NBER WORKING PAPER SERIES

# LINEAR PROBABILITY MODELS OF THE DEMAND FOR ATTRIBUTES WITH AN EMPIRICAL APPLICATION TO ESTIMATING THE PREFERENCES OF LEGISLATORS

James J. Heckman
James M. Snyder, Jr.

Working Paper 5785

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1996

Heckman's work was supported by NSF Grant SES-91-11455, NSF Grant SBR-93-02348 and the American Bar Foundation. Snyder's work was supported by NSF Grant SBR-91-22488. We thank Henry Brady, Tim Bresnahan, Steven Donald, Mel Hinich, Hidehiko Ichimura and two anonymous referees for very helpful comments. Earlier versions of this paper circulated under the title "A Linear Latent Factor Model of Voting," and was presented at the Quandtfest, Princeton University, May 17, 1995. This paper is part of NBER's research program in Public Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper 5785
October 1996

# LINEAR PROBABILITY MODELS OF THE DEMAND FOR ATTRIBUTES WITH AN EMPIRICAL APPLICATION TO ESTIMATING THE PREFERENCES OF LEGISLATORS

**ABSTRACT**

This paper formulates and estimates a rigorously-justified linear probability model of binary choices over alternatives characterized by unobserved attributes. The model is applied to estimate preferences of congressmen as expressed in their votes on bills. The effective dimension of the attribute space characterizing votes is larger than what has been estimated in recent influential studies of congressional voting by Poole and Rosenthal. Congressmen vote on more than ideology. Issue-specific attributes are an important determinant of congressional voting patterns. The estimated dimension is too large for the median voter model to describe congressional voting.

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
jimh@cicero.spc.uchicago.edu

James M. Snyder, Jr.
Department of Political Science
Massachusetts Institute of Technology
77 Massachusetts Avenue, E53-457
Cambridge, MA 02139-4307

Richard Quandt contributed actively to three important developments in rational choice theory. (1) In his 1956 paper, he pioneered the application of random utility models developed in psychology to the analysis of consumer choices. These models use preference shocks to explain why persons with observationally-identical characteristics and opportunities make different choices. (2) In joint work with Baumol (1966), he formulated and estimated models for discrete choice that apply and extend the random utility framework to the analysis of choices among modes of travel. Their work supplements the random utility model by introducing heterogeneity in decision-maker preferences, and omitted attributes of choice alternatives to explain differences in choices made by observationally identical persons.[1] (3) In the same research, Baumol and Quandt apply the attribute theory of Lancaster (1966; 1971) and Gorman (1980) to solve the problem of forecasting the demand for a new good. By reformulating the demand for goods in terms of the demand for attributes that comprise the goods, and by assuming that new goods repackage the same attributes, they demonstrate how it is possible to forecast the demand for a new good from information on the purchase behavior of old goods.

The theory of rational choice, extended to account for random preference shocks, discreteness in choices and heterogeneity in preferences among decision-makers and formulated in terms of attributes of choices, is widely applied in the social sciences. This paper builds on Quandt's research. In it, we formulate and estimate a linear probability model of rational discrete choice among two alternatives characterized by attributes that are unobserved by the analyst but observed and acted on by the decision makers we model. Our choice probabilities are linear both in parameters and attributes of choices and hence they are easily computed.

The linear probability model is widely used in empirical research as a simply-computed but *ad hoc* approximation to what are widely regarded as more theoretically appropriate non-linear discrete choice models. Our paper provides <u>two</u> distinct rigorous justifications for the linear probability model.

Prejudice against the linear probability model is based in part on results in Pratt and Gibbons (1981, p. 395 and 420, problems 64 and 65) who demonstrate that under conventional assumptions made in random utility theory, the linear probability model cannot be a representation for any random utility model. We show that relaxing one minor assumption in the random utility model - that preferences over outcomes receive independent shocks from a *common* distribution - makes it possible to justify the linear probability model as an *exact* representation of a well-formulated random utility model. The assumption made in random utility theory, that shocks to preferences are identically distributed, is an artificial convention. Relaxing it produces a simply-computed easily-interpreted discrete choice model. This extension of the classical random utility model is one of two basic justifications we offer in this paper for a rigorous defense of the linear probability model.

This paper extends conventional models of discrete choice in econometrics that recover preferences for attributes measured by the empirical analyst. Our approach builds on the pioneering work of Gorman (1980) who developed methods for recovering prices for unobserved attributes of goods in a continuous-choice setting. Our analysis extends Gorman's method to a discrete choice setting in which attention focuses on estimating individual valuation parameters for unobserved characteristics.[2] It provides a rigorous econometric foundation both for Gorman's approach and our own. We present methods for consistently estimating the preferences of consumers over unobserved (by the analyst) attributes. In this context, we provide a second and more basic rigorous justification for the linear probability model as an *exact* representation of any discrete-choice model formulated in terms of unobserved characteristics. We present formal methods for determining the effective dimensionality or rank of the attribute space. These methods have applicability to the estimation of Gorman-Lancaster hedonic models for the pricing of attributes and the demand for attributes. There are many potential empirical applications of this model in labor economics, industrial organization and political economy.

This paper applies our framework to analyze the voting decisions of legislators. The decision to vote yes or no on a bill is modelled as the outcome of a rational choice process in which legislators use their preferences to weigh the attributes or characteristics of each bill they consider. Outside observers do not measure these characteristics but legislators act on them.

The "dimensionality" - really the rank - of the attribute space of bills plays a central role in recent studies in quantitative political science. (See, *e.g.* Poole and Rosenthal 1985a, 1991a). In that literature, a low-rank attribute space for a variety of very different bills is interpreted as arising from a common ideology - or party affiliation - across legislators. In this view, all bills, no matter how diverse, are judged on a common ideological basis. Intrinsic attributes of bills above and beyond their location on the ideological scale such as their value to particular constituents, are considered to be secondary features of congressional voting. A high-dimensional attribute space with different attributes affecting the votes on different issues is interpreted as evidence that issue-specific features of proposed legislation affect the votes of Congressmen, and is also inconsistent with a median voter model which requires that at most one or two dimensions govern voting behavior. (Riker, 1990).

Recent studies in political science apply various extensions of random utility models or "ideal point" models such as those of Enelow and Hinich (1984, 1985, 1989) to address this question. Bartels and Brady (1993) provide an insightful review of this literature. The empirical work of Poole and Rosenthal has been the most influential. Using a nonlinear-in-parameters logit model, they estimate preferences of congressmen and the rank or effective dimension of the attribute space assuming - like Gorman - that none of the attributes of choices can be observed. Their evidence on the low dimensionality of the attribute space is interpreted as proving the importance of ideology in explaining congressional voting.

Using the econometric framework developed in this paper, we re-examine this issue. We demonstrate the fundamental lack of identification of decision-maker preferences in choice settings where attributes are not observed. This point is implicit in Gorman (1980) but has not received the attention it deserves. Certain statistical conventions are required to identify preferences and attributes and these conventions are only weakly justified by social or economic theory. However, the effective dimension of the attribute space can be estimated under much more general conditions.

In our model, application of these conventions produces an econometric framework in which it is possible to (a) consistently estimate decision-maker preferences, (b) produce a rigorous asymptotic distribution theory for our estimator and (c) derive a rigorous test for the effective dimension or rank of the attribute space. In the influential model of Poole and Rosenthal, application of the conventions used to identify the Gorman model does not lead to identified models that can be consistently estimated nor can the choice of the rank of the attribute space be rigorously justified. In addition, their model is computationally very difficult to estimate whereas ours requires only very simple calculations.

At the heart of the theoretical difficulty with the Poole-Rosenthal estimator is the "incidental parameters" problem first analyzed by Neyman and Scott (1948). In the study of congressional voting, attributes of any single bill are sampled only once. Hence it is impossible to consistently estimate the attributes of any particular bill, since consistency requires multiple observations on the same bill. This problem also arises in estimating the demand for attributes of a particular good which can be thought of as one draw from a distribution of potential attributes. With only one realization, it is not possible to consistently estimate the attributes of any good although it may be possible to estimate the distribution of the attributes of all goods if they come from a common distribution.

In nonlinear voting models, it is generally not possible to separate the estimation of preference parameters from the estimation of attributes of individual bills. Both preference parameters and attributes have to be estimated jointly. Inconsistency in estimating attributes is transmitted to estimators of preferences. In our linear voting models, it is possible to separate the estimation of preference parameters from the estimation of attributes. In this way we avoid the inconsistency that plagues the Poole-Rosenthal estimator. The nonlinear estimator proposed by Brady (1989) also avoids the inconsistency problem arising from estimating incidental parameters. He follows a suggestion of Kiefer and Wolfowitz (1956) and assumes a low-dimensional distribution for attributes of bills and simultaneously estimates the preferences of legislators and the *distribution* of bill characteristics. Our estimator is simpler and easier to compute than Brady's in part because we avoid the problem of estimating the distribution of attributes. The advantage of our approach over his is simplicity and ease of computation. Our model reproduces in a linear framework the predictive content of any nonlinear model. It formally justifies the application of simply-computed, factor-analytic models of voting that were once widely used in the analysis of voting behavior but which were driven out of use by apparently more rigorous nonlinear voting models. Our analysis suggests that it is the simple methods that are the rigorously-justified ones and their application fundamentally alters the interpretation placed on congressional voting patterns.

Using our apparatus, we reanalyze data on congressional voting previously analyzed by Poole and Rosenthal. Using formally-justified statistical testing procedures, we determine that at least five and perhaps as many as eight attributes are required to rationalize congressional voting patterns. In terms of the literature in political science, more than ideology is required to explain voting behavior. The extra effective dimensions of a model that we estimate are practically important in explaining the voting of congressmen. We demonstrate that congressional voting is based in part on issue-specific characteristics and not just ideology. However, our results do not support pressure group models that do not allow for logrolling or coalition building among groups, as in Stigler (1971) or Peltzman (1976). If each special interest lobbies alone, we would expect to find nearly as many voting dimensions as there are interest groups unless different interest groups merely repackage the same low-dimensional attributes. Instead, we estimate only a handful of dimensions. Thus our empirical analysis supports a model midway between a model of pure ideology and a model of pure interest groups. It does not support a median voter model for Congress.

We also demonstrate that if the rank of the model is fixed in advance of estimation, our estimator produces estimates that are very close to those obtained from the Poole and Rosenthal estimator. For legislatures as large as the U.S. House and Senate, the practical consequences of the statistical inconsistency of their estimator are slight so long as the rank of the model is pre-specified. Since our linear estimator is much simpler to compute, this evidence strongly supports its use in place of more complicated estimators of voter preferences. However, the problem of the inconsistency and the absence of a rigorous statistical theory for the more complicated estimators affects the choice of the rank of the model, and this has important interpretive consequences.

We further demonstrate that the first two factors that explain much of the variance in the voting data are closely related to party and region. We find that party and region predict about as well, and sometimes better, than two unobserved factor models. We present an analysis of voting on civil rights bills that dramatically illustrates the importance of looking at higher dimensions in interpreting conflict between committee preferences and entire House and Senate preferences.

This paper develops in the following way. Section I presents a model of discrete choice with both observed and unobserved attributes. Throughout, we assume that the persons we study observe all attributes relevant to choices so that the categories "observed" and "unobserved" refer to the information available to the analyst. We demonstrate that a variety of widely-used models of voting and the purchase of ordinary consumer goods are special cases of a model of quadratic utility. Section II discusses estimation of the model and the choice of the dimension of the characteristics space. Section III compares our estimator to two other estimators that have been proposed for estimating preferences when the attributes of choices are not known. Section IV presents an empirical analysis of congressional voting and contrasts the evidence obtained from our model with claims advanced in the empirical literature on voting.

## I. *The Model*

We assume that decision maker *j* has preferences defined over a fixed set of characteristics that describe *N* distinct binary choice problems. Associated with each potential choice is a pair of utility functions $U_0(i)$ and $U_1(i)$ whose structure is now characterized. For each choice problem, there is an associated pair of vectors of characteristics $X_{0i}$ and $X_{1i}$ that enter as arguments of $U_0(i)$ and $U_1(i)$ respectively. Characteristics may include direct physical attributes, as well as prices and costs associated with the two alternatives in any choice setting. Thus $(U_0(i), U_1(i))$ may be direct, indirect, or mixed direct-indirect utility functions. In the context of voting, ($X_{0i}$ and $X_{1i}$) are vectors of characteristics of a bill and the alternative. $(U_0(i), U_1(i))$ may also be random functions indexed by *j*. Persons facing the characteristics of choices may place different valuations on the two options in any particular choice problem. This source of randomness is sometimes called preference heterogeneity.

Decisions to purchase a good, or vote for a bill are thus the outcome of a comparison of two potential utilities that characterize the two alternatives as in Thurstone (1927). In the context of voting, one outcome for a choice trial is a "yea" vote and the other is a "nay" vote. In what follows, we denote the characteristics of the "yea" or "purchase" alternative on choice *i* by $X_{1i}$, and the characteristics of the corresponding "nay" or "no purchase" alternative by $X_{0i}$. Thus, in deciding whether to purchase good i or not, or to vote yea or nay on roll call i, agents compare the utility of the attributes of the two potential outcomes $X_{0i}$ and $X_{1i}$. A major simplification of the theory postulates that a common utility function *U* evaluates all options in all choice settings. Then rational actor *j* chooses option 1 in choice *i* if and only if the utility of option "1" exceeds the utility of option "0":

$$U(X_{1i};j) > U(X_{i0};j) \quad (1)$$

To simplify the argument we assume that ties are zero-probability events.

*(a) Sources of Differences in Behaviour*

Identical persons make different choices because of (1) fundamental randomness across people ("shocks") in the utility assessment of the same set of alternatives (Quandt, 1956), (2) utility-relevant characteristics not adequately captured by the measured ($X_{0i}$, $X_{1i}$) vectors, or because (3) persons differ in their evaluation of the same bundle of characteristics in a stable way across choice settings. (This is called "preference heterogeneity" in the discrete choice literature. (See Quandt and Baumol, 1966; Domencich and McFadden, 1975). The distinction between (1) and (3) is based on the distinction between externally-imposed preference shocks that are assumed to be independent across persons and across choice settings even for the same person (source 1) and differences across decision makers that are stable across choice settings (source 3). Let $\varepsilon_{\ell ij}$, $\ell = 0,1$ denote person-specific and choice-specific shocks (source 1). These shocks affect the utilities of the two options along with the attributes. We write the utility pair for choice problem $i$ as

$$(U(X_{1i};\varepsilon_{1ij};j),\ U(X_{0i};\varepsilon_{0ij};j))\ .$$

Person $j$ chooses to purchase option 1 in choice setting $i$ if

$$U(X_{1i};\varepsilon_{1ij};j) > U(X_{0i};\varepsilon_{0ij};j)\ . \tag{2}$$

In principle, with panel data and controlled variation in the characteristics ($X_{0}$, $X_{1i}$) these three sources of heterogeneity in responses among persons can be distinguished. Random utility shocks (1) cause persons with the same preferences who face the same attributes in different choice settings to make different choices in different settings. Omitted attributes of choice options that are common across persons (Source 2) produce dependence in the choices across persons with the same utility pairs and preference shocks. Preference variation arising from Source 3 causes persons confronting the same attributes of choice options and experiencing the same preference shocks to make different choices.[3] Although conceptually important, these distinctions are of little empirical consequence when some of the components of $X_{1i}$ or $X_{0i}$ are unobserved.

Assuming that the conditional expectation of $U$ exists, we may always write

(3) $$E[U(X_{\ell i};\epsilon_{\ell ij};j)\,|\,X_{\ell i})] = V(X_{\ell i};j) \qquad \ell = 0,1$$

and define

(4) $$\epsilon(X_{\ell i};j) = U(X_{\ell i};\epsilon_{\ell i j};j) - V(X_{\ell i};j) \qquad \ell = 0,1.$$

The $V$ are the mean scale utility functions featured in Baumol and Quandt (1966) and McFadden (1974).

### *(b) A Quadratic Specification*

Models that are quadratic in attributes dominate the empirical applications of the theoretical literature on discrete choice and voting. In this subsection, we abstract from preference shocks as defined in the preceding subsection. In voting models, differences among legislators in preferences are assumed to be captured by a finite-dimensional vector $\alpha_j = (a_j, A_j)$ producing the specification

(5) $$V(X;\alpha_j) = a_j'X - X'A_jX$$

where $A_j$ is symmetric. Actor j chooses the "1" option (votes yea) in choice setting $i$ if

(6) $$a_j'(X_{1i} - X_{0i}) + X_{1i}'A_jX_{1i} - X_{0i}'A_jX_{0i} > 0.$$

This model incorporate preferences heterogeneity since $a_j$ and $A_j$ may vary among decision makers. The "Euclidean Spatial Voting" model of Enelow and Hinich, (1984), writes

(7) $$V(X;\alpha_j) = -(X-L_{j\alpha})'(X-L_{j\alpha})$$

where $L_{j\alpha}$ is the "ideal point" (bliss point) for voter $j$. Utility declines the farther the decision maker is from the ideal point. Observe that setting $A_j = I$ and $a_j = 2L_\alpha$, in (5), inequality (6) becomes

$$2L_{\alpha j}'(X_{1i} - X_{0i}) + X_{1i}'X_{1i} - X_{0i}'\,X_{0i} > 0.$$

The "directed preference" model of Rabinowitz and McDonald (1989) writes $A_j = 0$. In the language of utility theory, the "$a_j$" are the marginal utilities of the attributes for decision maker $j$. This is the most commonly-utilized specification in the empirical literature on discrete choice.

*(c) Deriving Choice Probabilities: Our First Rigorous Justification for the Linear Probability Model*

Let $N$ be the number of binary choices made by the decision maker, and let $\{(X_{0i},X_{1i})\}_{i=1}^{N} = X$ denote the characteristics of the alternative options that characterize these choices settings. If $\alpha_j = (a_j,A_j)$ and $X$ are known, then choices are strictly deterministic if we condition on $\alpha_j$ and $X$ and abstract from random utility shocks. "Randomness" in this setting arises solely from components of the $X$ or $\alpha_j$ on which we do not condition.

The parameter $\alpha_j$ is fixed for each decision maker, although observers may not know its value. Randomness in choices from the point of view of the observer is obtained if he or she cannot condition on *all* components of $X$, or if there are, in fact, random "shocks" to preferences. Since all decision makers face common values of $X$, omitted $X$ components induce dependence in voting outcomes across decision makers whose preferences have non-zero weights on the unobserved $(X_{0i}, X_{1i})$. The preference shocks $\epsilon_{0ij}$, and $\epsilon_{1ij}$ are conventionally assumed to be distributed and (a) independent across decisions for the same decision maker, (b) independent of observed choice characteristics, and (c) independent across decision makers on the same choice.

Decompose $X$ into $((X_{0i}^{o}, X_{oi}^{u})\ (X_{1i}^{o}, X_{1i}^{u}))$, where "$o$" refers to observed attributes and "$u$" refers to unobserved attributes. We assume that preference shocks are unobserved. (Recall that "observed" and "unobserved" are statements about the analyst's information). Let $d_{ji}$ be a choice indicator for agent $j$ in choice setting $i$. $d_{ji} = 1$ if the agent chooses the "1" alternative. It is "0" otherwise. The probability that decision maker $j$ chooses option "1" in choice setting $i$ is

$$(8) \qquad Pr(d_{ji} = 1 | X_{1i}^{o}, X_{0i}^{o}) = Pr[U(X_{1i}, \varepsilon_{1i}; j) > U(X_{0i}, \varepsilon_{0i}; j) | X_{1i}^{o}, X_{0i}^{o}].$$

A choice setting can be a vote on a specific bill or a purchase decision where two goods are available to the decision maker. (No purchase or the default state of a no vote is viewed as one choice).

The conventional random utility model conditions on $(X_{1i}^{o}, X_{0i}^{o})$ and absorbs the omitted characteristics into error term (4). Only if observed and unobserved characteristics are statistically independent can the parameters of mean utility $V$ (for the observed characteristics) be identified up to scale. Such independence is typically assumed. In this paper we follow these conventions in the literature but note that the conditions required to justify them are strong.

For notational simplicity, we drop the "$o$" and "$u$" notation and henceforth define $X_{0i}$, and $X_{1i}$ as observed characteristics for options "0" and "1" on choice $i$. We define $\eta(i,j) = \epsilon(X_{1i},j) - \epsilon(X_{0i};j)$. Assuming $\eta(i,j)$ is statistically independent of $V(X_{1i}; j) - V(X_{0i};j)$, and iid,

$$Pr(d_{ji} = 1 | X_{0i}, X_{1i}, \alpha_j) = 1 - G_\eta[(V(X_{1i};\alpha_j) - V(X_{0i};\alpha_j))]$$

where $G_\eta$ is the distribution function for $\eta$.

