Consequently, no assumption of proportional odds needs to be made. Ordinality in the response variable is preserved because the odds are estimated for each pair of adjacent categories. If Y is a Likert-type survey item, then the local logits inform the analyst on how $\mathbf{x}$ is related to "climbing the scale." The major drawback of the model is it yields $(K+1)(J-1)$ log-odds ratios. The number of parameters could be reduced by placing restrictions over some of the adjacent category logits, but the model is multinomial nonetheless. It is useful to note the equivalence between the adjacent category logit model and the baseline category logit. Suppose we apply model (14) to these data, treating category 1 as the baseline. Were we to do so, the $J-1$ log-odds for the moral traditionalism scale would be: (2 vs. 1): $-.81$; (3 vs. 1): -2.33; and (4 vs. 1): -4.08. To obtain the adjacent category logits, one simply subtracts these ratios sequentially. Thus, the log-odds of category 3 vs. 2 is given by $-2.33 + .81 = -1.52$ and the log-odds of category 4 vs. 3 is $-4.08 + 2.33 = -1.75$. These are identical to the ratios obtained from Table 4. Given the equivalence of the two models, it makes no sense to think about adjudicating between them: the likelihood ratio is the same (and is -981.839). The major difference here is one of interpretation. The adjacent category logit gives rise to an "ordinal" interpretation (in terms of adjacent scores); the baseline category logit model does not.

Anderson's Stereotype Model

As an alternative to the baseline category logit (as well as, perhaps, the adjacent category logit), Anderson (1984) derived a model he named the "stereotype" ordinal regression model. The model's awkward sounding name comes from the perspective Anderson took towards ordinal response variables. Specifically, he noted that researchers often work with "assessed" ordinal scales (Anderson 1984, 2). Such a scale requires "judges" to assign ordinal ratings on an otherwise continuous scale. Because judgements require categorization, these scores may be "loose stereotypes" (Anderson 1984, 10), and be prone to error. Put differently, imagine a medical researcher assessing "pain" a patient is experiencing. The judge (the researcher) may rely on multiple pieces of information to assess the pain level. Once a judgement is made, a category is assigned. The category assigned to a given patient may be "stereotypical" for that kind of patient. Similarly, in survey research, respondents are called upon to answer questions. As noted throughout, it is common that survey researchers categorize otherwise ordinal concepts into a series of observable scale categories. In the language of Anderson, these categorizations might be stereotypical labels. The import of the model is in its recognition that observed scale scores from an ordinal item might be error prone because the categorizations are stereotypes. Anderson (1984) reasoned that because of this, a putatively ordinal scale may in fact *not* be ordered at all. Further, Anderson (1984) developed the stereotype logit model as an alternative to the baseline category logit. Because the number of terms estimated increase in both K and J, Anderson proposed a model that estimated a single set of regression parameters (i.e. β) and a series of weights (ϕ_j) that could be used to evaluate the ordinality assumption. In terms of the log-odds, the stereotype model is given by

$$\log\left[\frac{\Pr(Y = y_j \mid \mathbf{x})}{\Pr(Y = 1 \mid \mathbf{x})}\right] = \alpha_j + \mathbf{x}'\phi_j\beta, \qquad j = 1, 2, \ldots j - 1, \tag{16}$$

where the ϕ correspond to the estimated weights for each scale category. As with the baseline category logit, one of the J categories must serve as a referent. To identify the model, $\phi_J = 0$

and $\beta_J = 0$. To normalize the scale, a further restriction is placed on the model such that $\phi_1 = 1$. Thus for a four-category response variable, two ϕ parameters are estimated. It should be clear that inference on the β_k is not independent of the ϕ; the log-odds are given by $\phi_j \beta_k$.

