The Politics of Opinion Assignment: A Conditional Logit Model with varying Choice Set.*

Thomas Gschwend
Mannheimer Zentrum fuer Europaeische Sozialforschung (MZES)
University of Mannheim
68131 Mannheim, Germany
Thomas.Gschwend@mzes.uni-mannheim.de

Chad M. King
Department of Political Science
Stony Brook University
Stony Brook, NY 11794-4392
chking@ic.sunysb.edu

October 26, 2002

*The authors wish to thank Maltzman, Spriggs and Wahlbeck for providing their replication data. We also wish to thank Stanley Feldman, Bradford Jones, Jeffrey Segal, Florian Heiss, Paul Thurner, Ulrich Kohler, Jack Buckley, Jeff Gill and Scott Graves for their helpful comments and assistance. The authors' names are listed alphabetically and any errors or omissions are solely their responsibility.

ABSTRACT

This note replicates and extends Chapter 2 of Forrest Maltzman, James F. Spriggs and Paul J. Wahlbeck's (henceforth: MSW) "Crafting Law on the Supreme Court" (2000). Using a conditional logit model, the authors test the effects of both choice-specific and chooser-specific variables on majority opinion assignment on the United States Supreme Court during Chief Justice Burger's tenure. The authors find that the effect of ideology, as well as other variables, is conditioned on both case facts as well as justices' attributes. In this note, we take issue with the authors' specification of the model, specifically their failure to include choice-specific, i.e. the justices, constants. Below we argue for the statistical necessity of the inclusion of these controls and reassess the original theoretical model with the appropriate statistical specification. We first show that the failure to include these constants will yield biased estimates. We then test if the authors' substantive findings are robust to the correct specification of their original model. While we successfully replicate the original model (yielding biased estimates), we generally find that MSW's core findings, although confirmed, are diminished when correctly estimated.

1. INTRODUCTION

Conditional logit models have increased in both popularity and employment in political science because of the ease of their implementation in most statistical software as well as their appropriateness to fundamental theoretical questions. These models allow scholars to simultaneously test the effects of not only choice-specific characteristics, but chooser-specific characteristics as well. In this note, we raise a more general methodological issue with specifications of conditional logit models that is not commonly appreciated in the field of judicial politics and in the discipline at large. We show that the way these models are often specified yields biased estimates.

In an important and award-winning[1] work by MSW, the authors use this model to estimate the causes of majority opinion assignment on the Supreme Court. While the authors find that several conditional relationships influence assignment decisions, we question the manner in which they estimated their model. If the authors' substantive findings are robust, they should withstand an extension by correctly specifying their original model. While we successfully replicate the original model (yielding biased estimates), we generally find that the MSW's core findings although confirmed, are diminished when correctly estimated.

2. SUPREME COURT OPINION ASSIGNMENT

Why is the assignment process of majority opinions so important to understanding the Supreme Court? The majority opinion provides more than just the answer as to who "wins" the case; it provides the reasoning for that outcome. Often times the written opinion creates or interprets legal standards that will be important not only to the litigants but to the public as well as lower courts and other branches of government. The majority opinion affects not

[1]The 2000 C. Herman Pritchett Award.

only the case it is written for but how future cases will be decided as well as the future behavior of a variety of actors. Because the actual text and language of opinion matters, choosing which justice writes that opinion matters too (Epstein and Knight, 1998). Justices frequently differ in the legal standards and reasoning that they employ in their decisions. This suggests that given the same case, two justices who agree on the outcome of the case might differ in how to arrive at that common outcome. Thus, who is assigned a given opinion is extremely important. MSW's work provides a comprehensive and detailed examination of opinion assignment on the Burger Court. Their model reveals that ideology affects opinion assignment, but only under conditions of minimum winning coalitions and highly salient cases. They also find that non-ideological variables matter. The assigning Chief Justice was found to be concerned with equity of assignments, the expertise of justices in the issue area of the case, as well as the burden of the existing workload of justices at the time of the assignment. These findings lead the authors to conclude that the assignment process is best summarized as a "Collegial Game".