In their analysis of voting, Poole and Rosenthal (1985a, 1985b, 1991a, and 1991b) follow traditions in the econometric discrete choice literature. However, they work with a monotonic transformation $\rho: R^1 \rightarrow R^1$ of the $V$ functions, and define mean utilities by

$$V(X_{1i};\alpha_j) = \rho((X_{1i} - L_{\alpha j})'(X_{1i} - L_{\alpha j}))$$

where $\rho(d) = \beta\exp(-d/8)$, $\beta > 0$, which is strictly decreasing in $d$. They assume that $\eta$ has a logit distribution. Then, using the fact that for the logit, $G(\eta) = 1 - G(-\eta) = \exp(\eta)/[1+\exp(\eta)]$, their model for the probability of voting is

$$(9) \quad Pr(d_{ji}=1 | X_{1i}, X_{0i}, L_{\alpha j}) = \left[ \frac{\exp(\rho[(X_{1i}-L_{\alpha j})'(X_{1i}-L_{\alpha j})])}{\exp(\rho[(X_{1i}-L_{\alpha j})'(X_{1i}-L_{\alpha j})]) + \exp(\rho[(X_{0i}-L_{\alpha j})'(X_{0i}-L_{\alpha j})])} \right].$$

A linear probability alternative to their model, that is one of two linear probability models presented in this paper, assumes that $\eta$ is uniform and independently but not necessarily identically distributed across choices and decision makers. Letting

$$\overline{M}_{\alpha j} = Max\ [V(X_{1i};\alpha_j) - V(X_{0i};\alpha_j)],$$
$$\underline{M}_{\alpha j} = Min\ [V(X_{1i};\alpha_j) - V(X_{0i};\alpha_j)], \quad i = 1,\ldots,N, \ j = 1,\ldots,J$$

and letting

$$G(\eta_{\alpha j}) = \frac{\eta_{\alpha j} + \bar{C}_{\alpha j}}{(\overline{M}_{\alpha j} - \underline{M}_{\alpha j})}, \qquad j = 1,\ldots,J$$

where $\bar{C}_{\alpha j}$ is a location parameter, the probability that option 1 is selected on choice i is

$$(10) \qquad Pr(d_{ji} = 1 | X_{1i}, X_{0i}, \alpha_j) = [V(X_{1i};\alpha_j) - V(X_{0i};\alpha_j)]/[\overline{M}_{\alpha j} - \underline{M}_{\alpha j}] + C_{\alpha j}$$

where $C_{\alpha j} = \dfrac{\bar{C}_{\alpha j}}{(\overline{M}_{\alpha j} - \underline{M}_{\alpha j})}$.

The uniform distribution of utility differences arises trivially when one of the two error terms $\varepsilon(X_{1i},j)$ *or* $\varepsilon(X_{0i};j)$ is degenerate and the other is uniform. It can also arise in a random utility model as the difference between two nondegenerate random variables. If $\eta$ is $U(-M,M)$, its characteristic function is $E(e^{i\eta t}) = (Sin\ Mt)/Mt$. This can be written as the product of two characteristic functions, $[(Sin\ (Mt/2)/(Mt/2)]\cdot[Cos\ (Mt/2)]$. Thus a uniform random variable can be represented as the sum or difference of two nondegenerate random variables as required for it to be derivable from a random utility model.[4]

In the literature on discrete choice, model (10) is often viewed as *ad hoc* because it cannot be justified as a random utility model. The correct claim that uniformly-distributed random utility differences cannot be produced as the difference between two nondegenerate and independently and *identically* distributed utility shocks relies critically on the assumption that both shocks come from the same distribution.[5] If preference shocks are restricted to come from a common distribution, no random utility model rationalizes (10). However if preference shocks are permitted to come from different nondegenerate distributions, the linear probability model can be rationalized as a random utility model. In the context of voting choices or in the context of analyzing purchases of goods, there is no necessary symmetry between the "1" state and the "0" state so our assumption is not artificial. Indeed, it is the assumption of symmetry that is artificial.

Using representation (5) for the mean scale utility, and letting $\overline{M}_{\alpha j} - \underline{M}_{\alpha j} = M_{\alpha j}$ we may write (10) as

(11) $$Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j) = [a_j'(X_{1i}-X_{0i}) + X_{1i}'A_jX_{1i} - X_{0i}'A_jX_{0i}]/M_{\alpha j} + C_{\alpha j}.$$

This can be written as a regression equation with error $\tau_{ji}$:

$$\tau_{ji} = d_{ji} - Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j) = d_{ji} - E(d_{ji} = 1\,|X_{1i},X_{0i},\alpha_j),$$

where $E(\tau_{ji}) = 0$, and $Var(\tau_{ji}) = (1-Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j))\cdot Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j)$ where $\alpha_j = (a_j, A_j)$.

In this framework, the directional voting model of Rabinowitz and MacDonald may be written as a linear probability model:

(12a) $$Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j) = [a_j'(X_{1i} - X_{0i})]/M_{\alpha j} + C_{\alpha j}.$$

The Euclidean spatial voting model of Enelow and Hinich may be written as

(12b) $$Pr(d_{ji}=1\,|X_{1i},X_{0i},\alpha_j) = [2L_{\alpha j}'(X_{1i} - X_{0i}) + X_{1i}'X_{1i} - X_{0i}'X_{0i}]/M_{\alpha j} + C_{\alpha j}.$$

A more conventional justification for the linear probability model is as an approximation to (7) using a standard Taylor approximation argument. (Goldfeld and Quandt, 1972). Below we present a second and conceptually distinct justification for the linear probability model and demonstrate that it is one exact representation of any model of discrete choice when all of the characteristics are unobserved as they often are in the Gorman-Lancaster model of consumer choice or in the Poole-Rosenthal model of voting.

## II. *Formulating The Empirical Model*

We now formulate the population counterparts to the theoretical models presented in the preceding section. We present a second and more basic justification of (10) when all the attributes are not observed. This is the most challenging case. An important question that arises in that situation is what can be identified? We demonstrate that the rank of the model is identified even if particular parameters of the model are not. We consider both population definitions and sample inference procedures.

### (a) *The Model When X is Observed*

When $X$ is observed, the statistical analysis of the model (10) is straightforward. In inner-product form this probability can be written as,

(13) $$Pr(d_{ji}=1 \mid X_{1i}, X_{0i}, \alpha_j) = [a_j'(X_{1i}-X_{0i}) + (Vec\ A_j)'(Vec(X_{1i}X_{1i}' - X_{0i}X_{0i}'))]/M_j + C_{\alpha j}.$$

where Vec $A$ is a vectorization of matrix $A$ (converting $A$ from a matrix into a vector). There are many possible vectorizations, and the choice of a particular one is irrelevant provided the operation is performed consistently.

For each decision maker $j$, it is possible to estimate $a_j$ and $A_j$ by least squares or maximum likelihood (see, *e.g.* Goldfeld and Quandt, 1972), using data on votes ($d_{ji}$) and characteristics of bills (the $X_{1i}, X_{0i}$). Under standard rank conditions, ordinary least squares consistently estimates ($a_j, A_j$) up to the unknown scale parameter $M_j$ for each decision maker. Absorb the $C_{\alpha j}$ into the constant term. Asymptotic normality follows from standard conditions. Goldfeld and Quandt (1972) or Amemiya (1975, 1983) discuss more efficient maximum likelihood estimation in this model.[6]

Using standard regression analysis, one can test between the competing models (12a) and (12b) (see Theil, 1971). The "directed distance model" (12a) is the conventional linear probability model and assumes that $A_j = 0$, so no interaction or quadratic terms are present. The "spatial proximity model" (12b) assumes that $A_j$ = I. In terms of scaled parameters, this model implies that the coefficients of all interactions are zero and that all coefficients on quadratic terms are the same. If this restriction is not rejected, then the common coefficient on the quadratic terms provides an estimate of $M_{\alpha j}$. Unless $A_j$ is restricted to a known non-zero value as in the "spatial proximity model", it is not possible to decompose $a_j/M_j$ into its component parts.

The "rank" $R$ of the model is $K + K(K+1)/2$, the assumed rank of the regressor matrix. The number of "basic" variables determining choice behavior is $K$, but these variables may have nonlinear as well as linear effects on decisions. The additional $K(K+1)/2$ terms arise from squares and interactions among the variables. There is no requirement that the components of $X$ be orthogonal or independent, since multiple regression allows for interdependence among the components of $X$.[7]

In the case of a full-rank data matrix, it is possible to estimate $(a_j, A_j)$ up to the scale parameter $M_j$, and to test for similarity among the preferences of voters using the large-sample asymptotic normality of the estimators. One can consistently estimate the population distribution of preferences using the empirical distribution of the estimated $(a_j, A_j)$.

(b) *The Model When X is Unobserved*

The simplicity and familiarity of the standard linear model vanishes when X is unobserved. The model is not identified unless further restrictions are imposed. At issue is whether it is possible to learn *anything* about decision maker preferences under these conditions. We first consider issues of identification and model formulation, deferring discussion of inference to subsection (d).

One way to approach this problem suggested by the analyses of Gorman (1980) and Lancaster (1966; 1971) is to represent the characteristics of choices and preferences in factor-analytic form. The factors are the unobserved attributes of two potential options for choice setting $i$, $(X_{1i}, X_{0i})$ and the factor loadings are the preferences of the decision makers $(a_j, A_j)$. Our inference problem is similar to that studied by Gorman, except in this paper we deal with discrete choices whereas he considers continuous demands and pricing functions. Many of the identification problems that arise in our model arise in the setting of his model and are more simply discussed there. To analyze this problem, we briefly review Gorman's model.

He interprets the price paid for good $\ell$ at time $t(P_{\ell t})$ as the sum of the prices of the attributes or characteristics that constitute the good. Let $c_{\ell q}$, $q = 1,..,Q$ be the quantities of characteristics embodied in good $\ell$. (These correspond to our $(X_{0i}, X_{1i})$). Array these into a $1 \times Q$ dimensional vector $C_\ell$. Let $w_{qt}$ be the price of characteristic $q$ at time $t$. (These correspond to the preference parameters).The vector of prices is $W_t$. Then

$$P_{\ell t} = C_\ell' W_t \qquad \ell = 1,...,L$$
$$t = 1,...,T.$$

Allowing for mean zero measurement error in the prices, $\varepsilon_{\ell t}$, uncorrelated over goods ($\ell$) and time ($t$), and uncorrelated with $C_\ell' W_t$, produces a *factor structure* model for observed prices $P^*_{\ell t}$ *i.e.*

$$P^*_{\ell t} = C_\ell' W_t + \varepsilon_{\ell t} \qquad \ell = 1,...,L$$
$$t = 1,...,T.$$

For each good $\ell$, this model implies that the time series of prices is a $Q$ factor model with time-varying factor loadings. Formal tests for rank $Q$ factor structures can be based on the methods described in subsection (c). The Gorman model imposes testable restrictions on price data.[8]

Since only $P^*_{\ell t}$ is observed, it is not possible to uniquely identify $C_\ell$ or $W_t$. Any orthogonal transformation $F$ can be applied and the transformed prices and qualities describe the price data equally well. Thus

$$\tilde{C}_\ell = FC_\ell \qquad \text{and} \qquad \tilde{W}_t = FW_t$$

rationalize $P_{\ell t}$ if $C_\ell$ and $W_t$ rationalize the price data since

$$\tilde{C}_\ell' \tilde{W}_t = \tilde{C}_\ell' F' F W_t = C_\ell' W_t \ .$$

One statistical way to resolve this lack of identification of price and characteristics parameters is to impose mutual orthonormality on the characteristics vectors. (Mutual orthogonality and restriction to unit length.) Nothing in the economics of the problem produces this mutual orthonormality and it is clearly only an identifying convention. Some such convention is required to identify either the characteristics or their prices.

To apply these insights to our discrete-choice setting, the unobserved attributes or combinations of them, correspond to the "factors" or attributes in the Gorman model, and the "loadings" are the preference parameters of decision makers. Indeed, in a representative agent model, the "loading" or prices in the Gorman model are the marginal valuation parameters for attributes of the representative agent.[9] (Pudney, 1980). Thus, the Gorman model has the same formal structure as the model equation 12(a). Imposing mutual orthonormality on the attributes ($X_{1i}$-$X_{0i}$, Vec($X_{1i}X_{ii}'$) - Vec($X_{0i}X_{0i}'$)) in (13) produces an identified version of the model when the $X_{0i}$ and $X_{1i}$ are not observed. Thus far we have developed the parallelism between the special linear probability model (10) and factor structure models. We now present a second and more fundamental justification for factor models as an *exact* representation of any binary choice model formulated in terms of unobserved characteristics of choices.

We claim that *any* model formulated in terms of unobserved attributes *e.g.*

$$Pr(d_{ji} = 1 \mid X_{1i}, X_{0i}, \alpha_j) \qquad j = 1,\dots,J;\ i = 1,\dots,N$$

where $X_{1i}$ and $X_{0i}$ are unobserved, may be represented in a linear factor structure form. The specific functional form assumed in (13) is only an instance of the class of all probability models formulated in terms of unobserved attributes that can be approximated by linear latent-factor probability model.

Using the singular-value decomposition theorem (*e.g.* Rao 1973), any matrix of choice probabilities can be decomposed into an orthogonal-factor form. Let

$$\bar{P} = (Pr(d_{ji}=1 \mid X_{1ji}, X_{0ji}, \alpha_j)), \qquad j = 1,\dots,J; \quad i = 1,\dots,N.$$

A standard result in linear algebra states that any $J \times N$ matrix $\bar{P}$ can be represented by a min($J,N$) -dimensional set of basis vectors. The singular-value decomposition (SDV) theorem operationalizes this result to produce orthonormal basis vectors. More precisely the SDV theorem states that we can exactly represent any $\bar{P}$ matrix by

$$\text{(14)} \qquad \bar{P} = \sum_{\ell=1}^{\mu} \lambda_\ell^2\, \bar{\varphi}_\ell\, (Z_\ell)' \qquad \mu = Min(J,N)$$

where $\{\varphi_\ell\}_{\ell=1}^{\mu}$ is a collection of $J\times 1$ vectors, $\{Z_\ell\}_{\ell=1}^{\mu}$ is a collection of $N\times 1$ vectors, and $\{\lambda_\ell^2\}_{\ell=1}^{\mu}$ is a collection of non-negative scalars that are the eigenvalues of $\bar{P}\,\bar{P}'$. These vectors and scalars are obtained as solutions of the following pair of equations:

(15a) $$\bar{P} Z = \lambda\bar{\varphi}$$

(15b) $$\bar{P}'\varphi = \lambda Z.$$

Solving these equations, we obtain

(15c) $$\bar{P}'\,\bar{P}\,Z = \lambda^2 Z$$

(15d) $$\bar{P}\,\bar{P}'\varphi = \lambda^2\bar{\varphi}$$

so the $\lambda^2$ are the characteristic roots of $\bar{P}\,\bar{P}'$ *or* $\bar{P}'\,\bar{P}$, the $\{\bar{\varphi}_\ell\}_{\ell=1}^{\mu}$ are the orthonormal characteristic vectors of $\bar{P}\,\bar{P}'$, and the $\{Z_\ell\}_{\ell=1}^{\mu}$ are the orthonormal characteristic vectors of $\bar{P}'\,\bar{P}$ associated with the non-zero characteristic roots. Thus we obtain orthonormality just as in the Gorman model for prices. Observe that the $\lambda_\ell^2$ are real and non-negative since $\bar{P}\,\bar{P}'$ and $\bar{P}'\,\bar{P}$ are symmetric semipositive-definite matrices. The number of nonzero eigenvalues is the rank of the attributes or characteristics space. For distinct eigenvalues, the decomposition is unique.

(16) $$Pr(d_{ji}=1) = P_{ji} = \sum_{\ell=1}^{\mu} \lambda_\ell^2\, \bar{\varphi}_{j\ell} Z_{\ell i}\,.$$

We can absorb the $\lambda_\ell^2$ in (16) into the "preference parameters" $\bar{\varphi}_{j\ell}$ and define $\varphi_{j\ell} = \lambda_\ell^2\bar{\varphi}_{j\ell}$. Under Gorman-type orthonormality conventions, this version of the linear probability model is a *universal* representation of all models in the literature of discrete choice formulated in terms of unobserved attributes. The model is linear in the characteristics or attributes, $Z_{\ell i}$, $\ell = 1,\ldots,\mu$ and in the preference parameters $\varphi_{j\ell}$, $\ell = 1,\ldots,\mu$. In the population or in any sample, the constructed probabilities lie in the unit interval so the usual objection lodged against the linear probability model - that predicted probabilities may be outside the unit interval - does not apply in to this model provided a full-dimensional representation is used.

To gain some intuition about the relationship of this model to the standard linear probability model discussed in the preceding subsection, consider the special case of equation (13) where $A_j = 0$ for all $j$. For this case we can write (13) as

(17) $$P_{ji} = Pr(d_{ji}=1\,|\,X_{1i},X_{0i},\alpha_j) = a_j'(\Delta X_i) = \sum_{\ell=1}^{K} a_{j\ell}\,\Delta_{\ell i}$$

where $\Delta X_i = X_{1i} - X_{0i}$, and where we absorb $M_j$ into $a_j$ (an innocuous normalization in this case).

Representations (16) and (17) look very similar. If, in fact, the elements of $\Delta X_i$ are *mutually* orthogonal there would be an equivalence between $\Delta X_i$ in (17) and $Z_i$ in (16). However, nothing in choice theory justifies an assumption of orthogonality among attributes. In fact, it is more plausible that observed and unobserved attributes would be correlated. For example, in the context of a voting model, if the legislators who initiate bills attempt to trade off perceived losses on one component of a bill with gains on another to cater to diverse preferences, the attributes would be correlated.

The eigenvectors and eigenvalues that emerge from equation system (15) correspond to

(18a) $$\varphi_j = C_J a_j$$

and

(18b) $$Z_i = C_N(\Delta X_i)$$

where $C_J$ denotes a $J \times J$ orthogonal matrix, and $C_N$ is an $N \times N$ orthogonal matrix. Thus applying the singular value decomposition theorem to the model of equation (13) produces parameters that are scrambled. Fundamental identifying information is lost when one cannot observe the characteristics on which choices are based. This is precisely the problem discussed in identifying characteristics and their prices in the Gorman model.

Our latent-structure model can be viewed as one of many possible exact representations of the choice probabilities. It is a unique representation given the assumption of orthonormality of characteristics. Lower-dimensional representations with a rank less than $\mu$ have an optimal approximating feature as described in the following theorem.

*Theorem 1. Let the $(r,s)^{th}$ element of $\bar{P}\bar{P}'$ be $\delta_{rs}$ and let the $(p,q)^{th}$ element of $\bar{P}'\bar{P}$ be $\nu_{pq}$. Form $\bar{P}_M$, the M dimensional approximation from (16). Let the $(r,s)^{th}$ element of $\bar{P}_M\bar{P}_M'$ be $\hat{\delta}_{rs}$, and let the $(p,q)^{th}$ element of $\bar{P}_M'\bar{P}_M$ be $\hat{\nu}_{pq}$. Let "tr" denote the trace of a matrix. Then the $\bar{P}_M$ approximation based on the eigenvectors associated with the M largest eigenvalues, all assumed to be distinct, simultaneously minimizes*

$$\sum_{r=1}^{J} \sum_{s=1}^{J} (\delta_{rs} - \hat{\delta}_{rs})^2 = tr(\bar{P}\bar{P}' - P_M\bar{P}_M')$$

*and*

*among all approximating families of type (16) with dimension M.* ■ (See, *e.g.* Theil, 1971). ■

$$\sum_{p=1}^{N} \sum_{q=1}^{N} (\nu_{pq} - \hat{\nu}_{qp})^2 = tr(\bar{P}'\bar{P} - \bar{P}_M'\bar{P}_M)$$

Using only a subset of M factors, and recalling the definition that $\varphi_{j\ell} = \lambda^2\bar{\varphi}_{j\ell}$, we obtain

(19)
$$P_{ji} = \sum_{\ell=1}^{M} \varphi_{j\ell}Z_{\ell i} + \{\sum_{\ell=M+1}^{\mu} \varphi_{j\ell}Z_{\ell i}\}$$

where the term in braces is the "error term" *constructed* to be orthogonal to the "regressors." This theorem provides a precise way to characterize how a low-dimensional model relates to a higher dimensional model as an approximation. The factors for the lower-dimensional model are invariant to the inclusion of the factors for the higher order model. It is the best-fitting *M*-dimensional model. Observe that equations (13) and (19) are of the same form. The best one-dimensional approximation is based on the largest eigenvalue and the associated eigenvectors.

If the largest eigenvalue accounts for the lion's share of the variance (as measured by the distances in Theorem 1), what has been learned about equation (13)? Certainly *not* that only one "real" variable (*e.g.* one component of $\Delta X_i$) explains choice behavior, except in the unlikely case that the components of $\Delta X_i$ are mutually uncorrelated. By equations (18a) and (18b), the leading latent factor so isolated is a linear combination of the original variables. Evidence of a "one factor" model is consistent with the notion that components of $\Delta X_i$ are highly intercorrelated. The same analysis applies to the more general model where $A_j \neq 0$ in (13).

(c) *A More General Formulation With Unobserved and Observed Characteristics*

Any probability model formulated in terms of unobserved characteristics and preferences can be written as

(20a)
$$d_{ji} = P_{ji} + \zeta_{ji}, \quad j = 1,\ldots,J;\ i = 1,\ldots,N$$

where

(20b)
$$E(\zeta_{ji} \mid P_{ji}) = 0, \quad E(\zeta_{ji}\, \zeta_{j'\ell} \mid P_{ji}, P_{j'\ell}) = 0 \text{ all } j,j',\ i,\ \ell$$

(20c)
$$Var(\zeta_{ji} \mid P_{ji}) = P_{ji}(1 - P_{ji})$$

where

(20d)
$$P_{ji} = (\varphi_j)'\, Z_i.$$

so the variance for each $j$ is a stable function of the $Z_i$ (it depends on $i$ only through its dependence on $Z_i$).