The stereotype logit model has not been widely applied. As Liu and Agresti (2005) note, this is probably because Anderson died just before his 1984 paper was published. Nevertheless, the model has some attractive features. First, far fewer parameters need to be estimated than when compared to the multinomial models discussed earlier. Further, the model can be thought of as providing a test of ordinality. Specifically, since the ϕ_j give information about the scale categories, if the condition given by

$$1 = \phi_1 > \phi_2 > \ldots \phi_J = 0$$

hold, then an ordered regression model is obtained (Ananth and Kleinbaum 1997, Lall et al 2002); if this condition does not hold, then there is some evidence of a lack of ordinality (given the covariates in the model). Thus, a nice feature of the model is that ordinality is not a property of the model and so (16) provides an explicit test of the property. Yee and Wild (1996) and Yee and Hastie (2003) refer to the stereotype model as a "reduced-rank" model. This is because only a single parameter vector is estimated for the β (again, in contrast to the baseline category logit). For similar reasons, McCullagh (1984) calls this model a canonical regression model.

We illustrate and interpret the stereotype model using the data on immigration attitudes discussed in the previous section. We applied the stereotype model using the same set of covariates as in Table 4. The parameter estimates and standard errors are given in Table 5. For purposes of identification, category $J = 4$ ("not very likely") is treated as the baseline category. The ϕ_j parameters correspond to the weights just discussed. The α_j correspond to the intercepts for the $J - 1$ categories. To interpret the model, first consider the α_j. The estimates suggest the log-odds ratios are ordinal, as they monotonically increase from $J = 4$

(the baseline category) to $J = 1$. The estimated standard errors relative to the parameter estimates are small and so there is evidence suggesting the categories are distinguishable from one another. Moreover, the ϕ_j give information about the "distance" between categories. For example, the difference between the baseline category ($J = 4$) and category $J = 3$ ("somewhat likely") is largest (.51). In contrast, the distance between scale score $J = 1$ and $J = 2$ is the smallest (.15). Given the nature of the scale, this is not surprising: category four is distinct from the remaining three categories as it is the only response option allowing respondents to claim immigrants will *not* likely take away jobs. Since the other categories record negative views toward immigrants and the chances they will take away jobs, the difference between the α_j are smallest for them.

Turning now to the odds ratios, consider again the moral traditionalism scale. The coefficient estimate of 4.11 gives the log-odds ratio for responding in category 1 ("extremely likely to take away jobs") versus category 4 ("not very likely"). The associated odds ratios is $\exp(4.11) \approx 61$. This implies the odds for a respondent scoring highest on the moral traditionalism scale saying immigrants are extremely likely to take away jobs is about 60 times that of someone scoring lowest on the moral traditionalism scale. The confidence interval around this estimate is very tight, as the standard error is very small. The remaining odds ratios can be derived from Table 5 by accounting for the scale score weights. Thus, the odds of responding "very likely" ($J = 2$) versus "not very likely" ($J = 4$) is $\exp(\beta\phi_2) = \exp(4.11 \times .85) \approx 33$, or roughly half of the previous contrast. Finally, the odds of answering "somewhat likely" versus "not very likely" (i.e 3 vs. 4) are about 8.02, or roughly a quarter of the previous contrast.

It is clear that the interpretation of the stereotype model is akin to the baseline category logit model with two very important differences. First, ordinality in Y is explicitly tested for here in the ϕ_j; second, there is only a single parameter vector estimated for K (making it a reduced rank model). In comparing the stereotype logit model to a baseline category logit (the full model is not reported here), we find that the odds ratios for the same set of contrasts computed in the previous paragraph are approximately, 59, 26, and 6. The estimates are similar to, though not identical, to the estimates gotten from the stereotype model. Since the stereotype model is a reduced rank version of the baseline category, standard likelihood ratio tests can be applied. We find that the log-likelihood for the 15 parameter baseline category logit is -981.839 and for the 10 parameter stereotype logit is -983.586. On five degrees of freedom, the difference in likelihoods between the baseline category and the stereotype logit is not significant. We would probably, therefore, prefer the stereotype model on the grounds of parsimony (use of the AIC and BIC criteria would reenforce this decision).