3. REPLICATION OF MSW'S RESULTS

In Table 1, we replicate the results from MSW's original analysis of Burgers opinion assignments based upon their conditional logit estimation. As just discussed, ideology affects opinion assignment, but only when it is conditioned on a few independent variables; political salience, the closeness to the end of the term and as the size of the original majority coalition increases. Workload and equity matter as well, as the Chief Justice seems to prefer assignment to justices who have fewer decisions to write at the time of assignment and those who receive fewer assignments the same day from associate justices. Lastly, expertise is significant as justices who are experts in the issue areas of the case are more likely to be assigned the opinion. These findings lead MSW to conclude that the Chief Justices assignments

Table 1: *Replication of MSW's Results:* Conditional Logit Model of Opinion Assignment on the Burger Court.

Independent Variables	Coef.	Std.Err.	*p*-value
Ideology	0.017	0.006	0.002
Self-Assignment (= j1)	-0.167	0.086	0.051
Ideology X Winning Margin	-0.004	0.001	0.001
Ideology X Political Salience	-0.003	0.001	0.001
Ideology X Legal Salience	-0.006	0.006	0.357
Self-Assignment X Political Salience	-0.018	0.021	0.392
Self-Assignment X Legal Salience	0.287	0.246	0.243
Equity	-0.095	0.035	0.006
Expertise	0.066	0.026	0.010
Freshman	-0.072	0.096	0.454
Freshman X Case Complexity	-0.048	0.097	0.617
Workload	-0.107	0.027	0.000
Ideology X End of Term	0.000	0.000	0.060
N	12873		
PCP	0.190		
PRE	0.042		

are based upon not just ideological concerns, but also upon institutional and political factors.

4. CORRECTION OF THE MSW MODELING STRATEGY

In this section we extend and improve upon MSW's modelling strategy. In general we agree with the authors that a conditional logit (CL) model is a reasonable strategy to model opinion assignment for the following two reasons. First, a CL model allows the choice set – the set of particular justices in the majority of the initial conference vote – to vary from case to case. Second, compared to alternative discrete choice modeling strategies, at the estimation stage a CL model brings to bear *a priori* more substantive information about the choice process under investigation – namely that for every case there is one and only one justice who is assigned to write the opinion – therefore mirroring the process driving opinion assignment more realistically.

The authors' theory is operationalized by two types of covariates: (1) legal case-specific variables that vary only across cases in the data set but are fixed across justices (e.g., *Political and Legal Salience, Winning Margin, Case Complexity and End of Term*), and (2) alternative-specific variables that also vary across choice alternatives in the model, i.e., the justices in the majority (e.g., *Ideology, Equity, Expertise, Workload*).

Typically, scholars employ a CL model in order to estimates alternative-specific effects on the likelihood of observing a particular choice behavior (Alvarez and Nagler, 1998, 66-71). The traditional conditional logit set-up, however, can be modified and "tricked" into a mixed version to model characteristics of the individual (cases) along with characteristics of the alternatives (justices) at the same time.

Following random utility theory we assume that individuals have utility functions in order to describe and compute their gain from choosing one alternative over the other. The utility for individual i ($i = 1, \ldots, N$) associated with the choice of j ($j = 1, \ldots, J$) is given by

$$U_{ij} = z'_{ij}\alpha + x'_i\beta_j + \epsilon_{ij} \tag{1}$$

whereby z_{ij} (including 1 as its first element) is the vector of measured characteristics of (alternative) justice j by individual case i. These characteristics vary across justices (choice alternatives), like *Ideology, Equity, Expertise.* Thus, we call them *justice-specific* characteristics. Furthermore, α is the vector of estimated coefficients indicating the impact of *justice-specific* characteristics on the likelihood of getting an assignment. Note there is only one set of coefficients (including $J - 1$ justice specific constants). We will call this the "CL-part" of the model since it resembles the typical conditional logit set-up.