Let $D_i$ record the choices on $i$ for all decision makers. $D_i = (d_{1i}, d_{2i},...,d_{Ji})$ where $d_{ji}$ is the choice of decision maker $j$ on issue $i$. $P_i = (P_{1i},...,P_{Ji})$, $\Phi = (\varphi_1,...,\varphi_J)$ where $\varphi_j$ is the $1 \times K$ ($K \leq N$) vector of preference parameters for legislator $j$, $Z = (Z_1,...,Z_N)$ where $Z_i$ is the $K \times 1$ vector of characteristics of bill $i$. Finally let $\zeta_i = (\zeta_{1i},...,\zeta_{Ji})$ be the choice $i$ - specific shocks of the $J$ decision makers. In this notation

$$D_i = P_i + \zeta_i$$

or in factor structure form

(20e) $$D_i = \delta + \Phi Z_i + \zeta_i, \qquad i = 1,...,N,$$

where $\delta$ is a $1 \times J$ vector of intercepts for each decision maker.

If some characteristics are observed, we can redefine the model and denote by $X_i$ by the observed characteristics and $Z_i$ the unobserved characteristics and write

$$P_{ji} = (\beta_j)'X_i + \varphi_j' Z_i$$

where $Z_i$ corresponds to the principal components of the population residuals from $\beta_j' X_i$

(21) $$P_{ji} - (\beta_j)'X_i = (\varphi_j)'Z_i.$$

The $Z_i$ now have the interpretation as orthonormalized population residuals of the unobserved characteristics not predicted by $X_i$. With this modification, we can accommodate both observed characteristics factor-structure components. In vector form, let $X_{1i}$ be the ($B \times 1$) vector of $B$ observed characteristics of bill $i$, and let $\beta = (\beta_1,...,\beta_J)$, where $\beta_\ell$ is a $1 \times B$, then letting $\Phi = (\varphi_1,...,\varphi_J)$, we may write

(22) $$D_i = \delta + \beta X_i + \Phi Z_i + \zeta_i$$

where $Z_i$ is a vector of unobserved characteristics.

Observe that even though the $\zeta_{ji}$ are uncorrelated across issues $i$ and voters $j$, the components in $\Phi Z_i$ are not. This generates stochastic dependence across votes for the same voter and may also create dependence across voters who face the same attributes. The linear probability model analyzed in this paper is richer than the conventional linear probability models analyzed by Goldberger (1964), Goldfeld and Quandt (1972) and Amemiya (1975, 1983) that assume independence across choice outcomes. The methods discussed in the next subsection can be used to estimate the factor structure of the error terms in a linear probability model with dependence among outcomes and to determine the rank of the factor matrix generating the dependence in population residuals.

(e) *Inference and Choosing The Rank of The Model*

Drawing on the important analyses of Cragg and Donald (1995, 1996), it is possible to rigorously justify the estimation of $\Phi$ when $N \to \infty$, and to produce a rigorous statistical criterion for estimating the rank of the model - *i.e.* the number of unobserved attributes required to rationalize the voting data. For simplicity, assume that there are no observed factors in the model ($\beta = 0$). If not, replace $D_i$ with $D_i - \beta X_i$ and proceed conditionally on the $X$. Inference is based on the covariance matrix of $D_i$, or $D_i - \beta X_i$, corresponding to whether or not we use raw data or residuals.

The Appendix presents formal conditions under which it is possible to estimate $\Phi$ consistently and produce a rigorous asymptotic distribution theory for the estimator. Standard results on distribution theory in factor analysis do not apply because of the heteroscedastic error terms $\varepsilon_i$ that is central to our model specification. Most of the existing literature in factor analysis assumes normality of the parent population which does not apply in our context. This affects the formula used to estimate population covariances of certain important sample covariance matrices. It is possible to use the minimum distance methods of Ferguson (1958) to estimate the preference parameters using the sample covariance matrix of $D$

$$\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^{N} (D_i - \bar{D})(D_i - \bar{D})' \tag{23}$$

where $\bar{D}$ is the sample mean of $D_i$. This covariance matrix is formed by averaging the votes of congressmen over bills. The analysis presented in the Appendix extends conventional factor analysis to inference on data from non-normal populations. See Theorem A-1 in the appendix.

Preference parameters can be consistently estimated when the $Z_i$ are independently distributed across choices, or when more general dependence across successive $Z_i$ is permitted.

In the context of a voting model, it is possible that the $Z_i$ are stochastically dependent as a consequence of log rolling and coalition formation. In the context of a consumer choice model, some brands may be close to other brands as a consequence of strategic behavior by rivals. Unlike the traditional random effects model, in our framework there is no need to assume a specific distributional assumption for unobserved characteristics in order to estimate the preferences of decision makers.

Cragg and Donald (1996) present conditions under which it is possible to rigorously estimate the rank $R$ of the model *i.e.* the number of attributes required to rationalize the covariance matrix for $D$, that is the rank ($R$) of $\Phi\Phi'$. Using a pretest estimator described in detail in the Appendix, it is possible to estimate the rank R. Cragg and Donald (1996) also produce a strongly-consistent model selection criterion based on the Bayesian Information Criterion (BIC). Their Monte Carlo analysis reveals that the pretest estimator works well, which is why we use it in this paper.

These results provide a rigorous foundation for the estimation of $\Phi$ and the dimension of the model. Because the number of choices ($N$) is large relative to the number of decision makers ($J$), no comparable results are possible for the attributes of any particular choice since $J/N \rightarrow 0$. The methods of Cragg and Donald can also be applied to the estimation of the prices and characteristics of the Gorman model when prices are not normally distributed.

## III. *Comparison With Other Estimators*

In our model, because of an "incidental parameters problem", only a subset of the relevant parameters can be consistently estimated in the case where the number of choices ($N$) becomes large relative to the number of decision makers ($J$). We obtain consistent and asymptotically normally distributed estimators of preference parameters because we are able to eliminate the incidental parameters from a system of estimating equations for the model by averaging over choices. Our estimator is a member of a general class of estimators suggested by Neyman and Scott (1948) that eliminates incidental parameters from estimating equations. Although we can use sample analogs of equation (15c) to estimate the (normalized) attributes of individual choices, such estimates are inconsistent.

In our model, the inconsistency in estimating $Z$ is innocuous. A consistent estimator of

$\Phi$ exists that does not depend on deriving an estimate of $Z$. In a more general nonlinear discrete choice model

$$Pr(d_{ji} = 1) = G(\varphi_j' Z_i), \qquad j=1,\ldots,J, i=1,\ldots,N$$

joint estimation of $\varphi_j$ and $Z_i$ poses more difficult problems when $J/N \to 0$. The problem, as first noted by Neyman and Scott (1948), is that the estimator of $\varphi_j$ depends in part on estimates of $Z_i$. Since $Z_i$ is not consistently estimated, neither is $\varphi_j$.

The classical example of such inconsistency is due to Andersen (1973, 1980) for the Raasch model in psychometrics. In that model, the probability of choosing "1" on choice $i$ by decision maker $j$ is

$$\Pr(d_{ji} = 1) = P_{ji} = \frac{\exp(\theta_i - \alpha_j)}{1 + \exp(\theta_i - \alpha_j)}$$

where $\theta_i$ is a choice-specific attribute and $\alpha_j$ is a decision-maker preference parameter. In the context of voting models, $\alpha_j$ is an ideal point. Assume that $(N/J) \to \infty$ so the number of decision makers is small relative to the number of decisions taken. Then if we estimate $(\theta_i)_{i=1}^{N}, (\alpha_i)_{j=1}^{J}$ by maximum likelihood, we obtain inconsistent estimators of the $\theta_i$ and the $\alpha_j$. For the case $J = 2$ and $N \to \infty$, plim $\hat{\alpha}_j = 2\alpha_j^0$, $j = 1,2$, where $\alpha_j^0$ is the true parameter. In the general case for any $J \geq 2$, we obtain

$$plim\, \hat{\alpha}_j = \left(\frac{J}{J-1}\right) \alpha_j^0, \qquad j=1,\ldots,J.$$

Our method avoids this type of inconsistency by solving out the $\alpha_j$ without using the $Z_i$ (the $\theta_i$ in our example) to construct the estimator.[10]

The estimator of Poole and Rosenthal is based on the highly nonlinear probability model (9). It is not the same as the Raasch model so the specifics of the preceding example do not directly apply to it, although the basic methodological point does. Poole and Rosenthal simultaneously estimate $(X_{1i}, X_{0i})$, i = 1,...,$N$, $\beta$ and $(\alpha_j)$, $j$ = 1,...,$J$. The inconsistency in estimating the $(X_{1i}, X_{0i})$ is transmitted to their estimates of preference parameters $L_{\alpha j}$. Because of the inconsistency of their estimators, it is not possible to perform valid statistical inference in their model. In particular, it is not possible to rigorously estimate the rank of their model, although rank plays a key role in their interpretation of the voting data.[11]

Brady (1989) was the first to point out the inconsistency in the Poole-Rosenthal estimator. Drawing on a basic paper by Kiefer and Wolfowitz (1956), Brady avoids the inconsistency problem by estimating $\alpha_j$ and the distribution of $Z$. By estimating the distribution of Z, and not any particular realization from it, it is possible to consistently estimate $\alpha_j$ up to some normalizations.

Integrating out $Z$, using distribution $F(Z \mid \sigma)$, where $\sigma$ is a finite-dimensional parameter vector, Brady bases his inference on the average probability over choices:

$$Pr(d_j = 1 \mid \varphi_j) = \int G(\varphi_j' Z) dF(Z \mid \sigma)$$

where he assumes knowledge of the functional form of $F$ up to an unknown finite-dimensional parameter vector $\sigma$ as well as knowledge of $G$. Under his stated conditions, his method consistently and efficiently estimates $\varphi_j$, $j = 1,\ldots,J$ and $\sigma$.[12]

Our method has several advantages over Brady's. (1) It is much simpler to compute. Brady reports numerical difficulties in obtaining estimates of preferences from his model. (2) Brady is forced to assume that the $Z_i$ are independent and identically distributed. Our method permits strong dependence across the $Z_i$. (3) Brady makes assumptions about the functional form of the distribution of $Z$ ($F(Z \mid \sigma)$). Our method requires no such assumptions. While in principle $F(Z \mid \sigma)$ can be nonparametrically estimated, so that a distributional assumption is not central to the application of this method, in practice nonparametric estimators of mixing distributions like $F(Z \mid \sigma)$ perform poorly when the dimension of the model is large.

As previously noted, our method provides an exact representation of *any* choice probability defined over unobserved preferences and characteristics. This representation is not unique. It is, however, computationally convenient and can be rigorously statistically justified. Our estimator, like Brady's, exploits the key idea in Kiefer and Wolfowitz and averages out over the $Z_i$ to solve the inconsistency problem. We average out the $Z_i$ by averaging over the $D_i$. This approach is computationally simpler and robust under more general conditions. Brady averages over population probabilities defined conditionally on $Z_i$. This is a computationally more difficult task that requires invocation of much stronger assumptions than the ones we use. Given that the attributes are not observed, and that basic parameters are identified by convention, the simpler approach seems the preferable one.

### *IV. An Application: Roll Call Voting in the U.S. Congress, 1947-1988*

In this section we apply the model developed above to roll call voting in the U.S. Senate and House of Representatives over the 80th-100th Congresses (1947-1988). For each roll call, the choice facing each senator or represenstative is to vote "yea" or "nay".[13] We estimate the congressmen's preference parameters, or ideal points (the $L_\alpha$'s). We also estimate the characteristics of the roll-calls (the $X_{0i}$'s and $X_{1i}$'s) but under our assumptions, these estimators are inconsistent. The number of roll calls ranges from 184 to 1144 for the Senate, and from 120 to 1217 in the House. On average, the number of roll calls in a congress is 572 for the Senate and 494 for the House. In each congress we drop a few senators and representatives from the analysis who served for only part of the time and voted on a relatively small number of roll calls. On average, there are 99 senators and 431 representatives in our samples.

Four main conclusions emerge from our research. First, assuming that the preference parameter space or attribute space has one or two dimensions, estimated preference parameters of our linear model are virtually identical to those obtained from Poole and Rosenthal's "NOMINATE" model which jointly estimates the characteristics ($X_{0i}$, $X_{1i}$) and ideal points $L_{\alpha j}$ of equation (9) as well as the $\beta$ that is part of the definition of $\rho$. Given the enormous computational savings entailed in estimating the linear model, this finding is of practical value even for those who prefer the nonlinear model. This evidence also suggests that the statistical inconsistency that plagues their estimator is of little consequence provided the dimension of the model is specified in advance of estimating the practical model. Second, *ad hoc* application of different goodness-of-fit measures produces different inferences about the dimension, $R$, of the model. For example, while a one-dimensional model sometimes appears adequate when evaluated in terms of classification success, (predicting the frequency of yea or nay votes), adopting a proportional reduction in error criterion usually indicates that higher dimensions are important. When we apply the formal testing procedure described in section IIe and in the Appendix, we overwhelmingly reject one- and two-dimensional models in favor of models with at least five effective dimensions

($R \geq 5$). This evidence is bolstered by evidence that estimated preferences over these higher dimensions are stable over time for the same congressmen. Third, we find that the first two estimated "factors" essentially are stand-ins for party and region. A model with two *observed* factors predicts as well as a model with two unobserved factors. Estimated "ideology factors" have prosaic names after all, and their empirical importance simply indicates the political importance of party and regional affiliations. Fourth, a multi-dimensional model is definitely required to adequately predict votes on certain issues, such as abortion, agricultural programs, defense spending and foreign aid. In the next two subsections we report estimates from a two-factor model. In this, we follow much of the congressional roll call voting literature, which argues that roll call voting is largely "ideological" and therefore one or perhaps two dimensional (*e.g.*, Kritzer 1978, Poole 1981, 1988, Poole and Daniels 1985, Poole and Rosenthal 1985b, 1987, 1991a).[14] In subsection (c) we analyze models with more factors, and also rigorously test for dimensionality. In subsection (d) we illustrate the value of looking at higher dimensions by studying the voting of committees and congressmen on civil rights.

### *(a) Estimating the Linear Factor Model*

We first compare the empirical performance of different methods for estimating the preferences of legislators. Computationally, the simplest methods are principal components and principal factor analysis. Principal factor analysis, and unweighted least-squares using the sub-diagonal terms of the covariance matrix, are also computationally straightforward. The generalized least squares approach of Cragg and Donald (1995) is the most appealing in principle, but the most difficult in practice. The problem is that the weighting matrix $\Omega^{-1}$, the inverse of the covariance matrix of $\hat{\Sigma}$ defined in (23), is very large. This covariance matrix is required to estimate preference parameters using the minimum distance method. The estimates for the elements of $\Omega$ involve the sample fourth moments of the original roll call data matrix. With 100 senators, the covariance matrix of the roll call data has $(100)(99)/2 = 4950$ sub-diagonal elements, so $\Omega$ is 4950 by 4950. For the House of Representatives, $\Omega$ is even larger, 94395 by 94395. Moreover, unless there are an enormous number of roll calls, $\Omega$ will be singular or nearly so, and the estimates of its elements will be imprecise. Thus, in practice we can only obtain weighted least-squares estimates for small subsets of senators or representatives. We use the subdiagonal estimator of Cragg and Donald (1995) that is discussed below Theorem A-1 in the Appendix.

The correlations between the principal factor estimates[1] and the estimates obtained from other methods are shown in tables 1a and 1b. Table 1a presents results for the Senate, and table 1b presents results for the House. For principal components analysis, and unweighted least squares applied to the sub-diagonal covariance terms, these correlations are all virtually indistinguishable from unity to three decimal places. The correlations between principal factor estimates and the generalized least squares estimates are lower. This is partly due to the fact that the generalized least squares model could only be estimated for subsets of legislators, but it probably also reflects the difficulty of estimating the covariance of $\Sigma$. Overall, the choice of estimation method does not make much difference, so it seems best to choose one that is computationally inexpensive. We use the principal factor method to secure estimates for the results reported in this section and the next. This estimator is consistent but not efficient. We base our statistics for estimating the dimension of the model on the theoretically appropriate minimum distance estimates as required by Theorem A-2 in the Appendix.

Figure 1 presents plots of the estimated preference parameters $\Phi$, for a few Houses, and Figure 2 presents plots for a few Senates. The symbols are "D" for northern Democrats, "S" for southern Democrats, "R" for northern Republicans, "T" for southern Republicans, and "I" for independents. Several patterns are worth noting. Most obvious, perhaps, is the large gap between the two parties in both chambers, especially in the 85th, 90th and 100th Congresses. The degree of partisan division is not constant, however. In the 95th Congress (1977-78), there were considerably more "centrists" in both chambers. This is true for the 93rd-96th Congresses as well, but by the 100th Congress (1987-88) the center was again largely vacant, especially in the Senate. Second, in most years, the main "axis of polarization" between the parties is not the same as the dimension defined by the largest eigenvalue (factor 1), but a combination of the first two factors. In the 100th House, however, the main axis of partisan conflict was fairly well defined by the first factor. Finally, in most Congresses there is a clear distinction between northern and southern Democrats, and northern Democrats generally appear to have been considerably more cohesive than southerners. Interestingly, a similar pattern appears among Republicans, as the number of

[1]Principal factor analysis imposes orthonormality conditions on the estimated factors.

southern Republicans grows. While the division is not as clean as it is in the Democratic party, by the 100th Congress the southern Republicans are all in the right wing of their party. Similar patterns have been identified previously by various political scientists using other methods. These findings illustrate the "reasonableness" of estimates obtained from our model.

*(b) A Comparison of the Linear Model and the "NOMINATE" Model*

In this section we compare the preference parameter estimates obtained from the linear model with those from "NOMINATE".[15] We examine "NOMINATE" because of its prominence in the empirical roll call voting literature. Poole and Rosenthal have used the model to study a variety of important questions about congressional roll call voting (1985a, 1985b, 1987, 1990, 1991a, 1991b), and an increasing number of other researchers are using "NOMINATE" estimates as inputs in their work (*e.g.*, Kiewiet and McCubbins 1991, Cox and McCubbins 1993).

As discussed above, maximum likelihood estimation of the "NOMINATE" model does not produce consistent estimates of either congressmen's preference parameters or roll call parameters. Nonetheless, we find that the "NOMINATE" estimates of the two dimensional preference parameters are quite similar to those obtained by consistent estimation of our linear model. This suggests that the inconsistency that plagues the Poole-Rosenthal estimator may not be a practically important problem when applied to large political bodies like both houses of the U.S. Congress. For smaller bodies like the U.S. Supreme Court, application of their estimator is likely to be more problematic.

Tables 2a and 2b present summaries of the results for the two dimensional models. In each table, columns 2 and 3 contain the Pearson correlations between the preference parameters using the linear model and the estimates obtained from "NOMINATE".[2] For the first dimension the average correlation for 21 congresses is over .98 in the Senate and is nearly .97 in the House. The Spearman rank-order correlations (columns 4 and 5) are even higher, ranging from .98 to 1.00 for the Senate and .96 to 1.00 in the House. For the second dimension, the correlations are only slightly lower. In the Senate, the average Pearson and rank-order correlations are both greater than .97, and in the House they are both greater than .96. Clearly, the two models produce

[2]"NOMINATE" estimates were computed using Poole and Rosenthal's W-NOMINATE FORTRAN program.

strikingly similar results.

Column 8 gives the percentage of roll-call votes correctly classified, averaged across representatives, using our two-factor linear model. Column 9 gives the classification success rate for the two-dimensional "NOMINATE" model.[16] The differences between the models are again negligible - both yield classification success rates between 83 and 88 percent.

The results are also important for practical reasons. Poole and Rosenthal required a supercomputer to calculate their "D-NOMINATE" estimates for the whole House for its entire history, and even on smaller data sets "NOMINATE" and "D-NOMINATE" are computationally expensive. On computational grounds, the linear model is clearly preferred. Also, since most commonly-used statistical packages implement factor analysis and principal components analysis, researchers who desire specialized roll call voting indices and have access to roll call data can construct such indices quickly and easily using a linear model. Computational speed is also important for estimating standard errors using bootstrap methods, and for conducting Monte Carlo studies to explore the small-sample properties of estimators.

*(c) Estimating the Rank of The Model and Interpreting The Evidence From a High Dimensional Model*

In this section, we use the sub-diagonal minimum distance estimators described in the Appendix to estimate the number of factors underlying congressmen's revealed roll call preferences. We find that in all of the post-war congresses, there are at least five statistically-significant factors, and that these factors are also important substantively. These results contrast sharply with the conclusions of scholars who claim that congressional roll call voting is largely one-dimensional (*e.g.*, Schneider 1979, Poole 1981, 1988, Poole and Daniels 1985, and Poole and Rosenthal 1985b, 1987, 1991a). Poole and Rosenthal (1991a) in particular argue that congressional roll call voting can be described by a model consisting of one or two spatial dimensions plus "noise". They estimated models with many dimensions and found that classification success rates increased only slightly with extra dimensions. The following quote illustrates their general conclusion: "The 97th House is at most two dimensional with the second dimension being very weak. After two dimensions the added classifications are minuscule. There is a clear pattern of noise fitting beyond two dimensions" (p. 243). Our results challenge this interpretation.