4 Discussion and Conclusion

Numerous alternatives to the standard approaches for ordinal response variables exist. Rarely in political science applications are these alternatives considered. A variety of reasons, both statistical and theoretical, were forwarded here to lend motivation for thinking about these alternative strategies. We should note that the survey of models provided here is by no means exhaustive. For example, we have made no mention of the continuation ratio logit model (McCullagh 1980; Cole and Ananth 2001), the multidimensional Anderson stereotype model (Anderson 1984, Ananth and Kleinbaum 1997, Lall et al 2002), and the constrained partial proportional odds model (Peterson and Harrell 1990). The continuation ratio logit model, in our view, has less applicability to the issues considered here: ordinal survey items. This is a compelling model when the ordering in the data is sequential, such that once category J is "passed through," it cannot be returned to (for example, progression of an illness). There is little reason to believe that data on survey responses to ordinal variables satisfies this sequentiality condition. The multidimensional stereotype model is most applicable when the response variable can be thought of as an amalgam of multiple indicators (Anderson 1984, Ananth and Kleinbaum 1997). The "dimensionality" refers to the number of possible dimensions underlying the observed scale. Because of space constraints, we opted to discuss the simpler one-dimensional model (given its connection to the other multinomial models considered herein) and forgo discussion of this model for another paper. Finally, we did not discuss the constrained partial proportional odds model forwarded by Peterson and Harrell. This model is a straightforward extension of the model discussed in (12) subject to the constraint that prespecified scalars are included in the model to describe the nature of nonproportional odds. Because these constraints are rarely known (and the ability to "know" them decreases as the number of parameters increase in the model), our view is the UPP model will be more naturally applicable to the kinds of problems political scientists work with.

Nevertheless, the models covered herein address a variety of issues that routinely emerge with ordinal response variables. The natural question to ask is 'which model should I choose?'. As should be clear, there is no obvious answer to this question because a variety of issues must be sorted through. Among the most important issues here is thinking about how the data were generated. If use of an ordinal response variable can be motivated in terms of random utility, then one is implicitly saying that underlying the scale is a continuous and unidimensional latent factor. Further, given this, it is implicitly assumed that the observed scale scores are ordered. To the extent these assumptions hold—that there exists some theoretical reason to expect individuals to "think" in terms of the latent factor—then the proportional odds model seems the natural starting point for the analysis. Unfortunately, as noted previously, it seems many researchers' starting point obviates the random utility perspective and instead opts for assuming equal-interval scoring. This assumption, which most likely impossibly holds for scales with few categories, leads to the use of normal-theory models, like OLS. In our "casual" overview of applied work, this strategy was quite prevalent. We hope we have made it clear that starting from this perspective is generally not a good idea. It is also important to note that starting and *ending* with the proportional odds model is likewise, a potentially bad strategy. The parallel regressions assumption built into the proportional odds model is a rigid restriction. Indeed, if the property holds, then the model has an attractive interpretation. Unfortunately, for many of the kinds of models political scientists like to postulate—models with many parameters—the chances proportionality holds is unlikely. Minimally, any use of the proportional odds model should be coupled with tests of the proportional odds assumption. Should the assumption not hold, we have considered a variety of alternatives to the proportional odds model, ranging from fully generalized ordinal logit (Fu 1998) to partial proportional odds models (Peterson and Harrell 1990, Williams 2005, 2006). These latter models seem particularly compelling, particularly if only a few covariates exhibit nonproportionality. Because these kinds of models can be easily estimated in software like R, Stata, and SAS, there is no particular reason we can think of to forgo considering them.

In contrast to the random utility perspective, if one is unclear about the dimensionality of the response variable or if one has reason to suspect ordinality may *not* hold given the covariates in the model, alternative strategies may be in order. We discussed a variety of multinomial models that could be derived from the baseline category logit. Anderson's stereotype logit seemed particularly compelling, as it provided a direct test of ordinality in the response variable. Further, if the model holds, far fewer parameters need to be estimated (and interpreted) in his model than when compared to the fully generalized multinomial models (like the baseline category logit).

Another consideration to be made involves the issue of inference and interpretation. As we hope we have made clear, "what you get" from these models varies substantially over different parameterizations. This is most clearly evidenced in the adjacent category logit. If the theoretical interest centers on how respondents move from one scale location to the next, for example, moving from below a scale's midpoint to above the midpoint, this kind of modeling strategy seems attractive because the contrasts are explicitly modeled in terms of the local logits. However, if interest centers on how respondents score on the scale in reference to some baseline scale category, then alternative parameterizations may be of interest. The baseline category logit (and the stereotype) model serves as an example of this.