Moreover the second part of the utility function represents what we call the "MNL-part" of the model since it resembles a typical multinomial logit (MNL) set-up. x_i (including 1 as its first element) is a vector of measured characteristics of the individual case i. These *case-specific* characteristics, for instance *Political and Legal Salience, Winning Margin, Case*

Complexity and End of Term, vary across cases but are fixed across the choice alternatives – namely the justices. For every case specific variable the model estimates $J-1$ sets of coefficients β_j, since one set of coefficients is (typically for discrete-choice models) set to 0 in order to identify the model (Alvarez and Nagler, 1998; Long and Freese, 2001). Thus β_j represents the vector of estimated *case-specific* characteristics of justice j relative to justice Burger. Overall for J justices we get $J-1$ sets of coefficients (including $J-1$ *justice-specific* constants) as in a typical MNL model set-up.

It is further assumed that the selection rule is simply to choose the alternative from which one gets the highest utility gain. The probabilities associated with assigning the opinion in case i to justice j can be written as

$$\Pr(y_i = j \mid z_{ij}, x_i) = P_{ij} = \frac{\exp(z'_{ij}\alpha + x'_i\beta_j)}{\sum_{j \in C_i} \exp(z'_{ij}\alpha + x'_i\beta_j)} \tag{2}$$

thereby allowing the choice set C_i to vary from case to case accounting for the fact that the likelihood of an opinion assignment is calculated *conditional* on the nature of the majority coalition that defines the choice set of the assignment process (hence the name *conditional logit*).

Simply specifying *case-specific* covariates as *justice-specific* variables does not do the "trick" because they do not vary across justices and hence will be dropped if estimated by standard software packages like STATA or LIMDEP. Nevertheless, as footnote 29 in the MSW text notes, this is apparently what the authors did. In order to set-up a mixed version of the CL model correctly, one has to specify the "MNL-part" appropriately. Since we are interested in how *case-specific* characteristics apply to the justices, we must interact these variables with j dummy variables for the justices (thereby using Chief Justice Burger as the baseline).

A mixed version of a CL model is a combination of a "standard CL-part", i.e. consisting of justice-specific covariates that vary across alternatives (justices), and a"MNL-part" consisting of $J-1$ dummy variables as choice specific constants together with its multiplicative terms based on all case-specific covariates (Greene, 2000; Long and Freese, 2001; Powers and Xie, 2000).

The estimated fixed effects represented by the justice-specific dummy variables indicate the average impact of unobserved (either case- or justice-specific) factors (Heiss, 2002, 229) on the decision makers' utility difference of assigning to himself versus any other alternative justices respectively that are not accounted for in the model. If the contribution of these unobserved factors on the likelihood of getting an assignment is non-zero and these fixed effects are not included in the systematic component of the model, then they are consequently absorbed into the stochastic component (error term) of the model thereby violating the assumption of a zero mean of an extreme value distributed error term (Train, 1986). This yields biased estimates.

More formally let us consider what happens if we exclude the fixed effects, represented by the *justice-specific* constants, from the systematic component. From the correct model specification in equation 1 we get

$$
\begin{aligned}
U_{ij} &= \tilde{z}'_{ij}\alpha + 1_i\alpha_j + \tilde{x}'_i\beta + 1_i\beta_j + \epsilon_{ij} && (3)\\
&= \tilde{z}'_{ij}\alpha + \tilde{x}'_i\beta + (1_i\alpha_j + 1_i\beta_j + \epsilon_{ij}) && (4)\\
&= \tilde{z}'_{ij}\alpha + \tilde{x}'_i\beta + \tilde{\epsilon}_{ij} && (5)
\end{aligned}
$$

