POLANA	mpm009	JF
Journal Name	Art. No.	CE Code

doi:10.1093/pan/mpm009

Expert Opinion, Agency Characteristics, and Agency Preferences

Joshua D. Clinton
Department of Politics, Princeton University, Princeton, NJ 08540
e-mail: clinton@princeton.edu (corresponding author)

David E. Lewis
Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08540
e-mail: delewis@princeton.edu

The study of bureaucracies and their relationship to political actors is central to understanding the policy process in the United States. Studying this aspect of American politics is difficult because theories of agency behavior, effectiveness, and control often require measures of administrative agencies' policy preferences, and appropriate measures are hard to find for a broad spectrum of agencies. We propose a method for measuring agency preferences based upon an expert survey of agency preferences for 82 executive agencies in existence between 1988 and 2005. We use a multirater item response model to provide a principled structure for combining subjective ratings based on scholarly and journalistic expertise with objective data on agency characteristics. We compare the resulting agency preference estimates and standard errors to existing alternative measures, discussing both the advantages and limitations of the method.

1 Introduction

The measurement of latent concepts occupies a critical place in political science, and progress has been made possible because of the development and application of measurement models. Not only have the methods of analysis been refined but the applications of measurement models also have greatly expanded. Whereas work originally focused on voting behavior in the U.S. Congress, work now regularly analyzes legislatures, courts, and deliberative bodies across the world (for a recent summary, see Poole 2005). The statistical models commonly used to analyze roll call voting behavior have far more applications than the analysis of elite voting behavior; the underlying statistical model provides a useful framework for the measurement of latent traits.

We demonstrate how the model commonly used to analyze legislative voting behavior can be extended to provide a statistical framework for analyzing the results of an expert survey on the policy preferences of political institutions. Our substantive focus is on

Authors' note: We thank Tom Hammond, George Krause, and Joshua Tucker for helpful comments. We are grateful to Simon Jackman and Shawn Treier for generously providing their code and our survey respondents for their time and expertise.

administrative agency policy preferences in the United States, but the generality of the approach is such that it is straightforward to extend it to other contexts. For example, our analysis of expert survey responses is directly analogous, and applicable, to the widespread use of experts to assess party platforms in comparative politics.

Several clever techniques have been used to characterize agency preferences and test theories of agency influence and action for a limited set of agencies in the United States. Nixon (2004) uses the service of bureaucratic officials both in an agency and in Congress as "bridges" to estimate the ideal points of members of Congress and agency commissioners on the same space. Bertelli and Grose (2006) use public positions of Labor secretaries on votes in Congress to estimate executive preferences. Moe (1985a) and Snyder and Weingast (2000) use votes to scale commission members (in a different space from other political actors), and the partisan identification of commission members and the policy preferences of appointing presidents are also sometimes used to measure bureaucratic preferences (Cohen 1986).

These techniques are difficult to apply to more than a handful of agencies. Very few agencies have appointed officials who have also served in other branches of government (and with observable actions in those capacities), and resource constraints limit our ability to identify appointee issue positions on congressional votes. Moreover, even if we could measure the preferences of every political appointee, agency preferences are not solely, or even largely, the product of appointee preferences (Lewis forthcoming).[1] Given these difficulties, studies of multiple executive agencies often use either subjective assessments or aspects of the political environment at the time of agency creation to proxy for agency preferences. For example, some categorize agencies by whether the agency's mission is closer to the core policy commitments of one party and others assume that agencies created under Democratic presidents or unified Democratic control are more liberal than others (e.g., Gilmour and Lewis 2006a, 2006b). Subjective assessments face difficulties, however, because the opinions of a single scholar are subject to bias, misinformation, and perhaps even moral hazard concerns. Using the politics at the time of agency creation to measure policy preferences is objective, but it produces very crude characterizations.