Tables 3a and 3b begin the analysis with a closer look at two commonly used goodness-of-fit measures: classification success and proportional reduction in error. As Weisberg (1978), Koford (1989) and others have pointed out, in most postwar congresses about 67-70 percent of all roll call votes are cast with the majority side[17], so even extremely naive models such as "everyone votes with the majority" will correctly classify at least 67-70 percent of the votes. Classification success rates of 80 percent are not so remarkable relative to this benchmark, and classification success may be therefore be a poor guide for choosing the dimension of a model. Focusing on "close" votes provides a somewhat tougher test. Columns 2 and 3 of tables 3a and 3b show the classification success rates on close roll calls -- where close votes are defined as those in which 60 percent or fewer congressmen voted with the majority -- for one-factor and six-factor models. (The six factor model was not chosen arbitrarily, as will be clear shortly.) Averaging across all congresses, the one-factor model achieves a classification success rate of 79 percent for the Senate and 81 percent for the House, even on the close votes. This is respectable, since the average size of the majority on these roll calls is only about 55 percent. For the six-factor model, the average classification success rate on close votes is 88 percent in both chambers. In the Senate this represents a nontrivial improvement of 9 percentage points over the one-factor model.

By way of comparison, columns 5-7 present the same results for the "lopsided" votes. Not surprisingly, in all but a few Senates and Houses the classification success rate is higher on lopsided votes than on the close votes, for both the one-factor and six-factor models. More interestingly, however, the "value added" of factors 2-6 is almost always greater on the close votes than on the lopsided votes. For the Senate the difference is substantial -- on close votes the extra factors increase classification success by 50 percent more than on lopsided votes.

An alternative goodness-of-fit measure is the proportional reduction in error (PRE), which measures classification success rates relative to the naive baseline model "everyone votes with the majority" (*e.g.*, Weisberg 1978, Poole and Rosenthal 1991a). On each roll call, PRE = 1 - errors/(number of votes on minority side). The individual PRE's are then averaged across all the roll calls in a congress. Average PRE's for the one-factor and six-factor models are shown in columns 8-11 of tables 3a and 3b. The one-factor model now appears much less impressive, explaining on average only about 35 percent of the variation unaccounted for by the naive model. Also, the "value added" of additional factors now appears much more substantial. In the Senate the six-factor model accounts for 57 percent of the variation unexplained by the naive model, while in the House the figure is 54 percent. The additional factors have an especially large impact on PRE in the Senate, both in absolute and relative terms. In all but three of the 21 Senates, the additional factors increase the PRE by more than 50 percent, and the average increase is 68 percent.

Tables 4a and 4b present results interpreting the unobserved factors that are estimated by our procedure. In place of a model cast only in terms of unobserved factors, we estimate a model with observed attributes using dummy variables for southern Democrats, northern Democrats, and Republicans. This model does nearly as well as the one-dimensional factor model, and sometimes performs better, both in terms of classification success and PRE (compare columns 2 and 3 of tables 4a and 4b with columns 2, 5 and 8 of tables 3a and 3b). Column 4 shows why. In it, we report the $R^2$ from regressions of the factors estimated from a model without any regressors, on dummy variables for party and region. The first preference factor is highly correlated with the region and party variables, especially for the House. The second factor also exhibits a high correlation with those variables. The column labelled "Average 3-6" shows the low *average* $R^2$ for the next four factors. These factors are not closely related to party or region.

While intuitively appealing, neither classification success nor proportional reduction in error are rigorous statistical criteria for choosing dimensionality. We now exploit Theorem A-2 in the Appendix to estimate the dimension of the model. Using the minimum distance estimates, we produce the sequential chi-square tests shown in tables 5a and 5b. As previously noted, it is impossible to accurately estimate the inverse of the covariance matrix of $\Sigma$ as required by Theorem A-2 using all 100 Senators or 435 representatives. We estimated the model using randomly drawn subsets of size 20. The results in the tables can be shown to provide lower bounds on the true number of dimensions.

Looking at the 87th Senate, for example, table 5a shows that there are at least six factors that represent something other than senator-specific "noise". The hypothesis that the number of factors is five or less is strongly rejected. The results are similar for other Senates and also for the House, as shown in table 5b. The tests overwhelmingly reject the hypothesis that the number of factors is one or two.

The Akaike Information Criterion (AIC) model selection rule discussed by Cragg and Donald (1996) tends to estimate an even larger number of factors than the sequential chi-square test - as many as seven or eight. These tables are deleted for the sake of brevity. On the other hand, the Bayesian Information Criterion (BIC) discussed by Cragg and Donald tends to estimate fewer factors, usually three or four. These differences are not too surprising in light of Monte Carlo evidence by Cragg and Donald who show that the BIC criterion often tends to underestimate the number of factors in finite samples. In any case, using any of the three procedures, we conclusively reject the hypothesis that the number of factors is less than three.

An informal method for measuring dimensionality is to examine the stability of estimated preferences (factor loadings) over time. Assuming that congressmen's preferences are stable, factors that measure preference parameters should be highly correlated across congresses, while factors that simply pick up "noise" should be transitory. Tables 6a and 6b present the correlations and multiple correlations between the estimated ideal point vectors of pairs of adjacent Houses and Senates. Clearly there is a striking degree of intertemporal stability. Even more important, this is true for more than just the first two factors. For example, in all but two Senates since the 88th, and all but one House, there are at least six factors with multiple correlations of .6 or higher across pairs of congresses. For more recent Houses there are eight such factors. Evidently, there are more than two factors and they capture something systematic and not just idiosyncratic random error.

It is possible that some of these factors are not associated with distinct issue dimensions, but instead pick up parameters such as those that reflect differences in the "intensity" of preferences across individuals and issues. In most cases, however, the additional factors seem to reflect special issue dimensions. These are dimensions which arise on only a few roll calls in most sessions, and therefore are easily missed when classification success is used to decide how many factors to retain. We can get a feel for the substantive issues associated with various factors by studying the pattern of factor scores. In many cases the roll calls with especially high factor scores on a given factor all deal with a particular issue or a small set of issues.[3]

Table 7 shows the roll calls with high factor scores on each dimension for the 90th House. The table shows clearly that factors 2, 3, 4, 5 and 7 capture dimensions of congressmen's preferences that are tapped by particular subsets of issues.

Lack of space prohibits us from presenting detailed tables for all of the postwar congresses. However, in table 8 we present a summary. The table reveals that certain key issues are consistently reflected in factors other than the first factor. During the 1950s and 1960s, these issues were civil rights and voting rights, agriculture, foreign aid, trade policy, and criminal law (rights of the accused, wiretapping, and so on). Note that foreign aid and trade policy are associated with the third or fourth factor. During the 1970s and 1980s, the issues with high scores on factors other than the first were agriculture, foreign aid, military spending, the debt ceiling, water projects, abortion, and congressional reform. In a large majority of cases the roll calls with high scores on one of the higher dimensions also have high scores on the first dimension (*e.g.*, see table 7). Thus, the patterns indicate that for many issues voting is a function of *both* a dominant, "ideological" or party component and an issue-specific component.

Table 9 briefly summarizes the predictive power of one-, two- and six-factor models on several of the issues shown in table 8. Adding factors 2-6 more than doubles the PRE on a number of issues, such as abortion for 93rd-95th congresses, agriculture for the 90th-100th congresses, civil rights for the 80th-90th congresses, defense spending for the 80th and 87th-97th congresses, foreign aid for the 80th-87th congresses, and water projects for the 90th-100th congresses. The table also shows that on issues such as abortion, civil rights, and foreign aid (but not agriculture and water projects), congressmen's revealed preferences became more correlated with the "main" dimension over time.

Overall, these results clearly demonstrate that roll call analysts should not focus their attention exclusively on *ad hoc* goodness of fit measures which have little rigorous justification. Doing so may cause analysts to overlook preference factors that are tapped by relatively few roll calls. While such factors may not add much to overall classification success, they are crucial for

[3]Note that this is even true of the *non*-rotated factors. Most factor-analytic studies rotate the factors based on some criterion for trying to produce "cleaner" interpretations. Because of disagreements over the pros and cons of various techniques for rotating factors, we examine the non-rotated factors.

accurate prediction on certain specific issues.

The results also allow some tentative conclusions about congressional decision making. Previous researchers have identified a variety of factors that might influence congressmen's voting decisions, including personal ideology, policy preferences, constituency preferences, pressure from party leaders or interest groups, and coalition building. Our results are inconsistent with "pressure group" models that do not allow for logrolling or coalition building among the groups (*e.g.*, Stigler 1971, Peltzman 1976, Denzau and Munger 1986). If each special interest lobbies alone, and each is identified with a distinctive attribute or characteristics, we should expect to find nearly as many voting dimensions as there are interest groups, and each dimension should isolate the roll calls of interest to a particular group. Instead, we find only a handful of significant dimensions, and most of these dimensions are tapped by legislation in several unrelated or loosely-related policy areas. Our results are also inconsistent with the ideological model, which asserts that on the vast majority of roll calls congressmen look to their own ideologies in deciding how to vote (Schneider 1979, Poole 1981, 1988, Poole and Daniels 1985). Drawing on Converse (1964), proponents of the ideological model argue that the existence of a stable, low-dimensional roll call voting space provides strong support for their view. However, we find that in virtually all of the post-war congresses, there are at least five and as many as eight statistically and substantively significant dimensions. While it is not clear what number of dimensions is low enough to qualify as supporting a model of pure ideology, five would seem to be too many. This also implies that the median voter model, which assumes a one-dimensional issue space, is an inappropriate framework for analyzing congressional voting (e.g., Riker 1990).

Our results suggest that some roll call voting is driven by constituency concerns and log-rolling. Table 8 shows that during the 96th-100th Congresses roll calls on defense spending, agricultural subsidies and rural development all had high scores on the second factor. It is also true that the preferences on factor two in these Houses are significantly correlated with the fraction of employed persons in the military, the fraction employed in agriculture, and the fraction living in rural areas (we do not report tables of these results in the interest of space). These groups obviously have special interests in defense, agricultural and rural development policies. There is also a positive relationship between support for defense spending and support for agricultural programs. Since there is no obvious logical link between military spending and agricultural subsidies, these facts provide indirect evidence of the existence of certain logrolls. Farm groups and groups interested in defense spending may have found a common interest in supporting one another's programs. This evidence is consistent with Mayhew's (1966) argument about the "inclusive" nature of the Democratic coalition, and deserve further exploration.

Finally, we do not find any striking changes in the number of statistically significant factors over time. This is interesting because during the time period under study there were a variety of congressional reforms (*e.g.*, the re-emergence of the Democratic Caucus, changes in committee structure, and the introduction of electronic roll call voting), variation in presidential agendas and leadership "styles," and changes in electoral patterns (*e.g.*, the emergence of the Republican South). As shown in table 8, there was some change over time in the types of issues captured by certain factors, probably reflecting changing agendas and coalition-building strategies. However, the institutional and environmental changes appear to have had little effect on the overall dimensionality of the underlying space of bill characteristics.

*(d) A Case Study of Committee Resistance to Civil Rights Legislation, 1949-1966*

Through most of the 1950s and 1960s, the chairmen and southern members of the Senate Judiciary Committee and the House Rules Committee were staunch opponents of civil rights legislation, and used what powers they had to kill, delay, and water-down such legislation. We show that this opposition can only be understood by studying a multidimensional model of congressional preferences, paying special attention to the second dimension.

The Senate Judiciary Committee reported only one civil rights bill between 1949 and 1964, and only when it was under orders from the Senate to do so. The committee successfully kept House-passed civil rights legislation from the Senate floor in 1949 and 1956, and consistently bottled up civil rights bills reported by the Judiciary Constitutional Rights Subcommittee. Senate leaders began bypassing the Judiciary committee in 1957 by sending civil rights bills passed by the House directly to the Senate floor rather referring them to the committee, or by encouraging Senators to add civil rights legislation as amendments to other, minor bills. The Judiciary Committee finally reported a civil rights bill favorably in 1965, although it was again under orders from the Senate, and the chairman, James Eastland of Mississippi, again opposed the bill and delayed it as long as possible.

The House Rules Committee was equally unfriendly to civil rights legislation, blocking and delaying action on bills in 1950, 1956, and 1959. Proponents of civil rights were forced to use a variety of informal procedures including suspension of the rules, and discharge petitions to bring their bills to the floor. In 1961, the House enlarged the committee in an effort to open the gates on a wide range of legislation, including civil rights. Since the committee was chaired by staunch conservative Howard Smith of Virginia, it still remained a bottleneck, blocking legislation in 1962 and 1963 and reporting a rule in 1964 only when threatened with a discharge petition.

Standard spatial models of committee power imply that regular battles between a committee and its parent chamber should occur only when the ideal point of the pivotal member of the committee is substantially different from the ideal point of the pivotal member of the floor, on at least one issue dimension included in the committee's jurisdiction (*e.g.*, Shepsle 1979, Denzau and Mackey 1983). Since the period under study was an era of "strong" committee chairs, it is likely that the pivotal committee members were the chairmen. This is even more likely when considering a committee's power to block (rather than pass) legislation.

Figures 3 and 4 present two-dimensional plots of the estimated ideal points in the House and Senate over the period 1949-1966. Each figure also contain two pairs of lines, one showing the location of the median member of the chamber on each dimension, the other showing the location of the chairman of the relevant committee under study -- the Judiciary Committee in the Senate and the Rules Committee in the House.

If we were to assume that the unidimensional, ideologicial view of congressional preferences is correct and focus only on the first dimension in figures 4, we would conclude that up until 1959 the ideal point of the Senate Judiciary Committee chairman was usually quite close to the ideal point of the median senator. In fact, in the 84th and 85th Congresses the Judiciary chairman was slightly more *liberal* than the floor median. From 1959 to 1962 the Judiciary chairman was more conservative than the floor median, but not by much. For example, only about ten senators had "ideal points" or preference parameters between the chair and the floor. The picture is similar in the House. On the first dimension the chair of the Rules Committee was considerably more *liberal* than the floor median from 1949-1952, and only slightly more conservative than the floor during the periods 1955-1958 and 1961-1964 (the Republicans controlled congress from 1953-1954). Viewing civil rights bills as "liberal" legislation, it is hard to see why the chairmen of the Senate Judiciary and House Rules committees were such strident opponents of these bills.

When we acknowledge that congressional preferences are multidimensional, however, the interpretation changes dramatically. Examination of the factor scores shows that during the period under study, civil rights was primarily a second-dimension issue, especially during the first part of the period (see table 8). For the set of civil rights roll calls taken in the House during the 81st-89th Congresses, the average of the absolute value of the factor scores on the second factor is about 4.8. The average absolute score on the first factor for these roll calls is just 3.0, and the average scores on the other dimensions are all much smaller.

Figure 4 shows that between 1949 and 1966 the chair of the Senate Judiciary Committee was consistently much more "southern", or "conservative", than the floor median on the second dimension (higher values on the second dimension means more support for civil rights). In fact, except during the Republican controlled 83rd Congress, the Judiciary chairman was always one of the *most* "conservative" senators on the second dimension. Figure 4 shows a similar pattern for the chairman of the House Rules Committee. From the 84th Congress on, the chair of the Rules Committee was consistently among the most "conservative" ten percent of the House on the second dimension. It is therefore no mystery that the Senate Judiciary and House Rules committees were strong opponents of civil rights. The two-dimensional pictures are informative in a way that the one-dimensional pictures based on ideology are not.

## *V. Summary and Conclusions*

This paper formulates and estimates a computationally tractable, statistically rigorous model of binary discrete choice for the case where some or all of the attributes of the options available in each choice setting are not observed by the analyst. This framework extends Gorman's formulation of the demand for goods expressed in terms of their attributes to a discrete-choice setting. The linear probability model for discrete choice is rigorously justified in two conceptually distinct ways: (a) as an exact representation of a class of random utility models with asymmetric shocks imposed on the utilities of the options confronting decision makers in a particular choice setting and (b) as an exact representation of any discrete choice model formulated in terms of unobserved attributes. We stress the second justification in this paper because of its wide applicability and because of its relevance to the problem in political economy that we study.

This paper presents rigorous methods for recovering decision-maker preferences when attributes of choices are not observed, and considers conditions under which these preferences can be identified. The effective dimension or rank of the attribute space can be identified and consistently estimated under fairly weak conditions. Recovery of preference parameters requires invocation of much stronger identifying assumptions.

The statistical model developed here justifies the application of low-cost factor-analytic methods for the analysis of discrete choices. The framework can be used to estimate conventional linear probability models with factor-analytic error structures. Our framework is one representation of any model formulated in terms of unobserved attributes of choices.

The model can be applied to a variety of discrete-choice problems. In this paper, we use it to study congressional voting patterns. Our model vindicates the application of factor analysis to the study of voting. There is a long history in quantitative political science and political economy that uses factor analysis to analyze voting data. This literature was discarded in deference to the apparently more rigorous analysis of Poole and Rosenthal. Those authors formulate and estimate nonlinear probability models of voting choices over unobserved attributes that are motivated by formal choice theory as in Enelow and Hinich (1984). We demonstrate that the Poole-Rosenthal estimator is statistically inconsistent and cannot be used to rigorously determine the dimension - or rank - of the unobserved attribute space. This is unfortunate, since this dimension plays a fundamental role in their interpretation of the voting data. Using *ad hoc* methods, Poole and Rosenthal claim that one or possibly two dimensions explain congressional voting. A low-dimensional model is interpreted as arising from legislators voting on the basis of ideology and not on the substance of specific bills. A high-dimensional model arises if the interest group models of Stigler (1971) and Peltzman (1976) explain the congressional voting data, assuming that different groups tap different policy attributes relevant to Congressmen.

The main advantages of our approach over predecessor models are computational simplicity and the ability to rigorously estimate the dimension - or rank - of the attribute space. If the dimension of the model is specified to be the same, there is close agreement between our estimates of preference parameters and those obtained from the method of Poole and Rosenthal.

This evidence suggests that the statistical inconsistency problem, while theoretically important, may be practically unimportant for estimating preferences of legislators in legislative bodies as large as the U.S. Senate or the U.S. House of Representatives *provided that the effective dimension of the attribute space is prespecified.*

An analysis of the two dimensions estimated by Poole and Rosenthal reveals that they are factor-analytic stand-ins for the party and region of congressmen. Those two observed attributes predict voting about as well, and sometimes better, than a model formulated in terms of unobserved attributes. "Ideology" in their model turns out to be just party and regional loyalty.

Using a more rigorous estimator for the dimensionality or rank of the attribute space, we determine that at least five and as many as eight dimensions are required to rationalize the voting data. The additional dimensions improve the predictive power of the estimated model, especially on close votes. Our evidence speaks against ideology as the sole determinant of congressional voting and indicates important issue-specific features of voting that have been missed in previous analyses. Our evidence also casts doubt on the validity of the median voter model as a description of congressional voting patterns. That model requires one-or at most two dimensions for issues (Riker, 1990) but we estimate at least five. On the other hand, the number of effective dimensions that we estimate is still relatively small suggesting that special interest groups do not act independently of each other as is assumed in Stigler and Peltzman. Rather, they appear to collude in forming logrolls on votes.

We also use our model to investigate Senate and House voting patterns on civil rights. We show that conflicts between committee preferences and the preferences of the entire chamber can only be understood if multidimensional models of preferences are estimated.

# ENDNOTES

1. In their otherwise excellent survey of discrete choice theory, Anderson, de Palma and Thisse (1992, p. 33) inappropriately assign credit for these ideas to later, derivative, work.

2. In the Gorman model, prices reflect valuation of characteristics. See Pudney (1980) and Heckman and Scheinkman (1987).

3. Panel data on the same persons confronting different controlled choice environments could distinguish among the three sources of heterogeneity. Such data are common in experimental psychology (see *e.g.* McFadden 1974 and Falmage, 1985) but are rare in social science.

4. Clearly, we may decompose the Sin term recursively and produce a countable collection of alternative, non-degenerate representations for the model. Other decompositions are possible *e.g.* $\varepsilon_1 = U + \eta$; $\varepsilon_0 = \eta$, where $\eta$ is arbitrarily stochastically dependent on $U$ and $U$ is uniform.

5. The square root of the characteristic function of a uniform random variable does not satisfy Bochner's Theorem and hence cannot be a characteristic function. See, *e.g.* Pratt and Gibbons, 1981, p. 395 and problems 64 and 65, p. 420.

6. The rank conditions are not entirely trivial. If the $(X_{1i}, X_{0i})$ includes binary variables, then the diagonals of $(X_{1i}X_{1i}'-X_{0i}\ X_{i0}')$ and $(X_{1i}-X_0)$ are redundant. Although $(a_k+a_{kk})$, $k = 1,..,R$, are identified it would not be possible to separate out these components. further. Partial observability of $(X_{1i}, X_{0i})$ creates the usual problem of specification bias.

7. If the regressor matrix is of less than full rank, however, one cannot estimate all of the coefficients. One solution to the problem is to use the generalized inverse to compute all coefficients. Theil (1971) demonstrates that this is equivalent to forming principal components of the regressors and estimating data-defined linear combinations of the true parameters of interest.