Thus, the (unobserved) factors $1_i\alpha_j$ and $1_i\beta_j$ are consequently absorbed into the stochastic component (error term) of the model, whereby $\tilde{\epsilon}_{ij}$ denotes the new error term. By assumption the stochastic component of a CL model is distributed extreme value, i.e., it has zero mean. Because equation 1 represents the correct specification by assumption we know that $E(\epsilon_{ij}) = 0$. But what is the expectation of the new error term $\tilde{\epsilon}_{ij}$ including the variance usually picked up by $1_i\alpha_j$ and $1_i\beta_j$?

$$E(\tilde{\epsilon}_{ij}) = E(1_i\alpha_j + 1_i\beta_j + \epsilon_{ij}) \tag{6}$$

$$= E(1_i\alpha_j + 1_i\beta_j) + E(\epsilon_{ij}) = 1_i\alpha_j + 1_i\beta_j + 0 \tag{7}$$

Hence, we get unbiased estimates, if and only if

$$E(\tilde{\epsilon}_{ij}) = 0 \iff 1_i(\alpha_j + \beta_j) = 0 \tag{8}$$

Again, if *choice-* or here *justice-specific* constants are not explicitly in a CL model, all unobserved factors get absorbed into the error term. The model will then violate the assumption of a zero mean of the extreme value distributed error term of (conditional) logit models and yields biased coefficients if the unobserved factors are significantly different from zero (Train, 1986, 24-25). Including these fixed effects in the model relaxes an overly restrictive assumption that the contribution of all unobserved factors on the likelihood is non-zero. Using a similar argument one can show that one also gets nonsensical estimates if the main effects of included interactions are omitted.

In fact, MWS apply such a mixed model estimation strategy to the data because they derived hypotheses that relate to characteristics of the cases, attributes of the justices in the majority coalition as well as conditional relationships thereof. However, the MSW modeling strategy is plagued by both problems – as is related work on opinion assignment in the Rehnquist Court published in the *American Journal of Political Science* by two of the authors (Maltzman and Wahlbeck, 1996). They neither include the justice-specific constants nor do they correctly set-up the "MNL-part" of their model. Particularly, their *Self-Assignment* dummy is obviously collinear with an alternative specific constant for Chief Justice Burger. Therefore, in order to test hypotheses about self-assignment in political and legally salient cases (the so-called *Case Importance Hypotheses* 2a and b in Maltzman and Wahlbeck (1996)), for instance, the model must have the correctly specified "MNL-part". *Political* and *Legal Salience* are *case-specific* variables, therefore the product terms with (13-1) alternative specific dummy variables (taking justice Burger as baseline) are expected to be negative in all 12 sets of estimates of the "MNL-part" of the model.

Furthermore, in order to specify interaction terms, for instance, *Ideology* and *Winning Margin*, the component parts of these interaction terms must also be included. Setting up the "MNL-part" of their model correctly by multiplying the *case-specific* component parts by each *justice-specific* dummy variable will create the necessary *justice-specific* main effects (for 13-1 alternatives). Thus, these main effects vary across alternatives (justices) and, hence, do not get dropped out. The same logic holds for all the other interactive effects included in the original model.

Every CL model can be written as a generalized linear model (GLM). Since the (multinomial) link function (Liao, 1994, 60-61) is smooth and invertible it can be always passed back to the left-hand side of the equation. Therefore, interpreting interaction effects becomes analogous to the case of linear models (Friedrich, 1982; Gill, 2001) keeping in mind that the right-hand side is then the log odds of getting an assignment for a particular justice versus the baseline (here: CJ Burger). The conditional slope of an independent variable x_1 (say, *Winning Margin*) if interacted with x_2 (say, *Ideology*) does not only depend on the size of its main effect $\hat{\beta}_1$. Rather the slope of x_1 is a composite of its main effect and x_2 weighted by the size of the interaction effect $\hat{\beta}_3$, i.e. $(\hat{\beta}_1 + \hat{\beta}_3 x_2)\ x_1$.