The overriding issue of this paper, however is this. Forgoing alternative modeling strategies in favor of either the OLS or proportional odds model may result in either 1) ill-fitting models or 2) models that ignore possibly useful information. With respect to this second issue, suppose that proportionality in the odds ratios does not hold. If it does not, this implies that a covariate's effect will likely vary over the scale's location. A partial proportional odds model (or other models we have discussed herein) would be able to capture this effect. Now imagine a theory of political behavior that postulated some factor should have a stronger impact, say, above the midpoint on Likert-type item and weaker effect below. For example, consider Feldman and Huddy's (2005) important finding that the relationship between racial resentment and attitudes substantially differs for liberal versus conservative self-identifiers. If the racial resentment scale were treated as a covariate in a model of racial attitudes (measured with an ordinal response variable), one might theorize that the relationship between resentment interacted with ideology of the respondent would exhibit sharply different odds ratios over the scale. Applying a garden variety proportional odds model—or worse, OLS—would simply not pick this relationship up. The kinds of models discussed herein, could. Indeed, in the applications presented here, we found substantial differences in the relationship between symbolic racism and affirmative action attitudes over the range of the scale. Similar remarks apply to the relationship between moral traditionalism and attitudes toward immigrants. In short, on theoretical grounds, one might be naturally led to the multinomial models discussed here.

References

Agresti Alan. 1996. *An Introduction to Categorical Data Analysis.* New York: Wiley.

Agresti Alan. 2002. *Categorical Data Analysis.* 2nd Edition. New York: Wiley.

Ananth, Cande V. and David G. Kleinbaum. 1997. "Regression Models for Ordinal Responses: A Review of Methods and Applications." *International Journal of Epidemiology* 26: 1323-1333.

Anderson, J.A. 1984. "Regression and Ordered Categorical Variables (with Discussion)." *Journal of the Royal Statistical Society, Series B* 46: 1-30.

Brant, Rollin. 1990. "Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression." *Biometrics* 46: 1171-1178.

Branton, Regina P., Bradford S. Jones and Chad Westerland. 2006. "All Along the Watchtower: Acculturation, Fear, and Immigration." Manuscript.

Branton, Regina and Bradford S. Jones. 2005. "Reexamining Racial Attitudes: The Conditional Relationship Between Diversity and Socioeconomic Environment." *American Journal of Political Science* 49: 359-372.

Cameron, Charles, David Epstein, and Sharyn O'Halloran. 1996. "Do Majority-Minority Districts Maximize Substantive Black Representation in Congress?" *American Political Science Review* 90: 794-812.

Cole, Stephen R. and Cande V. Ananth. 2001. "Regression Models for Unconstrained, Partially or Fully Constrained Continuation Odds Ratios." *International Journal of Epidemiology* 30: 1379-1382.

Feldman, Stanley and Leonie Huddy. 2005. "Racial Resentment and White Opposition to Race-Conscious Programs: Principles or Prejudice?" *American Journal of Political Science* 49: 168-183.

Fu, Vincent. 1998. "Estimating Generalized Ordered Logit Models." *Stata Technical Bulletin* 8: 160-164.

Hill, Mark E. 2002. "Skin Color and the Perception of Attractiveness among African Americans: Does Gender Make a Difference?" *Social Psychology Quarterly* 65: 77-91.

Jones, Bradford and Michael Sobel. 2000. "Modelling Direction and Intensity in Semantically Balance Ordinal Scales: An Assessment of Congressional Incumbent Approval." *American Journal of Political Science* 44: 174-185.

Loftus, Jeni. 2001. "America's Liberalization in Attitudes toward Homosexuality, 1973 to 1998." *American Sociological Review* 66: 762-782.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage.

Long, J. Scott, and Jeremy Freese. 2003. *Regression Models for Categorical Dependent Variables Using Stata.*

Liu, Ivy and Alan Agresti. 2005. "The Analysis of Ordered Categorical Data: An Overview and Survey of Recent Developments." *Sociedad de Estadstica e Investigacion Operativa TEST* 14: 1-30.