We propose an alternative method for measuring agency preferences. We conduct an expert survey on agency preferences, and we use a multirater item response model to jointly analyze the responses and objective information about agency characteristics; we use expert opinions to refine the inferences that result from using observable agency characteristics. We use a statistical model to provide a principled structure for combining objective data on agency characteristics with subjective ratings based on scholarly and journalistic expertise that determines the relative contribution of each expert and accommodate expert differences in the threshold defining a liberal or conservative agency. We proceed deliberately. After presenting the data and the statistical measurement model, we compare our estimates to existing alternatives and discuss limitations of the method before concluding.

2 Measuring Agency Policy Preferences Using Expert Opinion

Expert surveys provide a means of assessing quantities that are not easily quantifiable. Defining the precise characteristics of a "liberal" agency in terms comparable across time

[1] Surveys of federal managers that include questions to ascertain manager partisanship or ideology provide another measure of agency preferences (e.g., Aberbach and Rockman 2000). Existing federal surveys of federal employees do not ask about personal policy views. Other surveys occur irregularly, with different questions, for different purposes, and often with samples too small to get sufficient numbers of respondents across agencies. Surveying federal managers is involved and can be expensive. Surveys of managers also do not capture aspects of agency policy preferences unrelated to employee ideology including agency mission and culture. Information from surveys of federal managers, however, could be usefully incorporated into the method we propose here.

and agencies is likely to prove difficult, if not impossible. Much like U.S. Supreme Court's Justice Potter Stewart's "I know it when I see it" definition of pornography, the same is likely true of agency preferences. Rather than attempting to define what observable traits describe liberal and conservative agencies, we instead use an induced definition; we use expert assessments of agency preferences rather than trying first to define liberal outcomes and then classifying agencies according to the posited definition.

To assess expert opinion, we identified a sample of 37 experts in bureaucratic politics among academics, journalists, and Washington think tanks in May and June of 2006.[2] We sent each expert a list of 82 departments and agencies in existence between 1988 and 2005 and asked:

> Please see below a list of United States government agencies that were in existence between 1988–2005. I am interested to know which of these agencies have policy views due to law, practice, culture, or tradition that can be characterized as liberal or conservative. Please place a check mark (√) in one of the boxes next to each agency—"slant **L**iberal, **N**either Consistently, slant **C**onservative, **D**on't **K**now."

We received 26 responses from the 37 requests (70%). To minimize contemporaneous bias in the responses, we explicitly prompt considerations of long-standing agency tendencies (policy views due to law, practice, culture, or tradition). The question is intentionally ambiguous and allows experts to use whatever definition of agency preferences they feel is most appropriate. Because assigning a precise policy position to so many agencies is difficult, we only ask whether an agency "slants" liberal or conservative and we provide a "Don't Know" option to encourage informed assessments.

3 Mean Agency Preferences

One possible estimator of agency preferences is the mean opinion. If we assign the designations of "slant liberal," "neither consistently," and "slant conservative" the values $\{-1, 0, 1\}$, respectively, we get the estimates of Fig. 1. The 82 agencies are rank ordered based on the average opinion and arranged from the most liberal to the most conservative. The number of ratings is noted to the right (left) of the graph of the most liberal (conservative) agencies (a maximum of 26 if every expert rated the agency).

The results have strong face validity. The most liberal agency, with a perfect score of -1 based on 12 responses is Action—an agency overseeing government-sponsored volunteer efforts whose functions were eventually merged into President Clinton's Corporation for National and Community Service. The most conservative agency is the Department of Defense (.962), followed closely by the Department of the Navy (.960), the Department of the Army (.920), and the Department of the Air Force (.920) based on 26, 25, 25, and 25 responses, respectively. Several agencies have scores close to 0, including the Office of Government Ethics, the Federal Emergency Management Agency, the Federal Maritime Commission, the Federal Trade Commission, and the Farm Credit Administration. These "neutral" agencies contain one agency designed to monitor ethics, two bipartisan commissions, and a disaster relief agency. (The commissions, created with fixed and staggered terms and party balancing limitations on presidential appointments, were explicitly insulated from political intervention so it is reassuring, but not surprising to find them rated close to 0.)