What does this mean substantively? Excluding, for instance, the main effect for *Winning Margin* from the equation – as the authors' do – effectively constrains its estimated coefficient $\hat{\beta}_1$ to zero. Thus the substantive impact of *Winning Margin* is underestimated if $\hat{\beta}_1$ has the same sign as $\hat{\beta}_3 x_2$ or is overestimated otherwise.

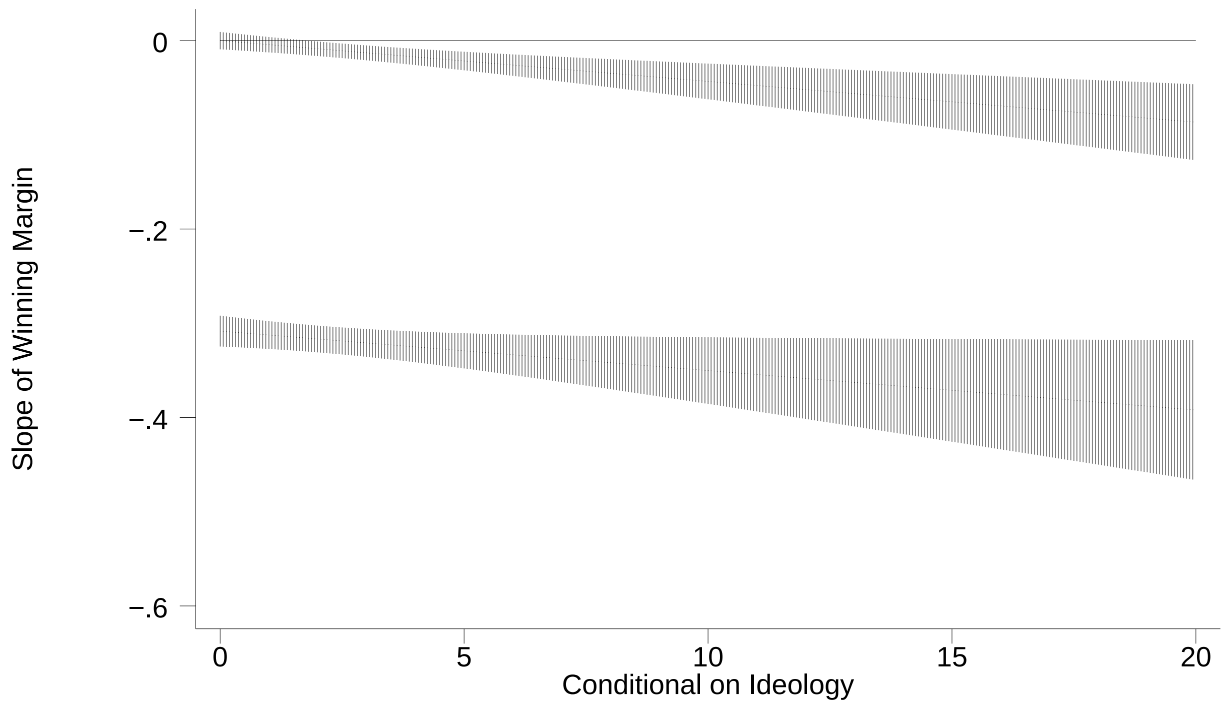

Figure 1: *A Comparison of the impact of the Size of Majority Coalition for the MSW model specification and our correct model specification:* The average simulated impact of the size of the majority coalition on the likelihood of getting an assignment conditional on ideological distance is graphed together with its estimated 90% confidence interval. The upper curve comes from the MWS model specification.