8. Conditions for uniform pricing of characteristics of goods are presented in Heckman and Scheinkman (1987).

9. In a cross-section, the roles of $C_t$ and $W_t$ are reversed. $W_t$ becomes the factor and $C_t$ is the factor loading.

10. Andersen (1973, 1980) develops an estimator that conditions out $\theta_i$ to obtain a consistent estimator of $\alpha_j^0$.

11. Not only is their estimator inconsistent, it is also very costly to compute. One reason for the enormous computational cost is that the likelihood is not concave in the parameters. Thus if $G$ is a logit, $G(\kappa(Z'\beta))$ is concave in $\beta$ or $Z$ if $\kappa$ is the identity function. (Amemiya, 1985). For the $\kappa = \rho$, defined in the text surrounding the discussion of equation (9), the log-likelihood is not in general concave in the parameters. Note, however that their choice of $\rho$ enables them to separately estimate $X_{1i}$ and $X_{0i}$.

12. Heckman (1981), Appendix A, presents a comprehensive analysis of factor-analytic discrete-choice models.

13. We ignore abstentions which are, less than one percent of the vote, and we assume that abstentions are random with respect to outcomes and preferences.

14. A variety of roll call studies use one- or two-dimensional models or measures of voting. This is true of representation studies (Markus 1974, Schwarz and Fenmore 1977, Schwarz, Fenmore and Volgy 1980, Bullock and Brady 1983, Glazer and Robbins 1985, Erikson 1990), committee studies (Kiewiet and McCubbins 1991, Cox and McCubbins 1993, Londregan and Snyder 1994), and studies of the links between roll-call voting and election outcomes (McAdams and Johannes 1983, Bond, Covington and Fleisher 1985, Whitby and Bledsoe 1986, Abramowitz 1988, Shaffer and Chressanthis 1990).

15. As in Poole and Rosenthal (1991a), we included all roll calls except those for which fewer than 2.5 percent of the votes were cast on the minority side. Also, for each Congress we included all representatives except those who cast votes on fewer than half of the roll calls.

16. Representatives are predicted to vote "yea" if their estimated probability of voting "yea" is greater than 1/2, they are predicted to vote "nay" if their estimated probability of voting "yea" is less than 1/2, and they are randomly assigned if their probability of voting "yea" is equal to 1/2.

17. This is true even after dropping the roll calls where 97.5 percent or more voted with the majority side.

# REFERENCES

Abramowitz, Alan. "Explaining Senate Election Outcomes." *American Political Science Review* Vol. 82 (2) June (198)8, pp. 385-403.

Amemiya, T. "Some Theorems on The Linear Probability Model", *International Economic Review*, 18(3) October 1977, 645-50.

_____. "Estimation of the Linear Probability Model." *International Economic Review*, 1983, pp. 310-341.

_____. *Advanced Econometrics*. Cambridge, MA: Harvard, 1985.

Andersen, E.B. *Conditional Inference and Models For Measuring*. Copenhagen: Mentalhygiejnisk Forlag, 1973.

_____. *Discrete Statistical Models With Social Science Applications*. Amsterdam: North Holland, 1980.

Anderson, Simon, De Palma, Andre and Thisse, Jacques-Francois, *Discrete Choice Theory of Product Differentiation*, Cambridge, Massachusetts: MIT, 1992.

Bartels, Larry and Henry Brady. "The State of Quantitative Political Methodology", Ada Finifter, *Political Science: The State of the Disciplines*, (1993), Washington, D.C. American Political Association.

Bond, John R., Gary Covington, and Richard Fleisher. "Explaining Challenger Quality in Congressional Elections". *Journal of Politics*, Vol. 47(2) (1985), pp. 510-529.

Brady, H. 1989. "Factor and Ideal Point Analysis For Interpersonally Incomparable Data". *Psychometrika*, Vol. 54 (1989), pp. 181-202.

Bullock, Charles S. III, and David W. Brady. "Party, Constituency, and Roll-Call Voting in the U.S. Senate". *Legislative Studies Quarterly*, Vol. 13 (1983), pp. 29-43.

Converse, Philip E. "The Nature of Belief Systems in Mass Publics." In David Apter (ed.), *Ideology and Discontent*. New York: Free Press, 1964.

Cox, Gary W., and Mathew D. McCubbins. *Legislative Leviathan*. Berkeley: University of California Press, 1993.

Cragg, John and Stephen, Donald. "Factor Analysis Under More General Unknown Form", in Advances in Econometrics and Quantitative Economics, pp. 291-310, G. S. Maddala, P.C.B. Phillips and T.N. Srinivasan, eds., Oxford, England: Basil Blackwell, 1995.

_____ . "Inferring The Rank of A Matrix". Forthcoming in *Journal of Econometrics*, 1996.

Denzau, Arthur T., and Robert J. Mackey. "Gatekeeping and Monopoly Power of Committees: An Analysis of Sincere and Sophisticated Behavior." *American Journal of Political Science*, Vol. 27 (1983), 432-447.

Denzau, Arthur T., and Michael C. Munger. "Legislators and Interest Groups: How Unorganized Interests Get Represented." *American Political Science Review*, Vol. 80(1) March (1986), pp. 89-106.

Domencich, T. and D. McFadden. *Urban Travel Demand: A Behavioral Analysis*. North Holland: Amsterdam, 1975.

Enelow, James and Melvin Hinich. *The Spatial Theory of Voting: An Introduction*, New York: Cambridge University Press, 1984.

_____. "Estimating The Parameters of A Spatial Model of Elections": An Empirical Test Base on The 1980 National Election Study", *Political Methodology*, Vol. 11, (1985), p. 249-268.

_____. "A General Probabilistic Spatial Theory of Elections", *Public Choice*, Vol. 61 (Vol. 1989), 101-113.

Erikson, Robert S. "Roll Calls, Reputations, and Representation in the U.S. Senate". *Legislative Studies Quarterly*, Vol. 15(4) November (1990), pp. 623-642.

Falmage, J. C. *Elements of Psychophysical Theory*. Oxford: Clarendon Press, 1985.

Ferguson, Thomas. "A Method of Generating Best Asymptotically Normal Estimates With An Application To The Estimation of Bacterial Densities", *Annals of Mathematical Statistics*, Vol. 29 (1958), pp. 1046-62.

Glazer, Amihai and Marc Robbins. "Congressional Responsiveness to Constituency Change". *American Journal of Political Science*, Vol. 29(2) May (1985), pp. 259-72.

Goldberger, Arthur S. *Econometric Theory*, Wiley: New York, 1964.

Goldfeld, Stephen and Richard Quandt. *Nonlinear Methods in Econometrics*. North Holland: Amsterdam, 1972.

Gorman, W. M. "A Possible Procedure For Analyzing Quality Differentials in the Egg Market", *Review of Economic Studies*, Vol. 47(5) October (1980), pp. 843-856, written in 1956.

Heckman, James. "Statistical Models For Discrete Panel Data", in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*, pp. 114-178, Cambridge, Ma: MIT Press, 1981.

Heckman, James and Jose Scheinkman. "The Importance of Bundling in a Gorman-Lancaster Model of Earnings", *Review of Economic Studies*, Vol. 54(2) April (1987), pp. 243-255.

Kiefer, Jack and Jacob Wolfowitz. "Consistency of The Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters", *Annals of Mathematical Statistics*, Vol. 27, (1956), pp. 887-906.

Kiewiet, D. Roderick, and Mathew D. McCubbins. *The Logic of Delegation*. Chicago: University of Chicago Press, 1991.

Koford, Kenneth. "Dimensions in Congressional Voting: Congressional parties and the Appropriations Process." *American Political Science Review*, Vol. 83(3) September (1989), pp. 949-62.

_____. "Dimensions, Transactions Costs and Coalitions in Legislative Voting." *Economics and Politics*, Vol. 2 (1990), pp. 59-82.

Kritzer, Herbert M. "Ideology and American Political Elites." *Public Opinion Quarterly*, Vol. 42 (1978), pp. 484-502.

Lancaster, K. *Consumer Demand: A New Approach*. New York: Columbia University, 1971.

_____. "A New Approach to Consumer Theory". *Journal of Political Economy*, Vol. 74 (1966), pp. 132-157.

Londregan, John, and James M. Snyder, Jr. "Comparing Committee and Floor Preferences." *Legislative Studies Quarterly*, Vol. 19, (1994), pp. 233-266.

Markus, Gregory B. "Electoral Coalitions and Senate Roll Call Behavior: An Ecological Analysis". *American Journal of Political Science*, Vol. 18 (1974), pp. 595-607.

Mayhew, David R. *Party Loyalty Among Congressmen*. Cambridge: Harvard University Press, 1966.

McAdams, John C. and John R. Johannes. "Determinants of Spending by House Challengers, 1974-1984". *Journal of Politics*, 31(3) August (19873), pp. 457-483.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka, ed., *Frontiers In Econometrics*, pp. 184-210, New York, Academic Press, 1974.

Neyman, J. and Elizabeth L. Scott. "Consistent Estimates Based on Partially Consistent Observations". *Econometrica*, Vol. 16 (1948), pp. 1-32.

Peltzman, Sam. "Toward a More General Theory of Regulation," *Journal of Law and Economics*, Vol. 19,(2) (August 1976), pp. 211-240.

Poole, Keith T. "Dimensions of Interest Group Evaluation of the U.S. Senate, 1969-1978". *American Journal of Political Science*, Vol. 25 (1981), pp. 49-67.

____. "Recent Developments in Analytical Models of Voting in the U.S. Congress". *Legislative Studies Quarterly*, Vol. 13(1) February (1988), pp. 117-133.

Poole, Keith T. and R. Steven Daniels. "Ideology, Party, and Voting in the U.S. Congress". *American Political Science Review*, Vol. 79(2) June (1985), pp. 373-399.

Poole, Keith T. and Howard Rosenthal. 1985a. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science*, Vol. 29(2) May (1985a), pp. 357-384.

_____. "The Unidimensional Congress: 1919-1984." Unpublished manuscript. Carnegie Mellon University, 1985b.

_____. "Analysis of Congressional Coalition Patterns: A Unidimensional Spatial Model." *Legislative Studies Quarterly*, Vol. 12(1) February (1987), pp. 55-75.

_____. "Patterns of Congressional Voting." *American Journal of Political Science*, Vol. 35 (1991a), pp. 228-278.

_____. "The Spatial Mapping of Minimum Wage Legislation." In *Politics and Economics in the Eighties*. Alberto Alesina and Geoffrey Carliner, eds. Chicago: University of Chicago Press, 1991b.

Pratt, John and Jean Gibbons. *Concepts of Nonparametric Theory*, NY: Springer-Verlag, 1981.

Pudney, Stephen, "Disaggregated Demand Analysis: The Estimation of A Class of Non-Linear Demand Systems", *Review of Economic Studies*, 47(5) (1980), p. 875-892.

Quandt, Richard E. "A Probabilistic Model of Consumer Behavior". *Quarterly Journal of Economics*, Vol. 70 (1956), pp. 507-536.

Quandt, Richard E. and William Baumol. "The Demand For Abstract Transport Modes: Theory and Measurement", *Journal of Regional Science*, Vol. 6, (1966), pp. 13-26.

Rabinowitz, George and Stuart McDonald. "A Directional Theory of Issue Voting". *American Political Science Review*, Vol. 83(1) March (1989), pp. 93-121.

Rao, C. R. *Linear Statistical Inference and Its Applications*. Second Edition. New York: Wiley, 1973.

Riker, William H. "Political Science and Rational Choice," pp. 163-181, in James Alt and Kenneth Shepsle, eds., *Perspectives and Positive Political Economy*, Cambridge, England: Cambridge University Press, 1990.

Schneider, Jerrold E. *Ideological Coalitions in Congress*. Westport, CT: Greenwood, 1979.

Schwarz, John E. and Barton Fenmore. "Presidential Election Results and Congressional Roll Call Behavior: The Cases of 1964, 1968, and 1972". *Legislative Studies Quarterly* Vol. 2 (1977), pp. 409-422.

Schwarz, John E, Barton Fenmore and Thomas J. Volgy. "Liberal and Conservative Voting in the House of Representatives: A National Model of Representation". *British Journal of Political Science*, Vol. 10 (1980), pp. 317-339.

Shaffer, Stephen D. and George A. Chressanthis. "Accountability and U.S. Senate Elections: A Multivariate Analysis", *Western Political Quarterly*, Vol. 44 (1990), pp. 625-639.

Shepsle, Kenneth. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science*, Vol. 23 (1979), pp. 27-60.

Stigler, George, "Theory of Economic Regulation", *Bell Journal of Economic and Management Science*, Vol. 2, No. 1 (1971), pp. 3-21.

Theil, Henri, *Principles of Econometrics*, Wiley: New York, 1971.

Thurstone, Louis, "A Law of Comparative Judgement", *Psychophysical Review*, Vol. 34, (1927), pp. 273-286.

Weisberg, Herbert F. "Evaluating Theories of Congressional Roll-Call Voting." *American Journal of Political Science*, Vol. 22 (1978), pp. 554-577.

Whitby, Kenny and Timothy Bledsoe. "The Impact of Policy Voting on the Electoral Fortunes of Senate Incumbents", *Western Political Quarterly*, Vol. 39(4) December (1986), pp. 690-700.

Table 1a

Comparing Different Methods of Estimating Preference Parameters of the Linear Model
(Two-Factor Model)

Senate

Correlations Between Principal Factor Estimates and:

| | Iterated Principal Factor | | Unweighted Least Squares | | Weighted Least Squares* | |
|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 1 | 2 | 1 | 2 |
| 80 | 1.0000 | 0.9998 | 1.0000 | 0.9996 | | |
| 81 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | | |
| 82 | 1.0000 | 0.9999 | 1.0000 | 0.9998 | | |
| 83 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | | |
| 84 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 85 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 86 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | |
| 87 | 1.0000 | 0.9999 | 1.0000 | 1.0000 | 0.9975 | 0.9643 |
| 88 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9874 | 0.9706 |
| 89 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | 0.9649 | 0.9888 |
| 90 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | 0.9954 | 0.9927 |
| 91 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | 0.9965 | 9.9831 |
| 92 | 1.0000 | 0.9999 | 1.0000 | 0.9998 | 0.9981 | 0.9646 |
| 93 | 1.0000 | 0.9998 | 1.0000 | 0.9999 | 0.9990 | 0.9504 |
| 94 | 1.0000 | 0.9996 | 1.0000 | 0.9998 | 0.9983 | 0.9655 |
| 95 | 1.0000 | 0.9986 | 1.0000 | 0.9996 | 0.9995 | 0.9899 |
| 96 | 1.0000 | 0.9998 | 1.0000 | 0.9998 | 0.9996 | 0.9970 |
| 97 | 1.0000 | 0.9999 | 1.0000 | 0.9998 | 0.9992 | 0.9937 |
| 98 | 1.0000 | 0.9999 | 1.0000 | 0.9999 | 0.9974 | 0.9937 |
| 99 | 1.0000 | 0.9996 | 1.0000 | 0.9998 | 0.9899 | 0.9818 |
| 100 | 1.0000 | 0.9997 | 1.0000 | 0.9995 | 0.9954 | 0.9648 |

Note: Average of 25 trials. A random subsamples of 20 senators was drawn for each trial. This estimator is the sub-diagonal method of Cragg and Donald as described below Theorem A-1 in the appendix. WLS results for the 80th-86th Congress are not included because too few roll calls were taken.

Table 1b

Comparing Different Methods of Estimating Preference Parameters of the Linear Model
(Two-Factors Model)

House of Representatives

Correlations Between Principal Factor Estimates and:

| | Iterated Principal Factor | | Unweighted Least Squares | | Weighted Least Squares* | |
|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 1 | 2 | 1 | 2 |
| 80 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | | |
| 81 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 82 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | | |
| 83 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 84 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 85 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | | |
| 86 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 87 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | | |
| 88 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | | |
| 89 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9975 | 0.9368 |
| 90 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9874 | 0.9443 |
| 91 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9849 | 0.9761 |
| 92 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9954 | 0.9762 |
| 93 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9965 | 0.9936 |
| 94 | 1.0000 | 1.0000 | 1.0000 | 0.9988 | 0.9981 | 0.9827 |
| 95 | 1.0000 | 1.0000 | 1.0000 | 0.9982 | 0.9990 | 0.9942 |
| 96 | 1.0000 | 1.0000 | 1.0000 | 0.9991 | 0.9983 | 0.9964 |
| 97 | 1.0000 | 1.0000 | 1.0000 | 0.9993 | 0.9995 | 0.9971 |
| 98 | 1.0000 | 1.0000 | 1.0000 | 0.9989 | 0.9996 | 0.9872 |
| 99 | 1.0000 | 1.0000 | 1.0000 | 0.9987 | 0.9992 | 0.9832 |
| 100 | 1.0000 | 1.0000 | 1.0000 | 0.9990 | 0.9978 | 0.9363 |

*Note: Average of 25 trials. This estimator is the sub-diagonal method of Cragg and Donald as described below Theorem A-1 in the appendix. A random subsamples of 20 representatives was drawn for each trial. WLS results for the 80th-88th Congress are not included because too few roll calls were taken.

Table 2a

Comparing the Estimated Preference Parameters of the Linear Model and NOMINATE Model (Two-Factor Model)

Senate

| | Correlations | | Rank-Order Correlations | | Multiple Correlations | | Classification Success | |
|---|---|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 1 | 2 | 1 | 2 | Linear Model | NOMINATE |
| 80 | 0.99 | 0.95 | 0.99 | 0.96 | 0.99 | 0.95 | .87 | .87 |
| 81 | 0.98 | 0.99 | 0.99 | 0.98 | 0.98 | 0.99 | .85 | .84 |
| 82 | 0.98 | 0.97 | 0.99 | 0.97 | 0.98 | 0.97 | .85 | .85 |
| 83 | 0.98 | 0.93 | 0.98 | 0.95 | 0.98 | 0.94 | .87 | .86 |
| 84 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 | 0.98 | .87 | .87 |
| 85 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 | 0.98 | .86 | .85 |
| 86 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | .86 | .87 |
| 87 | 0.96 | 0.97 | 0.98 | 0.96 | 0.96 | 0.97 | .86 | .86 |
| 88 | 0.97 | 0.95 | 0.98 | 0.95 | 0.97 | 0.97 | .87 | .86 |
| 89 | 0.99 | 0.98 | 0.98 | 0.98 | 0.99 | 0.98 | .85 | .86 |
| 90 | 0.99 | 0.96 | 0.99 | 0.97 | 0.99 | 0.98 | .84 | .84 |
| 91 | 0.99 | 0.97 | 0.99 | 0.98 | 0.99 | 0.98 | .85 | .85 |
| 92 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | .85 | .86 |
| 93 | 0.99 | 0.97 | 0.99 | 0.97 | 0.99 | 0.97 | .85 | .85 |
| 94 | 0.99 | 0.97 | 0.99 | 0.97 | 0.99 | 0.97 | .85 | .86 |
| 95 | 0.98 | 0.98 | 0.99 | 0.98 | 0.99 | 0.98 | .83 | .83 |
| 96 | 0.98 | 0.99 | 1.00 | 0.99 | 0.98 | 0.99 | .83 | .83 |
| 97 | 0.99 | 0.98 | 0.99 | 0.98 | 0.99 | 0.98 | .85 | .86 |
| 98 | 0.99 | 0.99 | 1.00 | 0.99 | 0.99 | 0.99 | .84 | .84 |
| 99 | 0.99 | 0.97 | 1.00 | 0.98 | 0.99 | 0.97 | .84 | .84 |
| 100 | 0.98 | 0.95 | 0.99 | 0.94 | 0.98 | 0.95 | .86 | .87 |

Table 2b

Comparing the Estimated Preference Parameters of the Linear Model and NOMINATE Model
(Two-Factor Model)

House of Representatives

| | Correlations | | Rank-Order Correlations | | Multiple Correlations | | Classification Success | |
|---|---|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 1 | 2 | 1 | 2 | Linear Model | NOMINATE |
| 80 | 0.95 | 0.87 | 0.96 | 0.91 | 0.96 | 0.92 | .89 | .86 |
| 81 | 0.97 | 0.98 | 0.98 | 0.97 | 0.98 | 0.98 | .87 | .87 |
| 82 | 0.94 | 0.98 | 0.96 | 0.97 | 0.94 | 0.98 | .87 | .83 |
| 83 | 0.97 | 0.96 | 0.98 | 0.96 | 0.98 | 0.96 | .86 | .85 |
| 84 | 0.95 | 0.97 | 0.97 | 0.98 | 0.96 | 0.97 | .84 | .83 |
| 85 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 | 0.99 | .85 | .84 |
| 86 | 0.96 | 0.99 | 0.98 | 0.99 | 0.97 | 0.99 | .87 | .84 |
| 87 | 0.95 | 0.97 | 0.99 | 0.98 | 0.96 | 0.98 | .88 | .86 |
| 88 | 0.97 | 0.98 | 0.98 | 0.98 | 0.97 | 0.98 | .89 | .88 |
| 89 | 0.98 | 0.96 | 0.97 | 0.97 | 0.98 | 0.98 | .89 | .87 |
| 90 | 0.98 | 0.98 | 0.99 | 0.98 | 0.99 | 0.98 | .88 | .88 |
| 91 | 0.97 | 0.94 | 0.98 | 0.94 | 0.98 | 0.95 | .87 | .87 |
| 92 | 0.96 | 0.93 | 0.97 | 0.93 | 0.97 | 0.94 | .85 | .85 |
| 93 | 0.97 | 0.97 | 0.99 | 0.97 | 0.97 | 0.98 | .84 | .83 |
| 94 | 0.97 | 0.95 | 0.99 | 0.94 | 0.97 | 0.95 | .84 | .84 |
| 95 | 0.98 | 0.96 | 1.00 | 0.97 | 0.98 | 0.97 | .84 | .85 |
| 96 | 0.97 | 0.95 | 1.00 | 0.95 | 0.97 | 0.95 | .85 | .85 |
| 97 | 0.97 | 0.93 | 0.99 | 0.93 | 0.97 | 0.93 | .86 | .85 |
| 98 | 0.95 | 0.93 | 0.99 | 0.92 | 0.96 | 0.93 | .86 | .85 |
| 99 | 0.97 | 0.92 | 0.99 | 0.91 | 0.97 | 0.92 | .86 | .85 |
| 100 | 0.96 | 0.91 | 0.99 | 0.89 | 0.97 | 0.91 | .88 | .86 |