Despite the estimates' plausibility, estimating agency preferences using the mean rating faces several shortcomings. One difficulty concerns the resolution of expert disagreement.

[2]Specifically, we contacted 30 political scientists specializing in American or bureaucratic politics, three journalists writing on topics related to the bureaucracy, two persons working in think tanks with expertise on bureaucracy, and two persons working in Washington for nonpartisan government agencies.

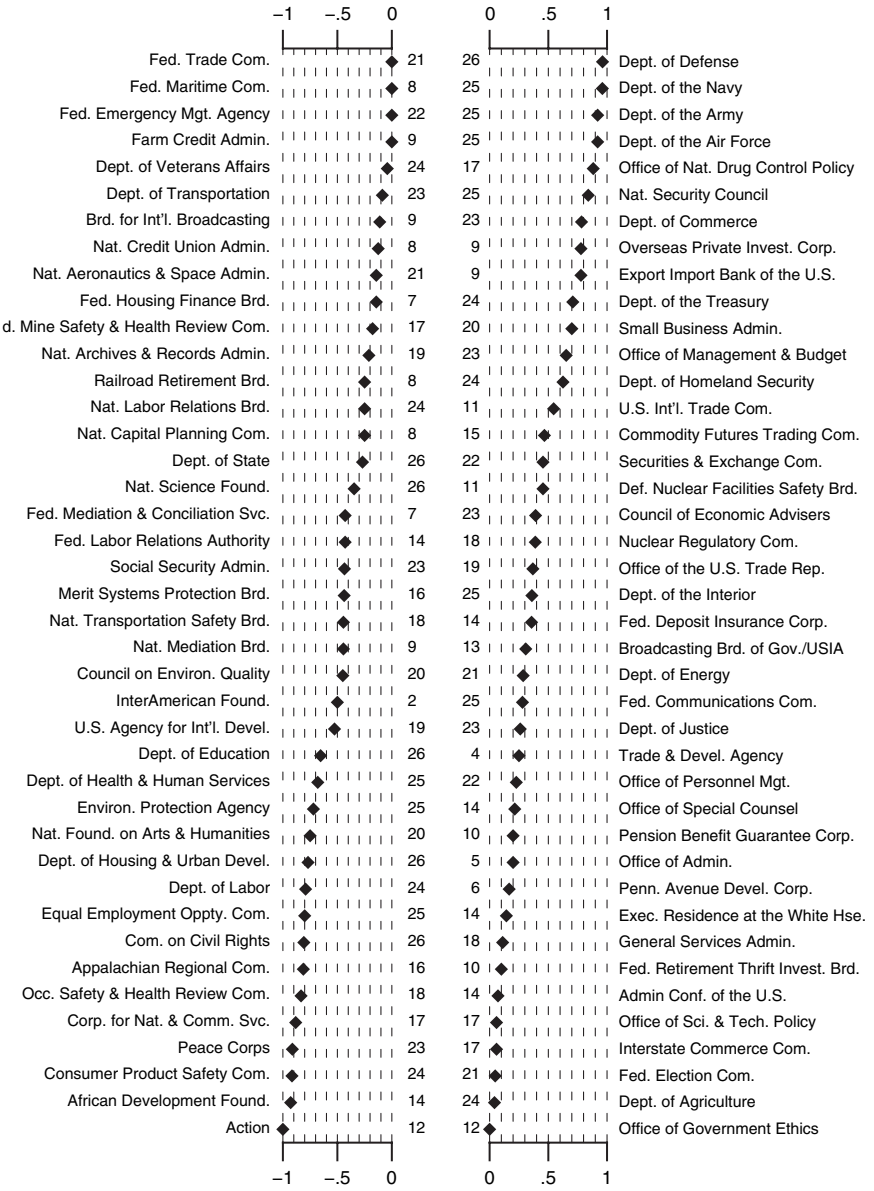

Fig. 1 Mean rating of 82 executive agencies. Number of experts rating each agency noted to the right of the more liberal agencies and to the left of the more conservative agencies.