In the following figure 1 we simulate the averaged slope coefficient[2] across levels of *Ideology* and graph it. The upper curve comes from the authors' model specification. The extent of the bias is transparent by comparing predictions with the correct model specification (the lower curve). The upper curve is not statistically different from zero even for a one-tailed test with $\alpha = .05$ for small ideological distances implying that *Winning Margin* has no effect on the (log odds) likelihood of getting an assignment. In general both the simulated slopes for both models do not vary too much across levels of *Ideology*. Thus, the substantive impact of the hypothesized interaction between *Winning Margin* and *Ideology* is small. The predictions of both models do differ significantly. In fact only at the end of the *Ideology*-scale (at about 50) do both curves overlap. All other interactions can be interpreted analogously.

In Table 2, we report the results of the correctly specified CL model where we included constants, the fixed effects created by justice specific dummy variables, and all necessary component parts of the interaction terms MSW specified in the original model. This allows us to estimate *political* and *legal salience* effects for all justices versus Chief Justice Burger as a baseline instead of excluding them. We, therefore, are able to directly test whether self-assignment is particularly predominant in these cases.

Starting with a correctly specified model we can easily interpret the estimation results and evaluate MWS's *Case Importance Hypotheses* that Burger is more likely to assign politically and legally salient cases to himself. If this were true we should find significantly negative coefficients for *political* and *legal salience* in every set of estimates. In general, however, this is not true. Politically salient cases, according to this model, are not treated differently than non-salient cases, holding everything else equal. None of the coefficients for political salience is significantly different from zero while eight out of twelve of them do not even have the

[2]The averaged conditional slope of *Winning Margin* is ($\hat{\beta}_1 + \hat{\beta}_3$ *Ideology*). It is averaged because $\hat{\beta}_1$ is the average coefficient across all 12 estimated coefficients of the comparison of every particular justice on the court versus the baseline (CJ Burger). Note $\hat{\beta}_3$ is not averaged because only one coefficient is estimated anyway.

Independent Variables		Coef.	Std.Err.	*p*-value									
	Ideology	-0.014	0.010	0.170									
CLM part	Ideology X Winning Margin	-0.004	0.002	0.089				N = 12873					
	Ideology X Political Salience	-0.003	0.001	0.041				PCP = 0.244					
	Ideology X Legal Salience	0.007	0.012	0.543				PRE = 0.107					
	Equity	-0.110	0.036	0.003									
	Expertise	0.066	0.028	0.017									
	Freshman	-0.026	0.109	0.814									
	Freshman X Case Complexity	-0.042	0.110	0.704									
	Workload	-0.095	0.030	0.001									
	Ideology X End of Term	0.000	0.000	0.778									
		Black vs. Burger			Douglas vs. Burger			Harlan vs. Burger			Brennan vs. Burger		
		Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value
	Constant	2.714	0.933	0.004	2.051	0.740	0.006	-0.509	1.128	0.652	1.930	0.492	0.000
MNL part	Political Salience	-0.068	0.192	0.725	0.083	0.094	0.374	0.027	0.133	0.841	-0.029	0.059	0.621
	Legal Salience	-1.359	1.203	0.259	-1.481	0.842	0.079	0.479	0.918	0.602	-0.657	0.597	0.272
	Winning Margin	-0.456	0.221	0.039	-0.135	0.174	0.439	-0.332	0.218	0.129	-0.238	0.117	0.042
	Case Complexity	0.002	0.272	0.995	-0.307	0.195	0.117	-0.491	0.354	0.165	-0.097	0.115	0.403
	End of Term	-0.006	0.005	0.225	-0.002	0.004	0.531	0.006	0.005	0.282	-0.004	0.002	0.139
		Stewart vs. Burger			White vs. Burger			Marshall vs. Burger			Blackmun vs. Burger		
		Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value
	Constant	1.690	0.377	0.000	1.498	0.311	0.000	2.039	0.463	0.000	0.982	0.350	0.005
MNL part	Political Salience	-0.091	0.053	0.088	0.009	0.033	0.790	0.007	0.050	0.881	0.040	0.035	0.249
	Legal Salience	-0.671	0.441	0.128	0.004	0.338	0.990	-0.521	0.560	0.352	-0.202	0.388	0.603
	Winning Margin	-0.549	0.094	0.000	-0.477	0.076	0.000	-0.090	0.110	0.412	-0.229	0.083	0.006
	Case Complexity	0.050	0.111	0.655	0.097	0.091	0.285	-0.150	0.104	0.148	-0.324	0.116	0.005
	End of Term	-0.001	0.002	0.462	-0.002	0.002	0.222	-0.005	0.002	0.025	-0.002	0.002	0.252
		Powell vs. Burger			Rehnquist vs. Burger			Stevens vs. Burger			O'Connor vs. Burger		
		Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value	Coef.	Std.Err.	*p*-value
	Constant	1.338	0.298	0.000	1.234	0.295	0.000	1.863	0.421	0.000	1.042	0.429	0.015
MNL part	Political Salience	0.040	0.026	0.125	0.012	0.030	0.690	0.003	0.042	0.934	-0.043	0.045	0.348
	Legal Salience	-0.050	0.326	0.878	-0.024	0.373	0.950	-1.642	0.628	0.009	-0.852	0.648	0.188
	Winning Margin	-0.389	0.072	0.000	-0.275	0.068	0.000	-0.345	0.100	0.001	-0.184	0.103	0.072
	Case Complexity	-0.174	0.108	0.107	0.176	0.092	0.056	-0.184	0.129	0.154	-0.018	0.155	0.908
	End of Term	-0.003	0.001	0.073	-0.003	0.001	0.078	-0.002	0.002	0.351	-0.003	0.002	0.208