McCullagh, Peter. 1980. "Regression Models for Ordinal Data (with Discussion)." *Journal of the Royal Statistical Society, Series B*, 42: 109-142.

McCullagh, Peter. 1984. "Response to Anderson" *Journal of the Royal Statistical Society, Series B* 46: 22-23.

McKelvey, Richard D. and William Zavoina. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *Journal of Mathematical Sociology* 4: 103-120.

Oliver, J. Eric and Tali Mendelberg. 2000. "Reconsidering the Environmental Determinants of Racial Attitudes." *American Journal of Political Science* 44: 574-589.

Peterson, Bercedis and Frank E. Harrell, Jr. 1990. "Partial Proportional Odds Models for Ordinal Response Variables." *Applied Statistics* 39: 205-217.

Powers, Daniel A. and Yu Xie. 2000. *Statistical Methods for Categorical Data Analysis.* San Diego, CA: Academic Press.

Simon, Gary. 1974. "Alternative Analyses for the Singly-Ordered Contingency Table." *Journal of the American Statistical Association* 69: 971-976.

Sobel, Michael E., Mark P. Becker, and Susan M Minick. 1998. *American Journal of Sociology* 104: 687-721.

Walker, Strother H. and David B. Duncan. 1967. "Estimation of the Probability of an Event as a Function of Several Independent Variables." *Biometrika* 54: 167-179.

Williams, O. Dale and James E. Grizzle. 1972. "Analysis of Contingency Tables Having Ordered Response Categories." *Journal of the American Statistical Association* 67: 55-63.

Williams, Richard. 2005. "Gologit2: Generalized Ordered Logistic Regression/Partial Proportional Odds Models for Ordinal Dependent Variables." Working Paper, University of Notre Dame.

Williams, Richard. 2006. "Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables." *The Stata Journal* 6: 56-82.

Yee, Thomas W. 2004. "A New Technique for Maximum-Likelihood Canonical Gaussian Ordination." *Ecological Monographs* 74: 685-701.

Yee, Thomas W. and Trevor J. Hastie. 2003. "Reduced-rank Vector Generalized Linear Models." *Statistical Modeling* 3: 15-41.

Yee, Thomas W. and Trevor J. Hastie. Nd. "Reduced-rank Vector Generalized Linear Models." Working Paper, University of Auckland.

Zuckerman, Ezra W. and John T. Jost. 2001. "What Makes You Think You're so Popular? Self-Evaluation Maintenance and the Subjective Side of the "Friendship Paradox." *Social Psychology Quarterly* 64: 207-223.

Figure 1: *This figure gives the estimated odds ratios for the symbolic racism scale from the proportional odds model.*

Table 1: Support for Affirmative Action
OLS and Proportional Odds Estimates

Variable	Coefficient	
	OLS	Proportional Odds
Symbolic Racism	1.375	2.562
	(0.111)	(0.207)
Racial Prejudice	0.383	0.728
	(0.193)	(0.356)
Education	0.259	0.510
	(0.102)	(0.182)
Ideology	0.066	0.117
	(0.027)	(0.047)
Female	0.025	0.075
	(0.049)	(0.088)
Intercept	1.479	
	(0.127)	
α_1		0.652
		(0.233)
α_2		1.888
		(0.236)
α_3		3.268
		(0.245)
n	1744	1744
log-likelihood:		−2289.518

Data are from the 1991 Race and Politics Study. Column one gives O.L.S. coefficients and column two gives proportional odds estimates. Models were estimated in both `Stata` v.9 and `R` and the `VGAM` package (Yee 2003, Yee and Wild 1996). Code is available from authors. Signs on the proportional odds were switched to make it comparable with OLS estimates.