How should we interpret the fact that seven experts rate the Department of Veterans Affairs as liberal, six rate it conservative, and 11 rate it neutral? Is the true preference of the Department of Veterans Affairs liberal (as seven indicate), conservative (as six indicate), neither (as 11 experts suggest), or the mean of −.04? Put differently, the mean opinion assumes that the determination of every expert is equally informative even though some experts may be more knowledgeable than others.[3] Also, the sample

[3]It is possible to weight the experts' determinations in principle, but it is unclear how such weights might be determined.

mean ignores additional information plausibly related to agency preferences (e.g., a lengthy literature argues that the coalition creating the agency creates a structure that embeds the coalition's political preference [McCubbins, Noll, and Weingast 1987, 1989; Moe 1989]).

We use a statistical model to combine objective data on agency characteristics with subjective ratings drawn from scholarly and journalistic expertise to produce an alternative measure of agency policy preferences. The statistical model provides a principled structure for determining the relative contribution of each expert and the objective characteristics in the estimation of agency preferences.

4 A Statistical Model of Expert Opinion

The multirater item response model was developed, and is widely used, in educational testing research because it models how test questions discriminate between individuals on the basis of a latent trait such as ability, aptitude, or intelligence (see, e.g., Johnson and Albert 1999). Political scientists use these models to characterize elite voting behavior (e.g., Voeten 2000; Martin and Quinn 2002; Clinton, Jackman, and Rivers 2004; Bafumi et al. 2005), measure countries' democratic tendencies (Jackman and Treier 2006), respondents' political ideology (Hillygus and Treier 2006), countries' political-economic risk (Quinn 2004), the notability of legislative enactments (Clinton and Lapinski 2006), and to analyze graduate admissions decisions (Jackman 2004). The model describes how observed responses—expert opinions in our application—correspond to the latent dimension (i.e., agency policy preferences).

We assume all agencies $i \in 1 \ldots m$ possess an unobservable policy preference (ideal point) x_i. We further assume surveyed experts $j \in 1 \ldots n$ agree on what constitutes liberal and conservative in the following sense—if the true agency policy preferences were observed, all experts would agree on the relative preference ranking. Experts may differ in what constitutes a liberal or conservative agency, but no expert would think that agency i is more liberal than agency $i + 1$ if $x_i > x_{(i+1)}$.

Each expert rates whether agency i is liberal (1), conservative (3), or neither (2). Let $\mathbf{Y}$ be the M $\times$ N ratings matrix with element y_{ij} being agency i's rating by expert j. In our application, M = 82 and N = 26. To specify how agencies' policy preferences $\mathbf{x}$ map into expert ratings, we assume that the latent variable governing the classification of agencies into the experts' scale is: $\mu_{ij} = x_i\beta_j + \varepsilon_{ij}$.[4] This permits for the possibility that experts—perhaps because of differing levels of knowledge—imperfectly observe true agency preferences. β_j describes how agency preferences $\mathbf{x}$ map into the latent variable $\boldsymbol{\mu}$ for expert j; if $\beta_j = 0$, the ratings of expert j are unrelated to agency preferences $\mathbf{x}$ and are uninformative for distinguishing agency preferences.

In determining whether an agency is liberal, conservative, or neither, experts implicitly define a threshold for each response category such that $-\infty < \tau_{j\mathrm{L}} \leq \tau_{j\mathrm{N}} \leq \tau_{j\mathrm{C}} \equiv \infty$. Expert j thinks agency i "slants liberal" ($y_{ij} = 1$) if $\mu_{ij} < \tau_{j\mathrm{L}}$, if $\mu_{ij} \in [\tau_{j\mathrm{L}}, \tau_{j\mathrm{N}}]$ the agency is "consistently neither" ($y_{ij} = 2$), and the agency "slants conservative" ($y_{ij} = 3$) if $\mu_{ij} > \tau_{j\mathrm{N}}$. Experts are assumed to share a common conception of the underlying policy dimension, but what constitutes a "liberal" agency may differ by expert. This accommodates the

[4]Because not every rater evaluates every agency, we assume that the decision to not offer a rating is independent of both the preferences of the agency and the expert's quality. That is, a high degree of missingness does not indicate that the expert's ratings are less informative for agencies that are rated relative to experts who rate every agency.

possibility that although they may agree on what defines the ideological endpoints, a conservative ideologue may think almost every agency is liberal (i.e., τ_{jL} is large) and a liberal ideologue may think only a handful of agencies are sufficiently liberal (i.e., τ_{jL} is small).