Table 2: *Corrected Specification:* A Conditional Logit Model of Opinion Assignment on the Burger Court.

expected sign. For legally salient cases the Case Importance Hypotheses is also generally *not* supported because a test that all 12 Legal Salience coefficients in Table 2 are jointly zero cannot be rejected ($p > .25$). We only find significant self-assignment effects compared to Justice Stevens. Our estimate indicates that Burger is ($1/\exp(-1.642) = 5.2$) approximately 5 times more likely to assign to himself than to Stevens. Thus apart from Justice Stevens in legally salient cases, we cannot conclude that there is a significant difference between Burger's choice of assigning to himself compared to the associate justices, in either legally or politically salient cases.

Nevertheless, the political salience of cases has a moderating effect on ideology. Ideological distance from the assignor becomes more important in predicting assignment in politically salient cases. The same justice will be slightly more likely ($.02\% = 1 - exp(-.003)\%$) to get an assignment in political salient cases. Thus, the ideological distance between the justices and CJ Burger does matter substantively for opinion assignment in politically salient cases holding everything else constant.

Moreover, the justice specific constants indicate the impact of unobserved factors, either case or justice specific factors that are not accounted for in the model (Heiss, 2002, 229), on average on the likelihood of getting an assignment compared to self-assignment. Since almost all the coefficients for these constants are different from zero this indicates that this model is missing justice specific variables or certain case specific variables that describe the context and the nature of an assignment decision in a particular case. Clearly, this suggests the need for more theoretical consideration and empirical research in this area. Nevertheless, we have at least tested whether these unaccounted for factors might exist by including these constants in the model. Even though they exist, as the significant coefficients suggest, and despite the fact that we cannot model them directly, we achieve the important goal of producing unbiased estimates. The MWS estimates in Table 1, however, remain biased since the excluded constants are significantly different from 0.

The coefficients of the CL model part indicate the overall impact of justice specific characteristics on the likelihood of getting an assignment. For instance, if we take the equity coefficient: the higher the number of associate justice assignments a justice receives the less likely that justice is to get the assignment. Our estimates indicate that for one more assignment from an associate justice, holding all other variables constant, the odds of getting the assignments will decrease by 10% ($1 - exp(-.110) = .10$). Also we find that, holding everything else constant, the more expertise a justice has and the lower the justice's workload is, the more likely that justice is to get the assignment. Again, these interpretations only depend on justice specific characteristics.