Table 2: Testing the Proportional Odds Assumption

Variable	**Coefficient**		
	$y > 1$	$y > 2$	$y > 3$
Symbolic Racism	1.822	2.458	3.037
Racial Prejudice	1.384	0.219	0.664
Education	0.209	0.496	0.651
Ideology	0.116	0.105	0.139
Female	-0.118	0.144	0.052
Constant	-0.279	-1.651	-3.645

Variable	χ^2	$p > \chi^2$	**df**
All	29.33	0.001	10
Symbolic Racism	12.37	0.002	2
Racial Prejudice	8.15	0.017	2
Education	2.29	0.319	2
Ideology	0.33	0.849	2
Female	5.79	0.055	2

Data are from the 1991 Race and Politics Study. Column. The test for proportional odds was done in `Stata` using the Brant test (Brant 1990)

Table 3: Support for Affirmative Action
Partial Proportional Odds (UPP Model)

Variable	**Coefficient**	**Standard Error**
Symbolic Racism	1.788	(0.285)
γ_2	0.800	(0.257)
γ_3	1.345	(0.358)
Racial Prejudice	1.563	(0.495)
γ_2	-1.437	(0.440)
γ_3	-0.811	(0.586)
Female	-0.098	(0.125)
γ_2	0.277	(0.111)
γ_3	0.147	(0.145)
Education	0.479	(0.182)
Ideology	0.122	(0.047)
α_1	.273	(.22)
α_2	-1.635	(.203)
α_3	-3.263	(.236)
n	1744	
log-likelihood	-2274.293	

Data are from the 1991 Race and Politics Study. The model was estimated in `Stata` v.9 using `gologit2` (Williams, 2005, 2006). Code is available from authors.

Table 3a: Support for Affirmative Action
Partial Proportional Odds (Generalized Ordinal Logit)

Variable	**Coefficient**		
	C_1	C_2	C_3
Symbolic Racism	1.788	2.588	3.133
	(0.285)	(0.245)	(0.280)
Racial Prejudice	1.563	0.127	0.752
	(0.495)	(0.421)	(0.462)
Education	0.478	0.478	0.478
	(0.182)	(0.182)	(0.182)
Ideology	0.122	0.122	0.122
	(0.047)	(0.047)	(0.047)
Female	-0.098	0.180	0.049
	(0.125)	(0.101)	(0.113)
Constant	-0.509	-1.698	-3.639
	(0.283)	(0.261)	(0.299)
n	1744		
log-likelihood	-2274.293		

Data are from the 1991 Race and Politics Study. The model was estimated in `Stata` v.9 using `gologit2` (Williams 2005, 2006). A fully generalized model was also estimated in `R` using `VGAM` package (Yee 2003, Yee and Wild 1996). Code is available from authors.

Table 4:
Adjacent Category Logit

Variable	**Coefficient**		
	2 vs 1	3 vs 2	4 vs 3
Moral Traditionalism	-0.812	-1.517	-1.752
	(0.776)	(0.639)	(0.742)
Group Difference	-0.010	-0.028	-0.001
	(0.096)	(0.076)	(0.088)
Hispanic Traits	-0.417	-1.092	-1.928
	(0.948)	(0.779)	(0.876)
Ideology	0.490	0.031	0.027
	(0.317)	(0.262)	(0.304)
Economic Evaluation	-0.324	-0.179	-0.714
	(0.402)	(0.336)	(0.410)
Constant	1.161	2.235	1.353
	(0.748)	(0.603)	(0.639)
n	774		
log-likelihood	−981.839		

Data are from the 2004 National Elections Study. Models were estimated in `R` using the `VGAM` package (Yee 2003, Yee and Wild 1996). Code is available from authors.

Table 5:
Stereotype Logit Model

Variable	**Coefficient**
Moral Traditionalism	4.111
	(0.852)
Group Difference	0.044
	(0.093)
Hispanic Traits	3.444
	(0.990)
Ideology	-0.410
	(0.328)
Economic Evaluation	1.092
	(0.437)
α_1	-4.725
	(0.825)
α_2	-3.660
	(0.726)
α_3	-1.382
	(0.624)
ϕ_1	1
ϕ_2	0.848
	(0.131)
ϕ_3	0.507
	(0.102)
ϕ_4	0
n	774
log-likelihood	-983.586

Data are from the 2004 National Elections Study. Models were estimated in `R` using the `VGAM` package (Yee 2003, Yee and Wild 1996). Code is available from authors.