The model therefore allows for two sources of expert disagreement. First, perhaps because of the difficulty of assessing agency preferences, experts only imperfectly perceive true agency preferences and they may differ in how their classifications reflect true agency preferences ($\boldsymbol{\beta}$). Second, experts' thresholds ($\boldsymbol{\tau}$) for defining liberal and conservative agencies may differ.

To turn this model of expert decision making into a statistical model requires an expression for the probability that each event occurs—that is, an assumption about the error $\boldsymbol{\varepsilon}$. If we permit the precision of the experts' judgments to vary with variance δ_j^2, but assume that $\mathrm{E}[\boldsymbol{\varepsilon}] = 0$, the cumulative distribution function (CDF) of ε_{ij} is $\mathrm{F}(0, \delta_j^2)$. Mathematically:

$$\Pr(y_{ij} = 1) = \Pr(\mu_{ij} < \tau_{jL}) = \int_{-\infty}^{\tau_{jL}} \frac{1}{\delta_j} f\left(\frac{\mu_{ij} - x_i\beta_j}{\delta_j}\right)\partial x = F\left(\left(\frac{\tau_{jL}}{\delta_j}\right) - x_i\left(\frac{\beta_j}{\delta_j}\right)\right),$$

$$\begin{aligned}\Pr(y_{ij} = 2) = \Pr(\tau_L < \mu_{ij} < \tau_N) &= \int_{\tau_{jL}}^{\tau_{jN}} \frac{1}{\delta_j} f\left(\frac{\mu_{ij} - x_i\beta_j}{\delta_j}\right)\partial x \\ &= F\left(\left(\frac{\tau_{jN}}{\delta_j}\right) - x_i\left(\frac{\beta_j}{\delta_j}\right)\right) - F\left(\left(\frac{\tau_{jL}}{\delta_j}\right) - x_i\left(\frac{\beta_j}{\delta_j}\right)\right),\end{aligned}$$

$$\Pr(y_{ij} = 3) = \Pr(\tau_N < \mu_{ij}) = \int_{\tau_{jN}}^{\infty} \frac{1}{\delta_j} f\left(\frac{\mu_{ij} - x_i\beta_j}{\delta_j}\right)\partial x = 1 - F\left(\left(\frac{\tau_{jN}}{\delta_j}\right) - x_i\left(\frac{\beta_j}{\delta_j}\right)\right),$$

where $f(\bullet)$ and $F(\bullet)$ denote the logistic density (PDF) and CDF, respectively.[5]

The likelihood of the model is the product of the probabilities across both agencies and experts $\mathrm{L}(\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{x}, \boldsymbol{\delta}) = \prod_{i=1}^{m}\prod_{j=1}^{n}(1/\delta_j) f\left(\frac{\mu_{ij} - x_i\beta_j}{\delta_j}\right) I(\tau_{j,k-1} \leq \mu_{ij} < \tau_{jk})$, where $I(\bullet)$ is an indicator function, $\boldsymbol{\tau}_j$ is the set of thresholds ($k = \{\mathrm{L}, \mathrm{N}\}$) used by expert j, y_{ij} is agency i's rating by expert j, and only $\mathbf{y}$ is observed. A set of priors completes the specifications (see also Johnson and Albert 1999; Quinn 2004; Jackman 2004; Jackman and Treier 2006; Hillygus and Treier 2006), and we estimate the model using Markov chain Monte Carlo methods.[6]

To accommodate existing measures of agency preferences and possibly refine our expert-based estimates, we also incorporate information plausibly correlated with agency