Besides the improved model fit – an almost 11% reduction in error of correctly predicted cases over a best-guess null model – the important difference is that we get weaker statistical support for the *Majority Coalition Size Hypothesis* because the associated standard errors get much wider. Additionally, a likelihood-ratio test shows that controlling for *Winning Margin*, the inclusion of the interaction of *Winning Margin* × *Ideology* does not significantly improve the model fit ($p > .08$) casting further doubt on the authors' *Majority Coalition Size Hypothesis*.

To summarize this section, other than the non-findings for the *Majority Coalition Size Hypothesis*, our results from using a correctly specified model produce less support for the hypotheses put forward by MSW. We get particularly divergent results when interpreting predicted probabilities associated with the interaction terms. We have made clear the extent to which MSW's findings are biased. In addition, a correct model specification makes transparent several points that have not been previously detected. Our model shows that various cases specific variables or personality factors describing the relationship between justices on the court are missing from our theoretical models. While we do not have a theoretical prediction for it, the data reveals an interesting working relationship between Burger and Stevens that the original model would have pushed into the error term. Further theoretical consideration might shed more light on this particular relationship.

5. CONCLUSION

Choice models present useful tools to political scientists because of the frequency with which we encounter limited dependent variables. Conditional logit models provide opportunities to simultaneously consider the effects of independent factors of both the choices and the choosers. Despite this usefulness, however, the requirements and assumptions of these models must be fully contemplated when they are utilized. We believe that our discussion above illustrates the specific hazards of using a CL model without consideration of an important element of the model, namely choice-specific constants.

Formally, we have demonstrated that failure to include choice-specific constants will yield biased results in conditional logit models. These models have to set up correctly in order to yield unbiased estimates. Scholars should specify choice-specific constants and enter them into the deterministic part of their models. This assures that the stochastic part of the model has a zero mean, as assumed in these types of choice-models.

Substantively, we have demonstrated that in the context of Supreme Court opinion assignment, this same failure will lead to incorrect inference and prediction about the impact of a variety of variables. We uncovered these problems by correctly specifying a CL model to test the effects of not only choice-specific characteristics, but also chooser-specific characteristics in a situation were the choice set is allowed to vary from observation to observation or from respondent to respondent. We made transparent the bias in estimated coefficients analyzing an example in judicial politics.

Our discussion should serve as a warning to researchers who seek to use these models in their own work. It is imperative to include choice-specific constants to control for the non-included, and often non-measurable, factors represented by the choices. Computationally, the solution is costless, while the alternative, incorrect inference is costly.

REFERENCES

Alvarez, R. Michael and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multi-Candidate Elections." *American Journal of Political Science* 42:55–96.

Epstein, Lee and Jack Knight. 1998. *The Choice Justices Make*. Washington, DC: CQ Press.

Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26:797–833.

Gill, Jeff. 2001. "Interaction Hierarchies in Generalized Linear Models." Paper presented at the Annual Meeting of the American Political Science Association.

Greene, William H. 2000. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Heiss, Florian. 2002. "Structural Choice Analysis with Nested Logit Models." *The Stata Journal* 2:227–252.

Liao, Tim Futing. 1994. *Interpreting Probability Models: Logit, Probit, and other Generalized Linear Models*. London: Sage.

Long, J. Scott and Jeremy Freese. 2001. *Regression Models for Categorical Dependent Variables using STATA*. College Station, TX: STATA Press.

Maltzman, Forrest and Paul J. Wahlbeck. 1996. "May It Please the Chief? Opinion Assignment in the Rehnquist Court." *American Journal of Political Science* 40:421–443.

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

Train, Kenneth. 1986. *Qualitative Choice Analysis. Theory, Econometrics, and an Application to Automobile Demand*. Cambridge, MA: MIT Press.