Table 3a

Classification Success and Proportional Reduction in Error for the Linear Model
Comparison of One-Factor and Six-Factor Models

Senate

| | Classification Success | | | | | | Proportional Reduction in Error | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Close Roll Calls | | | Lopsided Roll Calls | | | All Roll Calls | | | |
| | # of Factors | | | # of Factors | | | # of Factors | | | Rel. |
| Cong. | 1 | 6 | Diff. | 1 | 6 | Diff. | 1 | 6 | Diff. | Diff. |
| 80 | .84 | .91 | .07 | .83 | .90 | .07 | .43 | .65 | .22 | 0.52 |
| 81 | .79 | .89 | .10 | .82 | .89 | .07 | .38 | .63 | .25 | 0.65 |
| 82 | .81 | .90 | .09 | .82 | .89 | .06 | .41 | .62 | .21 | 0.53 |
| 83 | .84 | .92 | .08 | .82 | .90 | .08 | .39 | .63 | .24 | 0.61 |
| 84 | .81 | .92 | .11 | .80 | .90 | .10 | .33 | .65 | .31 | 0.94 |
| 85 | .78 | .90 | .12 | .81 | .90 | .09 | .33 | .64 | .31 | 0.93 |
| 86 | .76 | .89 | .12 | .82 | .91 | .09 | .29 | .61 | .32 | 1.08 |
| 87 | .81 | .89 | .08 | .83 | .91 | .08 | .37 | .64 | .27 | 0.74 |
| 88 | .75 | .89 | .14 | .82 | .89 | .09 | .30 | .64 | .34 | 1.12 |
| 89 | .83 | .89 | .06 | .82 | .89 | .07 | .35 | .57 | .22 | 0.63 |
| 90 | .75 | .86 | .11 | .82 | .88 | .06 | .26 | .49 | .23 | 0.89 |
| 91 | .79 | .88 | .08 | .82 | .88 | .06 | .32 | .53 | .21 | 0.65 |
| 92 | .79 | .88 | .09 | .83 | .88 | .05 | .35 | .55 | .20 | 0.57 |
| 93 | .78 | .85 | .07 | .84 | .88 | .04 | .33 | .49 | .16 | 0.48 |
| 94 | .80 | .87 | .07 | .85 | .89 | .04 | .38 | .54 | .16 | 0.41 |
| 95 | .77 | .84 | .07 | .84 | .88 | .04 | .34 | .50 | .16 | 0.46 |
| 96 | .75 | .84 | .09 | .83 | .87 | .04 | .30 | .48 | .18 | 0.60 |
| 97 | .79 | .87 | .08 | .83 | .89 | .06 | .33 | .53 | .20 | 0.59 |
| 98 | .78 | .87 | .09 | .83 | .88 | .05 | .28 | .49 | .21 | 0.76 |
| 99 | .79 | .87 | .08 | .83 | .88 | .05 | .32 | .50 | .18 | 0.56 |
| 100 | .82 | .89 | .07 | .85 | .89 | .04 | .34 | .53 | .19 | 0.55 |
| Avg. | .79 | .88 | .09 | .83 | .89 | .06 | .34 | .57 | .23 | 0.68 |

Note: The columns labeled "Diff." give the difference in classification success (or proportional reduction in error) between a one-factor model and a six-factor model. The last column, labeled "Rel. Diff.", gives the relative difference in the proportional reduction in error between a one-factor model and a six-factor model.

Table 3b

Classification Success and Proportional Reduction in Error for the Linear Model
Comparison of One-Factor and Six-Factor Models

House of Representatives

| | Classification Success | | | | | | Proportional Reduction in Error | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Close Roll Calls | | | Lopsided Roll Calls | | | All Roll Calls | | | |
| | # of Factors | | | # of Factors | | | # of Factors | | | Rel. |
| Cong. | 1 | 6 | Diff. | 1 | 6 | Diff. | 1 | 6 | Diff. | Diff. |
| 80 | .87 | .92 | .05 | .86 | .92 | .06 | .43 | .65 | .22 | 0.52 |
| 81 | .81 | .91 | .09 | .82 | .90 | .08 | .36 | .60 | .24 | 0.68 |
| 82 | .84 | .91 | .06 | .84 | .90 | .06 | .50 | .66 | .16 | 0.33 |
| 83 | .82 | .90 | .09 | .82 | .89 | .08 | .36 | .59 | .23 | 0.65 |
| 84 | .79 | .89 | .10 | .80 | .89 | .08 | .31 | .57 | .26 | 0.84 |
| 85 | .77 | .88 | .12 | .80 | .88 | .08 | .32 | .57 | .25 | 0.79 |
| 86 | .78 | .89 | .10 | .84 | .91 | .07 | .38 | .61 | .23 | 0.60 |
| 87 | .84 | .91 | .07 | .85 | .90 | .05 | .42 | .60 | .18 | 0.44 |
| 88 | .87 | .92 | .05 | .85 | .90 | .05 | .43 | .58 | .15 | 0.35 |
| 89 | .87 | .91 | .04 | .86 | .90 | .05 | .44 | .59 | .15 | 0.33 |
| 90 | .83 | .90 | .07 | .86 | .89 | .03 | .33 | .47 | .15 | 0.45 |
| 91 | .78 | .87 | .09 | .86 | .90 | .04 | .28 | .46 | .18 | 0.65 |
| 92 | .78 | .86 | .08 | .85 | .89 | .04 | .32 | .49 | .17 | 0.54 |
| 93 | .77 | .84 | .07 | .84 | .88 | .04 | .27 | .43 | .16 | 0.62 |
| 94 | .79 | .85 | .06 | .84 | .88 | .04 | .31 | .46 | .15 | 0.46 |
| 95 | .78 | .84 | .06 | .85 | .88 | .04 | .29 | .43 | .14 | 0.48 |
| 96 | .80 | .85 | .05 | .85 | .88 | .03 | .33 | .46 | .13 | 0.38 |
| 97 | .80 | .86 | .06 | .85 | .89 | .04 | .30 | .47 | .17 | 0.57 |
| 98 | .81 | .87 | .05 | .87 | .90 | .03 | .37 | .53 | .16 | 0.44 |
| 99 | .83 | .88 | .05 | .85 | .89 | .04 | .43 | .56 | .13 | 0.31 |
| 100 | .86 | .90 | .04 | .87 | .90 | .03 | .46 | .58 | .12 | 0.26 |
| Avg. | .81 | .88 | .07 | .84 | .89 | .05 | .36 | .54 | .18 | 0.51 |

Note: The columns labeled "Diff." give the difference in classification success (or proportional reduction in error) between a one-factor model and a six-factor model. The last column, labeled "Rel. Diff.", gives the relative difference in the proportional reduction in error between a one-factor model and a six-factor model.

Table 4a

Predictions Using Party and Region as Observables

Senate

| Cong. | Using Party/Region to Predict Roll Calls: Classif. Success | PRE | Using Party/Region to Predict Preference Parameters: R-Squares 1 | 2 | 3-6 | Parameters from Votes vs. Parameters from Residuals: R-Squares 1 | 2 | 3-6 |
|---|---|---|---|---|---|---|---|---|
| 80 | .83 | .42 | .84 | .48 | .05 | .28 | .61 | .90 |
| 81 | .80 | .37 | .77 | .39 | .08 | .37 | .62 | .86 |
| 82 | .80 | .36 | .71 | .36 | .08 | .43 | .66 | .86 |
| 83 | .82 | .35 | .76 | .34 | .08 | .35 | .68 | .89 |
| 84 | .80 | .32 | .77 | .15 | .08 | .28 | .86 | .88 |
| 85 | .80 | .33 | .71 | .39 | .09 | .34 | .62 | .88 |
| 86 | .81 | .34 | .64 | .62 | .08 | .41 | .41 | .90 |
| 87 | .81 | .35 | .72 | .65 | .08 | .35 | .33 | .92 |
| 88 | .82 | .39 | .62 | .86 | .04 | .42 | .19 | .94 |
| 89 | .80 | .30 | .58 | .76 | .03 | .49 | .32 | .96 |
| 90 | .79 | .24 | .55 | .71 | .06 | .54 | .35 | .92 |
| 91 | .79 | .27 | .50 | .80 | .08 | .59 | .23 | .89 |
| 92 | .80 | .30 | .58 | .68 | .08 | .53 | .36 | .88 |
| 93 | .80 | .25 | .62 | .54 | .08 | .49 | .50 | .89 |
| 94 | .80 | .26 | .60 | .62 | .08 | .52 | .35 | .90 |
| 95 | .78 | .23 | .55 | .49 | .07 | .53 | .51 | .90 |
| 96 | .79 | .26 | .65 | .54 | .06 | .46 | .50 | .93 |
| 97 | .82 | .33 | .83 | .30 | .08 | .23 | .70 | .93 |
| 98 | .81 | .26 | .77 | .21 | .10 | .31 | .66 | .90 |
| 99 | .81 | .31 | .80 | .24 | .11 | .24 | .75 | .93 |
| 100 | .83 | .29 | .80 | .28 | .07 | .26 | .68 | .92 |
| Avg. | .81 | .31 | .68 | .50 | .07 | .40 | .53 | .90 |

Note: The last three columns give R-squares from regressing the preference parameters estimated using the roll call votes on the preference parameters estimated using the residuals from regressions of the roll call votes on party and region.

Table 4b

Predictions Using Party and Region as Observables

House of Representatives

| | Using Party/Region to Predict Roll Calls | | Using Party/Region to Predict Preference Parameters | | | Parameters from Votes vs. Parameters from Residuals | | |
|---|---|---|---|---|---|---|---|---|
| | Classif. | | R-Squares | | | R-Squares | | |
| Cong. | Success | PRE | 1 | 2 | 3-6 | 1 | 2 | 3-6 |
| 80 | .86 | .43 | .93 | .57 | .03 | .25 | .51 | .94 |
| 81 | .85 | .43 | .86 | .79 | .03 | .33 | .35 | .92 |
| 82 | .83 | .46 | .78 | .67 | .05 | .42 | .41 | .90 |
| 83 | .83 | .41 | .90 | .55 | .05 | .22 | .51 | .92 |
| 84 | .81 | .34 | .91 | .56 | .05 | .15 | .45 | .90 |
| 85 | .80 | .34 | .78 | .65 | .06 | .32 | .38 | .89 |
| 86 | .83 | .40 | .84 | .69 | .05 | .22 | .37 | .92 |
| 87 | .84 | .37 | .84 | .54 | .06 | .28 | .51 | .92 |
| 88 | .85 | .40 | .82 | .63 | .05 | .29 | .41 | .91 |
| 89 | .84 | .38 | .74 | .69 | .03 | .39 | .40 | .94 |
| 90 | .83 | .29 | .72 | .71 | .04 | .41 | .35 | .90 |
| 91 | .82 | .21 | .65 | .68 | .03 | .50 | .41 | .94 |
| 92 | .80 | .24 | .64 | .62 | .02 | .49 | .49 | .97 |
| 93 | .80 | .20 | .68 | .55 | .03 | .44 | .51 | .96 |
| 94 | .81 | .25 | .72 | .49 | .05 | .37 | .57 | .92 |
| 95 | .80 | .22 | .67 | .40 | .08 | .42 | .64 | .91 |
| 96 | .81 | .27 | .71 | .45 | .07 | .38 | .58 | .91 |
| 97 | .82 | .24 | .77 | .40 | .04 | .35 | .63 | .94 |
| 98 | .83 | .32 | .83 | .46 | .05 | .27 | .59 | .93 |
| 99 | .83 | .38 | .86 | .36 | .04 | .24 | .67 | .92 |
| 100 | .85 | .40 | .88 | .34 | .04 | .22 | .69 | .93 |
| Avg. | .83 | .33 | .79 | .56 | .05 | .33 | .50 | .92 |

Note: The last three columns give R-squares from regressing the preference parameters estimated using the roll call votes on the preference parameters estimated using the residuals from regressions of the roll call votes on party and region.

Table 5a

Lower Bounds on the Dimensionality of the Preference Parameter Space
Using Generalized Least Squares Estimates and Sequential Chi-square Test

Senate

| | Avg. Chi-Square for Model with Number of Factors = | | | | | | | | Avg. Predicted # of Factors | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | α=.05 | α=.01 |
| 87 | 47927 | 5541 | 834.2 | 423.8 | 236.4 | 110.7 | 55.3 | 27.6 | 6.55 | 6.46 |
| 88 | 44133 | 9928 | 1313.7 | 402.6 | 200.5 | 93.3 | 51.7 | 29.0 | 6.10 | 6.05 |
| 89 | 24268 | 3640 | 856.5 | 387.4 | 191.5 | 102.4 | 55.7 | 28.3 | 6.15 | 6.05 |
| 90 | 19777 | 3082 | 607.3 | 302.2 | 170.1 | 98.0 | 53.7 | 29.6 | 6.40 | 6.05 |
| 91 | 21574 | 3041 | 540.2 | 248.4 | 138.5 | 76.5 | 42.9 | 26.6 | 5.80 | 5.40 |
| 92 | 26997 | 4523 | 504.4 | 258.0 | 149.5 | 92.5 | 53.1 | 31.9 | 5.95 | 5.75 |
| 93 | 19737 | 1828 | 425.5 | 233.0 | 149.1 | 95.5 | 56.5 | 34.1 | 5.95 | 5.70 |
| 94 | 18269 | 1853 | 553.8 | 302.9 | 184.2 | 113.2 | 67.6 | 40.2 | 6.70 | 6.30 |
| 95 | 20240 | 1298 | 621.1 | 339.5 | 204.9 | 113.7 | 69.8 | 41.6 | 6.60 | 6.35 |
| 96 | 12839 | 1793 | 483.1 | 256.9 | 151.6 | 102.3 | 67.1 | 38.4 | 6.15 | 5.75 |
| 97 | 25567 | 3636 | 749.9 | 311.8 | 167.1 | 93.6 | 53.5 | 31.7 | 6.05 | 5.95 |
| 98 | 20122 | 2491 | 524.3 | 266.8 | 153.0 | 90.5 | 53.6 | 30.5 | 6.15 | 5.75 |
| 99 | 17375 | 1786 | 625.4 | 321.3 | 170.9 | 95.3 | 55.6 | 33.6 | 6.10 | 5.90 |
| 100 | 21635 | 1925 | 642.9 | 302.2 | 172.4 | 99.4 | 60.4 | 31.1 | 6.10 | 6.10 |
| $\chi^2(.05)$ | 201.4 | 180.7 | 160.9 | 142.1 | 124.3 | 107.5 | 91.7 | 76.8 | | |
| $\chi^2(.01)$ | 215.8 | 194.4 | 173.9 | 154.4 | 135.8 | 118.2 | 101.6 | 86.0 | | |

Note: Results based on 25 trials. A random subsamples of 20 senators was used in each trial.
Note: Results for the 80th-86th Congress are not included because too few roll calls were taken.

Table 5b

Lower Bounds on the Dimensionality of the Preference Parameter Space
Using Generalized Least Squares Estimates and Sequential Chi-square Test

House of Representatives

| | Avg. Chi-Square for Model with Number of Factors = | | | | | | | | Avg. Predicted # of Factors | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cong. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $\alpha$=.05 | $\alpha$=.01 |
| 89 | 31208 | 2741 | 669.3 | 284.1 | 142.7 | 80.2 | 43.7 | 24.5 | 5.85 | 5.55 |
| 90 | 27041 | 3737 | 487.6 | 248.2 | 133.0 | 75.8 | 45.2 | 24.2 | 5.60 | 5.45 |
| 91 | 18599 | 1916 | 543.9 | 243.1 | 129.5 | 74.0 | 45.2 | 26.8 | 5.50 | 5.40 |
| 92 | 17364 | 1556 | 467.2 | 228.9 | 134.1 | 78.8 | 46.0 | 26.4 | 5.70 | 5.45 |
| 93 | 23060 | 1650 | 451.6 | 221.1 | 131.4 | 81.4 | 46.0 | 28.3 | 5.50 | 5.20 |
| 94 | 25057 | 1724 | 579.7 | 256.0 | 136.9 | 81.8 | 50.5 | 30.6 | 5.77 | 5.46 |
| 95 | 24979 | 1425 | 480.6 | 254.9 | 153.5 | 89.2 | 55.7 | 33.5 | 5.95 | 5.75 |
| 96 | 31808 | 1576 | 538.8 | 250.7 | 141.4 | 85.5 | 52.3 | 32.8 | 5.70 | 5.55 |
| 97 | 31360 | 1261 | 443.2 | 223.0 | 130.5 | 79.3 | 45.9 | 29.9 | 5.65 | 5.35 |
| 98 | 27454 | 1570 | 591.9 | 295.9 | 152.9 | 90.4 | 52.8 | 30.8 | 6.00 | 5.75 |
| 99 | 22314 | 1483 | 510.2 | 275.9 | 158.2 | 87.8 | 52.4 | 32.0 | 6.00 | 5.85 |
| 100 | 29720 | 1543 | 591.0 | 296.7 | 152.9 | 81.1 | 48.7 | 30.5 | 6.00 | 5.70 |
| | | | | | | | | | | |
| $\chi^2(.05)$ | 201.4 | 180.7 | 160.9 | 142.1 | 124.3 | 107.5 | 91.7 | 76.8 | | |
| $\chi^2(.01)$ | 215.8 | 194.4 | 173.9 | 154.4 | 135.8 | 118.2 | 101.6 | 86.0 | | |

Note: Results based on 25 trials. A random subsamples of 20 representatives was used in each trial.
Note: Results for the 80th-88th Congress are not included because too few roll calls were taken.

Table 6a

Stability of the Estimated Preference Parameters Over Time

Senate

| Cong. Pair | Correlations Between Individual Factors | | | | | | | | Multiple Correlations Using All 8 Factors | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 80,81 | .96 | .83 | .77 | .14 | .11 | .03 | .04 | .18 | .97 | .91 | .92 | .75 | .63 | .47 | .45 | .43 |
| 81,82 | .97 | .95 | .93 | .70 | .53 | .11 | .14 | .21 | .97 | .97 | .95 | .75 | .69 | .33 | .32 | .52 |
| 82,83 | .93 | .76 | .78 | .56 | .15 | .10 | .19 | .16 | .97 | .85 | .83 | .67 | .44 | .36 | .59 | .32 |
| 83,84 | .98 | .82 | .56 | .54 | .34 | .06 | .05 | .21 | .99 | .91 | .69 | .72 | .68 | .40 | .29 | .54 |
| 84,85 | .97 | .89 | .70 | .56 | .51 | .03 | .00 | .15 | .98 | .94 | .88 | .70 | .70 | .60 | .52 | .32 |
| 85,86 | .93 | .88 | .74 | .09 | .23 | .36 | .36 | .11 | .98 | .98 | .85 | .78 | .87 | .76 | .53 | .48 |
| 86,87 | .97 | .95 | .47 | .41 | .03 | .37 | .16 | .32 | .98 | .98 | .71 | .81 | .60 | .68 | .65 | .57 |
| 87,88 | .84 | .84 | .64 | .60 | .40 | .40 | .24 | .18 | .97 | .98 | .87 | .85 | .76 | .62 | .63 | .52 |
| 88,89 | .95 | .94 | .79 | .70 | .57 | .28 | .39 | .30 | .98 | .97 | .88 | .85 | .74 | .64 | .52 | .37 |
| 89,90 | .97 | .93 | .84 | .76 | .19 | .40 | .10 | .25 | .98 | .96 | .90 | .84 | .69 | .69 | .38 | .70 |
| 90,91 | .97 | .96 | .66 | .31 | .41 | .43 | .06 | .38 | .98 | .98 | .83 | .75 | .76 | .58 | .63 | .46 |
| 91,92 | .98 | .96 | .78 | .51 | .19 | .22 | .18 | .03 | .99 | .97 | .88 | .87 | .82 | .67 | .44 | .37 |
| 92,93 | .98 | .95 | .79 | .31 | .53 | .60 | .28 | .03 | .99 | .96 | .89 | .74 | .84 | .78 | .52 | .42 |
| 93,94 | .99 | .94 | .88 | .47 | .34 | .22 | .14 | .18 | .99 | .97 | .93 | .82 | .80 | .64 | .74 | .42 |
| 94,95 | .96 | .56 | .41 | .02 | .56 | .37 | .38 | .09 | .97 | .93 | .84 | .59 | .80 | .64 | .51 | .70 |
| 95,96 | .97 | .77 | .58 | .72 | .65 | .05 | .01 | .27 | .99 | .96 | .90 | .81 | .79 | .60 | .34 | .41 |
| 96,97 | .95 | .86 | .77 | .33 | .33 | .10 | .39 | .31 | .98 | .94 | .91 | .78 | .75 | .36 | .57 | .58 |
| 97,98 | .97 | .88 | .79 | .65 | .41 | .20 | .11 | .51 | .98 | .94 | .90 | .80 | .78 | .80 | .49 | .67 |
| 98,99 | .98 | .91 | .55 | .64 | .28 | .60 | .33 | .25 | .98 | .96 | .84 | .83 | .77 | .77 | .44 | .55 |
| 99,100 | .97 | .91 | .82 | .29 | .21 | .47 | .07 | .14 | .99 | .95 | .86 | .84 | .31 | .77 | .69 | .32 |

Note: Correlations and multiple correlations are calculated using all representatives who served in both congresses in the pair. Multiple correlations are calculated using all 8 factors from the adjacent congress.

Table 6b

Stability of the Estimated Preference Parameters Over Time

House of Representatives

| Cong. Pair | Correlations Between Individual Factors 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Multiple Correlations Using All 8 Factors 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 80,81 | .95 | .83 | .69 | .53 | .30 | .10 | .02 | .02 | .97 | .93 | .80 | .64 | .50 | .41 | .30 | .32 |
| 81,82 | .97 | .92 | .69 | .29 | .21 | .12 | .14 | .02 | .98 | .95 | .84 | .63 | .55 | .48 | .39 | .27 |
| 82,83 | .92 | .84 | .72 | .39 | .02 | .00 | .00 | .08 | .97 | .91 | .80 | .56 | .60 | .43 | .40 | .23 |
| 83,84 | .97 | .92 | .50 | .19 | .38 | .02 | .32 | .07 | .98 | .93 | .68 | .72 | .51 | .52 | .44 | .16 |
| 84,85 | .92 | .91 | .54 | .15 | .30 | .26 | .16 | .11 | .97 | .96 | .80 | .64 | .58 | .35 | .34 | .41 |
| 85,86 | .97 | .96 | .78 | .43 | .34 | .24 | .23 | .25 | .98 | .97 | .86 | .58 | .55 | .46 | .58 | .33 |
| 86,87 | .97 | .93 | .62 | .03 | .21 | .19 | .37 | .10 | .98 | .95 | .75 | .59 | .56 | .30 | .49 | .27 |
| 87,88 | .98 | .95 | .77 | .40 | .55 | .08 | .15 | .00 | .98 | .96 | .82 | .65 | .72 | .39 | .44 | .43 |
| 88,89 | .95 | .88 | .51 | .15 | .20 | .05 | .14 | .21 | .98 | .96 | .79 | .81 | .63 | .50 | .40 | .28 |
| 89,90 | .98 | .96 | .80 | .67 | .51 | .46 | .10 | .35 | .99 | .97 | .84 | .78 | .77 | .60 | .42 | .58 |
| 90,91 | .95 | .90 | .79 | .48 | .57 | .09 | .08 | .38 | .98 | .95 | .87 | .74 | .68 | .44 | .70 | .58 |
| 91,92 | .98 | .94 | .87 | .76 | .73 | .51 | .53 | .15 | .98 | .96 | .91 | .77 | .79 | .66 | .59 | .29 |
| 92,93 | .98 | .94 | .88 | .78 | .69 | .69 | .24 | .18 | .98 | .96 | .92 | .82 | .79 | .79 | .53 | .34 |
| 93,94 | .98 | .90 | .76 | .84 | .74 | .55 | .25 | .14 | .99 | .97 | .92 | .89 | .88 | .63 | .59 | .35 |
| 94,95 | .98 | .96 | .92 | .67 | .66 | .45 | .01 | .13 | .99 | .97 | .94 | .85 | .88 | .86 | .39 | .48 |
| 95,96 | .98 | .93 | .88 | .79 | .71 | .72 | .41 | .33 | .99 | .97 | .94 | .89 | .92 | .90 | .70 | .78 |
| 96,97 | .98 | .94 | .89 | .67 | .24 | .12 | .06 | .40 | .98 | .96 | .92 | .78 | .67 | .74 | .63 | .61 |
| 97,98 | .98 | .93 | .85 | .70 | .34 | .10 | .57 | .50 | .99 | .96 | .92 | .91 | .84 | .87 | .77 | .73 |
| 98,99 | .99 | .92 | .78 | .76 | .36 | .33 | .43 | .40 | .99 | .93 | .90 | .88 | .59 | .72 | .78 | .75 |
| 99,100 | .99 | .93 | .91 | .26 | .28 | .70 | .66 | .37 | .99 | .95 | .95 | .82 | .91 | .86 | .85 | .70 |

Note: Correlations and multiple correlations are calculated using all representatives who served in both congresses in the pair. Multiple correlations are calculated using all 8 factors from the adjacent congress.

Table 7

Roll Calls Characteristics
Roll Calls with High Scores on Factors 2-7, 90th House (1955-1957)

Factor 2: civil rights; agriculture; debt ceiling; gun control; public works appropriations

| | | |
|---|---|---|
| 3.7 | 5.0 | Extend civil rights commission |
| 3.0 | 5.0 | Penalties for interfering with civil rights |
| 3.6 | 5.4 | Penalties for interfering with civil rights |
| 1.7 | 3.8 | Racial discrimination in selection of federal juries |
| 7.3 | 4.3 | Racial discrimination in housing |
| 6.2 | 4.9 | Civil rights and discrimination in housing |
| 5.5 | 4.2 | Federal funds and school desegregation |
| 5.5 | 4.8 | Federal funds and school desegregation |
| 5.7 | 5.5 | Uniform Congressional District Act |
| 7.1 | 4.4 | Congressional districting and at-large elections |
| 4.2 | 4.4 | Equal-time requirements for 1968 presidential election |
| 4.7 | 3.7 | Equal-time requirements for 1968 presidential election |
| 6.2 | 3.9 | Seating and censuring of Adam Clayton Powell |
| 5.9 | 3.5 | Seating and censuring of Adam Clayton Powell |
| 5.4 | 4.4 | Seating and censuring of Adam Clayton Powell |
| 2.2 | 4.2 | Counsel for Adam Clayton Powell |
| 5.9 | 4.9 | Agricultural appropriations |
| 4.1 | 4.4 | Peanut acreage allotments |
| 0.5 | 3.6 | Peanut acreage allotments |
| 0.4 | 5.6 | Amend Food and Agriculture Act of 1965 -- limit farm payments |
| 0.1 | 4.6 | Amend Food and Agriculture Act of 1965 |
| 0.5 | 4.3 | Amend Food and Agriculture Act of 1965 |
| 1.2 | 5.7 | Amend Food and Agriculture Act of 1965 |
| 4.6 | 4.8 | Soil conservation service |
| 4.9 | 3.8 | Marketing orders for cherries |
| 7.3 | 6.6 | Increase debt limit |
| 5.5 | 5.9 | Increase debt limit |
| 8.1 | 4.9 | Increase debt limit |
| 8.0 | 4.9 | Increase debt limit |
| 7.2 | 5.1 | Increase debt limit |
| 8.0 | 5.8 | Increase debt limit |
| 8.4 | 3.8 | Continuing appropriations |
| 8.4 | 5.3 | Continuing appropriations |
| 7.3 | 5.6 | Continuing appropriations |
| 6.2 | 5.0 | Defense appropriations |
| 5.9 | 5.0 | Defense appropriations |
| 6.7 | 3.9 | Investigation of U.S. military credit unions |
| 6.5 | 3.7 | Public works appropriations -- Dickey-Lincoln hydroelectric plant |
| 5.8 | 4.7 | Public works appropriations |
| 3.9 | 3.8 | Public works appropriations |
| 3.8 | 4.0 | Public works appropriations |
| 6.0 | 6.1 | Funds for committee on science and astonautics |
| 7.2 | 3.9 | Transportation appropriations |
| 6.9 | 3.8 | Post-office and Treasury appropriations |
| 7.0 | 5.8 | Housing and urban development appropriations |
| 5.6 | 5.2 | Amend national housing act -- regulation of savings and loans |
| 6.6 | 3.9 | Higher education expenditures -- teacher corps |
| 6.2 | 4.8 | Regulation of firearms |

Table 7 (continued)

3.1 5.1 Regulation of firearms
5.6 4.7 Regulation of firearms -- exempt persons in military training
4.8 4.5 Regualtion of fireamrs
6.8 4.5 Create bureau of narcotics and dangerous drugs
5.4 4.1 Interest equalization tax
0.8 4.8 Revision of copyright law

Factor 3: criminal law & civil liberties; revenue & expenditure act of 1968
5.6 4.6 Hearings by House Un-American Activities Committee
4.8 4.3 Funds for House Un-American Activities Committee
3.9 4.1 Legalization of wiretapping and admissibility of confessions
7.6 3.5 Oath of office to Adam Clayton Powell
2.2 3.9 Revenue and Expenditure Control Act of 1968
1.8 3.7 Revenue and Expenditure Control Act of 1968

Factor 4: foreign aid
8.1 3.0 Foreign Assistance Act of 1967
8.2 3.4 Foreign Assistance Act of 1967 -- trading with North Vietnam
7.7 3.6 Foreign aid appropriations
8.0 3.8 Foreign aid appropriations
8.1 3.8 Foreign aid appropriations
7.7 3.8 Foreign Assistance Act of 1968
7.2 3.7 Foreign Assistance Act of 1968
7.2 4.1 Foreign aid appropriations
6.7 3.8 Foreign aid appropriations
2.2 3.4 Revenue and Expenditure Control Act of 1968

Factor 6: agriculture; irrigation projects
0.8 2.8 Amend Food and Agriculture Act of 1965
0.4 3.8 Amend Food and Agriculture Act of 1965
0.1 4.9 Amend Food and Agriculture Act of 1965
0.5 3.6 Amend Food and Agriculture Act of 1965
1.2 4.6 Amend Food and Agriculture Act of 1965
0.1 2.9 Funding for Missouri River irrigation project
1.9 3.8 Funding for Missouri River irrigation project

Factor 7: federal school aid; NASA appropriations
4.5 3.3 Federal aid to schools
2.8 3.2 Federal aid to schools
3.7 2.8 Federal aid to schools
5.0 3.4 NASA appropriations, 1968
3.3 3.2 NASA appropriations, 1968
0.5 3.0 NASA appropriations, 1969

Note: For each roll call in the group labeled "Factor i", the number in the first column give the factor scores on the first factor, and the numbers in the second column give the score on factor i.

Table 8

Roll Call Characteristics
The Issue Content of Selected Factors, House of Representatives

| Cong. | Factor | Issues with High Scores on the Factor |
|---|---|---|
| 80 | 2 | abolish poll tax; D.C. home rule; labor-employer relations |
| 80 | 3 | military spending; foreign aid |
| 80 | 4 | agriculture; reclamation projects; mining policy |
| 81 | 2 | civil rights; voting rights; D.C. appropriations; agriculture; natural gas regulation |
| 81 | 3 | foreign aid |
| 81 | 4 | mining subsidies |
| 81 | 5 | WWII pensions |
| 82 | 2 | veterans hospitals for blacks; agriculture; appropriations; investigations |
| 82 | 3 | military training; foreign aid |
| 83 | 2 | immigration quotas; foreign aid; agriculture; housing |
| 83 | 3 | foreign aid; trade policy |
| 83 | 4 | St. Lawrence seaway |
| 84 | 2 | civil rights; school construction & desegregation; foreign aid; natural gas regulation |
| 84 | 3 | trade policy |
| 84 | 4 | foreign aid |
| 84 | 5 | natural gas regulation; airline regulation |
| 85 | 2 | civil rights; federal/state relations; foreign aid; agriculture |
| 85 | 3 | foreign aid |
| 85 | 4 | price supports for minerals; agriculture (corn, soil bank) |
| 85 | 5 | trade policy |
| 86 | 2 | civil rights; school construction; agriculture; foreign aid; labor policy; criminal law & civil liberties |
| 86 | 3 | foreign aid; foreign investment taxes |
| 86 | 4 | area redevelopment act |
| 87 | 2 | civil rights; poll tax; foreign aid; agriculture (including agricultural labor policy) |
| 87 | 3 | transfer of army food institute (from Chicago to Massachusetts) |
| 87 | 4 | nuclear electric plant in Hanford, Washington |
| 88 | 2 | civil rights; aid to education; federal courts and reapportionment; agriculture |
| 88 | 3 | foreign aid; peace corps |
| 89 | 2 | civil rights; voting rights; immigration reform; D.C. home rule; agriculture |
| 89 | 3 | criminal law & civil liberties |
| 89 | 4 | foreign aid |
| 90 | 2 | civil rights; agriculture; gun control; debt ceiling; public works appropriations |
| 90 | 3 | criminal law & civil liberties |
| 90 | 4 | foreign aid |
| 90 | 6 | agriculture; irrigation projects |
| 90 | 7 | federal school aid; NASA appropriations |

Table 8 (continued)

| | | |
|---|---|---|
| 91 | 2 | housing policy; labor & HEW spending; affirmative action; agriculture; trade policy |
| 91 | 3 | foreign aid; Internal Security Committee appropriations |
| 91 | 4 | agriculture |
| 91 | 6 | trade policy |
| 92 | 2 | amend House rules; agriculture & rural aid; education policy (including discrimination); affirmative action |
| 92 | 3 | debt ceiling; foreign aid |
| 92 | 4 | school busing & desegregation; school prayer |
| 93 | 2 | agriculture & rural aid; House rules & committees reforms; presidential impoundments; |
| 93 | 3 | debt ceiling; foreign aid |
| 93 | 4 | energy act; school busing; abortion |
| 93 | 5 | agriculture |
| 93 | 6 | aid to education |
| 94 | 2 | agriculture; water projects; House rules & committee reforms; campaign finance reform; energy conservation |
| 94 | 3 | debt ceiling; foreign aid |
| 94 | 4 | debt ceiling; motions to adjourn |
| 94 | 5 | police officers' bill-of-rights; agriculture |
| 94 | 6 | federal civil service |
| 95 | 2 | agriculture; water projects; maritime policy; House rules & televised proceedings; budget |
| 95 | 3 | foreign aid |
| 95 | 4 | federal funds for abortions; tuition tax credit & private schools |
| 95 | 5 | tuition tax credit & private schools |
| 95 | 7 | federal funds for abortions |
| 95 | 8 | sugar price supports |
| 96 | 2 | water projects; defense spending; federal budget; energy policy |
| 96 | 3 | federal revenue sharing |
| 96 | 4 | windfall profits tax; low-income fuel assistance |
| 97 | 2 | water projects; defense spending; agriculture |
| 97 | 3 | foreign aid |
| 97 | 4 | budget reconciliation; congressional compensation |
| 98 | 2 | defense spending; central american policy; water projects; agriculture & rural aid |
| 98 | 3 | debt ceiling; foreign aid & foreign policy |
| 98 | 7 | abortion |
| 99 | 2 | defense spending; central american policy; foregin aid & foreign policy; agriculture; water projects |
| 99 | 3 | social and supplemental expenditures |
| 99 | 4 | debt ceiling |
| 99 | 5 | approve journal ("protest" votes by a group of members) |
| 99 | 7 | abortion & family planning |
| 100 | 2 | defense spending; central american policy;nuclear energy regulation |
| 100 | 3 | housing & other social expenditures |
| 100 | 4 | approve journal ("protest" votes by a group of members) |
| 100 | 5 | public works, interior, & post-office expenditures |

Table 9

Proportional Reduction in Error for Selected Issues, House of Representatives

| Issue | Cong. | # Roll Calls | Average PRE<br># Factors in Model<br>1 | 2 | 6 |
|---|---|---|---|---|---|
| Abortion | 93-95 | 26 | .28 | .31 | .61 |
| Abortion | 96-100 | 26 | .44 | .45 | .59 |
| Agriculture | 80-89 | 165 | .39 | .49 | .61 |
| Agriculture | 90-100 | 349 | .19 | .30 | .41 |
| Alaska/Hawaii Statehood | 80-86 | 15 | .17 | .44 | .56 |
| Civil Rights | 80-90 | 101 | .22 | .71 | .76 |
| Civil Rights | 91-100 | 127 | .47 | .51 | .62 |
| Defense Spending | 80 | 6 | .14 | .14 | .64 |
| Defense Spending | 81-86 | 27 | .40 | .45 | .51 |
| Defense Spending | 87-97 | 372 | .24 | .35 | .48 |
| Defense Spending | 98-100 | 232 | .46 | .56 | .62 |
| Foreign Aid | 80-87 | 91 | .33 | .55 | .74 |
| Foreign Aid | 88-100 | 429 | .42 | .49 | .59 |
| Immigration | 80-100 | 80 | .25 | .37 | .46 |
| Water Projects | 80-89 | 53 | .42 | .46 | .54 |
| Water Projects | 90-100 | 114 | .10 | .24 | .32 |
| Trade Policy & Tariffs | 80-82 | 10 | .68 | .74 | .76 |
| Trade Policy & Tariffs | 83-86 | 19 | .14 | .21 | .55 |
| Trade Policy & Tariffs | 87-88 | 11 | .52 | .62 | .68 |
| Trade Policy & Tariffs | 89-97 | 148 | .25 | .30 | .41 |
| Trade Policy & Tariffs | 98-100 | 97 | .41 | .42 | .50 |

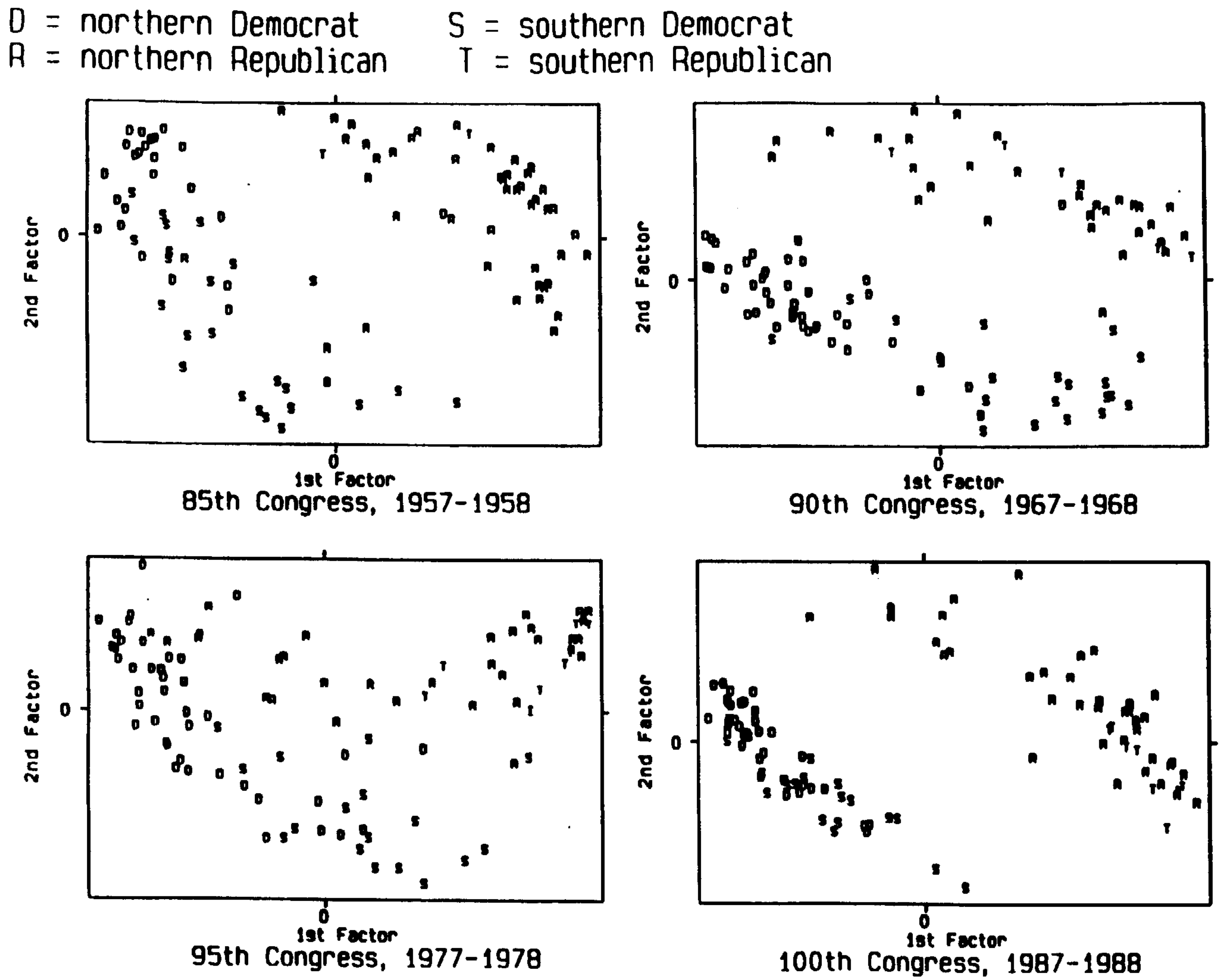


Figure 1

Estimated Preferences, Senate

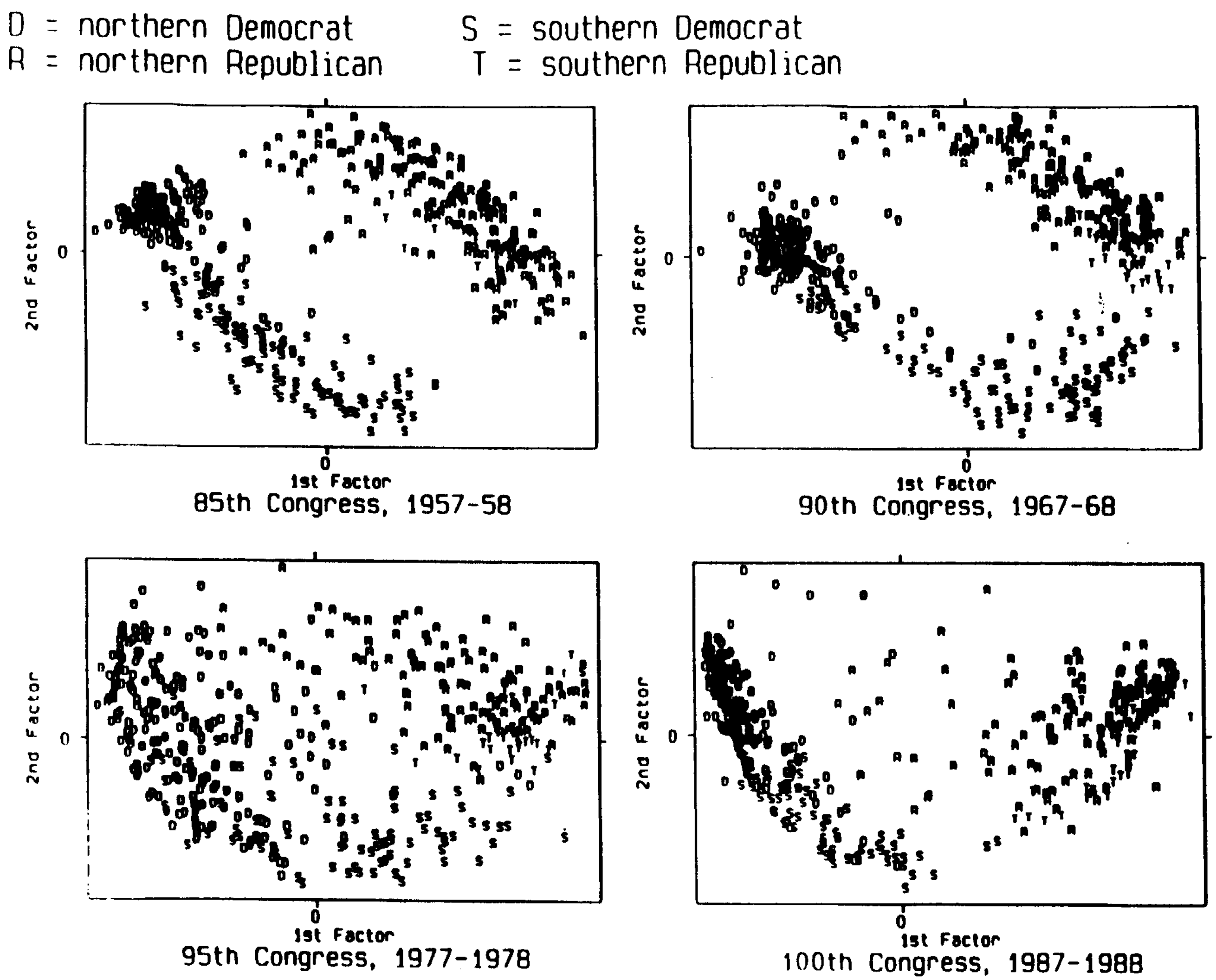


Figure 2

Estimated Preferences, House

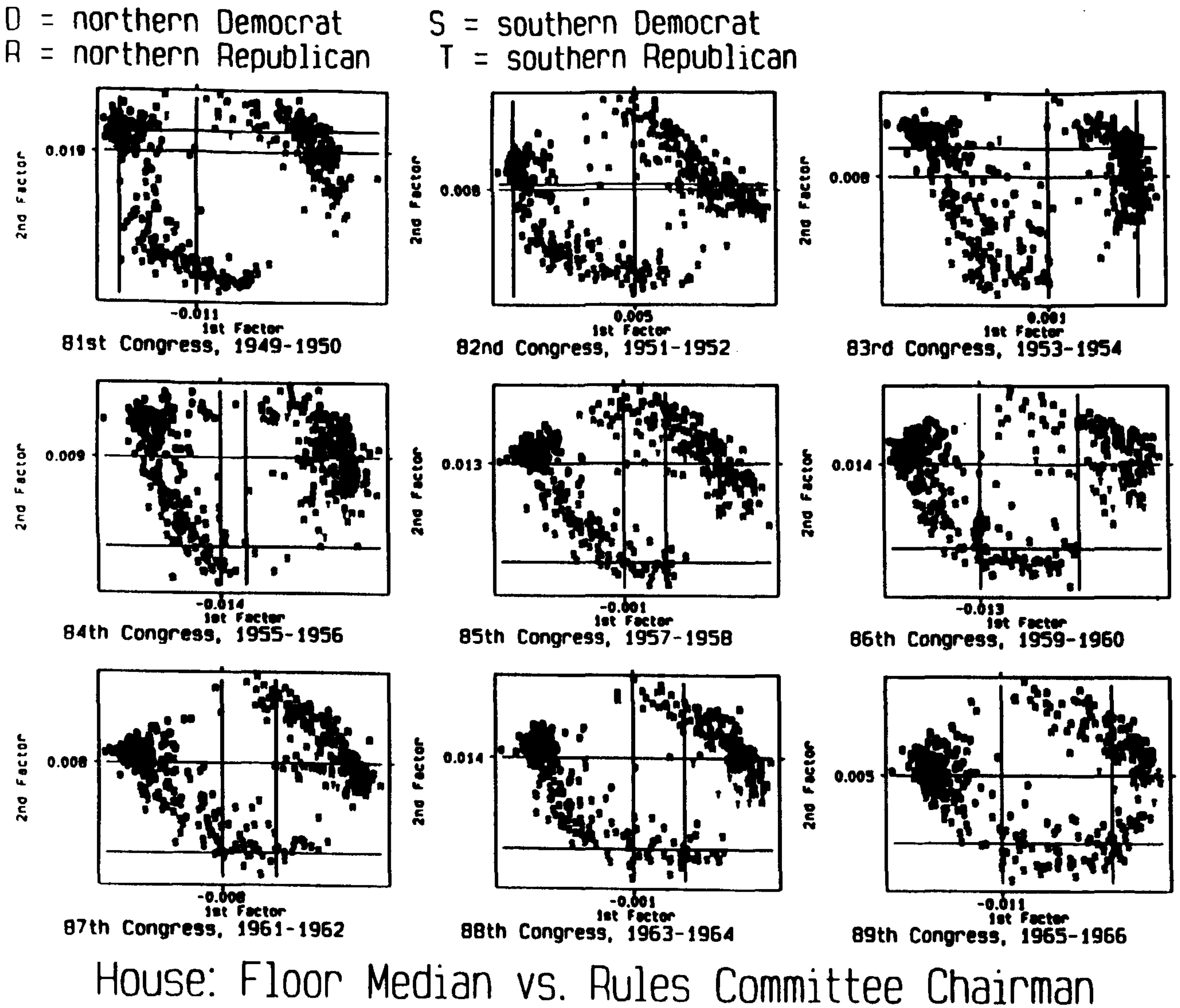


Figure 3

Preferences of Rules Committee and The House

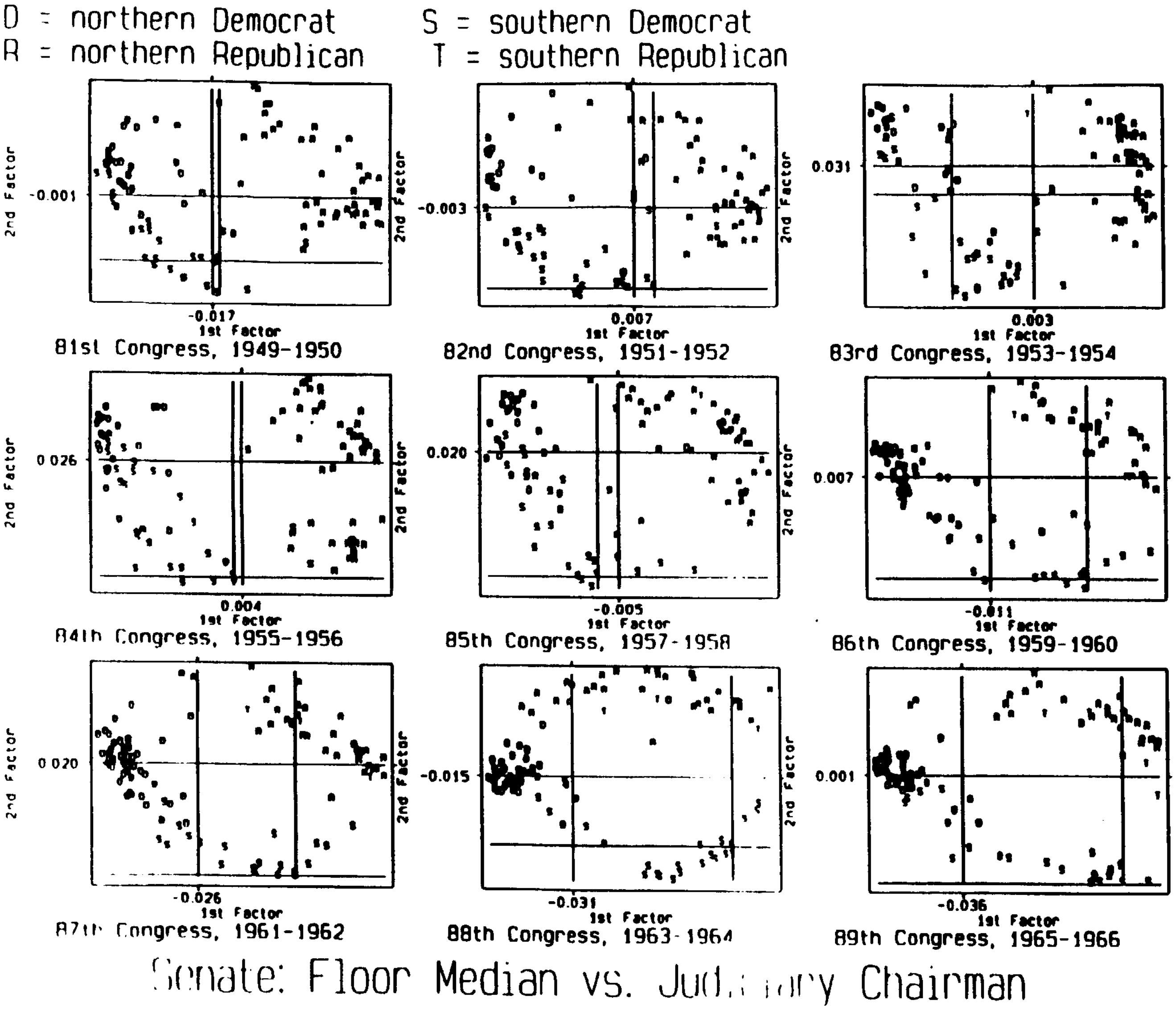


Figure 4

Preferences of Judiciary Committee and Senate

## Appendix: Formal Inference and Selecting the Effective Dimension of The Model

Assume that there are no observed factors in the model ($\beta \equiv 0$). If not, replace $D_i$ with $D_i - \beta X_{1i}$ in what follows and make all assumptions conditional on X. Inference is based on the covariance matrix of $D_i$. Assume further the following properties characterize $Z_i$, $\varepsilon_i$ and $\Phi$. Let $\Theta^0$ denote the true value of parameter vector $\Theta$.

(A-1) $Z_i$ is iid so $Z_i \perp\!\!\!\perp Z_{i'}$, $i \neq i'$.
Further $E(Z_i) = 0$, $Var(Z_i) = E(Z_i Z_i') = I_J$ and $E(\mid Z_{i\ell} \mid^{4+\sigma}) < \infty$ for all i, $\ell$, and some $\sigma > 0$.

(A-2) $\Phi\Phi'$ is positive definite with bounded characteristic roots.

(A-3) $E(\zeta_i \mid Z_{i'}) = 0$ for all i, i'. $E(\zeta_i \zeta_i' \mid Z_i)$ is diagonal and depends on i only through its dependence on $Z_i$, $E(\zeta_i \zeta_i' \mid Z_i) = \eta(Z_i)$, $0 < \gamma \leq \mid \zeta_{i\ell} \mid \leq \Delta < \infty$ all i, $\ell$. Removing the condition on $Z_i$, $E(\zeta_i \zeta_i') = \Psi$.

(A-4) Since the elements of the $\Phi$ are functions of L parameters $\omega$. Define $\Theta = \{\omega, \psi\}$ where $\psi$ is the parameter vector associated with $\Psi$. (In general J distinct values). Then $\Theta$ is a L + J vector and assume $\underline{\underline{\Theta}} \subset R^{L+J}$ is a compact parameter space - $\Theta^0 \in int(\underline{\underline{\Theta}})$ and that $\Sigma(\Theta)$ is a twice-continuously differentiable function of $\Theta$ and $\text{Rank}\left[\frac{\partial \text{Vech}\Sigma(\Theta)}{\partial \Theta}\right] = L + J$, where Vech C selects the unique elements of C. Assume that the model is identified (so $\Sigma(\tilde{\Theta}) \neq \Sigma(\Theta^0)$ if $\tilde{\Theta} \neq \Theta^0$).

The covariance matrix of D is consistently estimated by

$$\hat{\Sigma} = \frac{1}{N} \sum_{i=1}^{N} (D_i - \bar{D})(D_i - \bar{D})'$$

where $\bar{D}$ is the sample mean of the $D_i$. This is an average constructed over choices. Under assumptions (A-1) - (A-3), application of the Lyapunov central limit theorem implies that

$$\sqrt{N}\,(Vech\,\hat{\Sigma} - Vech\,\Sigma) \sim N(0,\Omega)$$

where $\Omega$ is a strictly positive definite symmetric matrix. (See Amemiya, 1985). The number of distinct elements of $\Sigma$ is $J(J+1)/2$.

A consistent estimator of the covariance between the (j,k) and ($\ell$,m) elements of $\Sigma$ is

$$\frac{1}{N}\sum_{j,k,\ell,m}^{N} (\hat{a}_j\, \hat{a}_k\, \hat{a}_\ell\, \hat{a}_m - \hat{\Sigma}_{jk}\, \hat{\Sigma}_{\ell m})$$

where

$$\hat{a}_j = D_{ij} - \bar{D}_j$$

and $\hat{\Sigma}_{jk}$ is the jk element of $\hat{\Sigma}$.

Under the assumed conditions, it is possible to use classical results in the minimum distance estimation literature to establish the following theorem.

*Theorem A-1. Let*

$$\tilde{\Theta} = \underset{\Theta}{argmin}\ (Vech\ \hat{\Sigma} - Vech\ \Sigma(\Theta))'\Omega^{-1}(Vech\ \hat{\Sigma} - Vech\ \Sigma(\Theta))\ .$$

*Then under assumptions (A-1) - (A-4),*

*(a)* $\quad plim\ \tilde{\Theta} = \Theta^0$

*(b)* $\quad \sqrt{N}(\tilde{\Theta} - \Theta^0) \overset{a.d.}{\sim} N(0,(C'\Omega^{-1}C)^{-1})$

*where*

*(c)* $\quad C = \dfrac{\partial Vech\ \Sigma(\Theta)}{\partial\Theta}$

*Proof.* Cragg and Donald (1995). This is an application of Ferguson (1958). ■

Cragg and Donald (1995) note that one can use only the subdiagonal elements (elements off the main diagonal) of $\Sigma$ to estimate the factor loadings. This subdiagonal estimator is asymptotically equivalent to the estimator considered in Theorem A-1. (See Proposition 2 and Corollary 2.1 of Cragg and Donald, 1995). Note further that $\Phi$ can also be estimated just as efficiently using this subdiagonal estimator. The subdiagonal estimator is also efficient and allows inference in cases where assumption A-3 fails (the unconditional variance-covariance estimator is not constant across i) whereas the estimator in Theorem A-1, while consistent, may no longer be efficient or provide valid inference in this case because the non-constancy of the variance in $\zeta_i$ makes it difficult to estimate the components of $\omega$ that correspond to the diagonal elements of $\Sigma$. Using their subdiagonal estimator, assumption A-3 can be relaxed and more general dependence of $E(\zeta_i \zeta_i' \mid Z_i)$ on i can be permitted.

Cragg and Donald (1995) present alternative assumptions on the data generating process under which Theorem A-1 is valid. Of special interest to the analysis of this paper is their case in which the $Z_i$ are a sequence of constants that satisfy Caesaro-summability conditions. This permits a much more general data generation process than the one assumed here. It is straightforward to relax the assumption A-1 that the $Z_i$ are iid. General forms of time series dependence such as m-dependence or $\phi$-mixing dependence can be introduced so long as a central limit theorem for Vech $\hat{\Sigma}$ can be produced. Open questions not investigated in this paper are (a) how much dependence among the $Z_i$ and the $\varphi_j$ can be permitted and still permit results like those in Theorems like A-1 and A-2 (below) to be proved and (b) how much dependence is induced by plausible voting models with logrolling and coalition formation and how does this dependence rule out the results reported in Theorems A-1 and A-2?

Theorem A-1 assumes that the rank of the model is known. Cragg and Donald (1996) investigate estimation of the effective dimension or rank (R) of the model i.e. the rank of $\Phi\Phi'$. R is the relevant number of relevant attributes on bills that influence votes of legislators. Define criterion T(R)

$$T(R) = \underset{\Gamma(R)}{Min} \ (Vech\ \hat{\Sigma} - Vech\ \Sigma)'\Omega^{-1}(Vech\ \hat{\Sigma} - Vech\ \Sigma)$$

over the set

$$\Gamma(R) = \{\Sigma: [\Sigma - \Psi(\psi)]V = 0 \text{ for some } V \text{ that } V'V = I_{J-R}, \text{ where } \psi \text{ is in a compact set and } \Sigma = \Sigma'\}.$$

Then Cragg and Donald (1996) establish the following proposition, where $R^0$ is the true rank.

*Theorem A-2. Under assumptions (A-1) - (A-3),*

$$NT \overset{a.d.}{\sim} \chi^2\left(\frac{(J-R_0)(J - R_0 + 1)}{2} - J\right)$$

*and if* $R \geq R^0$, *Plim* $T(R) = 0$, *and if* $R < R^0$, *Plim* $T(R) > 0$ *and* $N\,T(R) \rightarrow \infty$ *for* $R < R_0$ *and* $N\,T(R) < NT(R_0)$ *for* $R > R_0$.

*Proof.* See Cragg and Donald (1996).

They also investigate model selection procedures for determining the rank based on the criterion:

$$S(R) = \left[\frac{N}{f(N)}\right]T(R) - g(R)\ .$$

When $f(N) \equiv 1$ and $g(R) = ((J - R + 1)(J - R) - 2J)$, this criterion specializes to the Akaike criterion (AIC). The BIC (or Schwarz) criterion is produced when $f(N) \equiv \log(N)$ and $g(R) = (J - R + 1)(J - R)/2 - J$.

Assume that $\hat{\Sigma}$ satisfies the Law of the Iterated Logarithm:

(A-5) $$\lim_{N\rightarrow\infty} Sup\left(\frac{N}{k \log \log N}\right)^{1/2} \Omega^{-1/2}(Vech\ \hat{\Sigma} - Vech\ \Sigma) \leq \iota_{J(J+1)/2}$$ with probability one,

where k is some finite constant and the ordering is component-wise, and $\iota$ is a vector of

J(J + 1)/2 ones.

Cragg and Donald (1996) demonstrate that BIC is strongly consistent but AIC is not. There is a small (10% or less) asymptotic probability of overestimating the number of independent factors. They also establish the weak of consistency of a sequential pre-test estimator that selects R on the basis of the smallest value of NT(R) that does not reject the null hypothesis provided the size of the test $\alpha_N$ is adjusted with sample size so

$$(a) \quad \alpha_N \rightarrow 0 \qquad and \quad (b) \; \frac{-\ln\alpha_N}{N} \rightarrow 0 \; .$$

In a series of Monte Carlo experiments they indicate that BIC tends to underestimate the true rank $R_0$ while AIC and the pre-test estimator perform relatively well. The evidence on the performance of AIC clearly runs counter to the predictions of their theorem. This evidence leads us to use their pretest estimator.

Theorems A-1 and A-2 and the model selection analysis of Cragg and Donald demonstrate that it is possible to rigorously estimate the dimension of the model, and that consistent and asymptotically normal estimators of decision maker preferences can be achieved. Because the number of choices (N) is large relative to the number of decision makers (J), no comparable consistency results are possible for the attributes of individual choice attributes, $Z_i$, $i = 1,\ldots,N$. Since $J/N \rightarrow 0$, $Z_i$ cannot be consistently estimated.