[5]The first expression is derived as follows: $\Pr(y_{ij} = 1) = \Pr(\mu_{ij} < \tau_{jL}) = \Pr(x_i\beta_j + \varepsilon_{jL} < \tau_{jL}) = \Pr(\varepsilon_{jL} < \tau_{jL} - x_i\beta_j)$. Because $\mu_{ij} = x_i\beta_j + \varepsilon_{ij}$ implies $\varepsilon_{ij} = \mu_{ij} - x_i\beta_j$, $\Pr(\varepsilon_{ij} < \tau_{jL} - x_i\beta_j) = \int_{-\infty}^{\tau_{jL}} 1/\delta_j f((1/\delta_j)(\mu_{jL} - x_i\beta_j))\partial x = F((1/\delta_j)(\tau_{jL} - x_i\beta_j))$ and $f(\bullet)$ and $F(\bullet)$ denote the PDF and CDF of $\boldsymbol{\varepsilon}$, respectively. Clearly, only the ratios ($\boldsymbol{\tau}_j/\delta_j$) and ($\beta_j/\delta_j$) are identified. Assuming that experts rate with equal precision ($\delta_j = 1$ for all raters) simplifies the expression accordingly.

[6]We assume $\boldsymbol{\beta} \sim \mathrm{N}(0, 100)$, $\mathbf{x} \sim \mathrm{N}(0, 1)$, the first threshold $\tau_{jL} \sim \mathrm{N}(0, 1)$ for all raters j, $\tau_{jk} = \sum_{k=2}^{3}\theta_{jk}$ for $k = 2$ (N) and $k = 3$ (C) and the difference between successive cutpoints $\tau_{j,k} - \tau_{j,k-1} = \theta_{jk} > 0$. The prior for $\boldsymbol{\theta}$ is exponential with $\lambda = 2$. To identify the model we postprocess the output iteration-by-iteration constraining the estimates $\mathbf{x} \sim \mathrm{N}(0, 1)$ and the ideal point for Action to be liberal (for other applications in political science, see Jackman 2004; Jackman and Treier 2006; Hillygus and Treier 2006).

Table 1 Descriptive statistics of agency characteristics

Variable	*Sample mean*	*Values*
Social regulation	.232	{0, 1}
Economic regulation	.183	{0, 1}
Defense agency	.073	{0, 1}
International agency	.171	{0, 1}
Created under unified Democratic control	.500	{0, 1}
Created under unified Republican control	.049	{0, 1}
Created under unified Federalist control	.061	{0, 1}
Created under Democratic president	.573	{0, 1}

preferences. The benefits of incorporating additional information depends on how correlated the covariates are to agency preferences. Table 1 summarizes several plausible correlates used in the literature.[7]

There are two ways of including information about the politics at the time of agency creation and the tasks of each agency into the statistical model (see Johnson and Albert 1999, chapter 5.3). Because we are ambivalent about the proper weight to attribute to agency characteristics, we estimate both. First, covariates can be recoded and treated as "experts." For example, the politics at the time of agency creation can be treated as an expert whose determination is conservative if the agency was created under unified Republican control, liberal if created under unified Democratic control, and neither if created during divided government. This assumes that agency characteristics are as informative as a single expert, and it requires assuming a mapping from agency characteristics into the expert opinion scale.[8]

A second approach uses the covariates to specify the relationship of the latent trait μ_{ij}. Instead of assuming $\mu_{ij} = x_i\beta_j + \varepsilon_{ij}$, assume $\mu_{ij} = x_i\beta_j + \mathbf{W}\gamma + \eta_i + \varepsilon_{ij}$, where $\boldsymbol{\gamma}$ is a d $\times$ 1 matrix of regression coefficients, $\mathbf{W}$ denotes the T $\times$ d matrix of covariates related to agency preferences, and $\boldsymbol{\eta}$ is an independent and identically distributed N(0,1) error term (Jackman 2004). Modeling latent agency preferences as a function of covariate information effectively privileges the covariate information. Experts whose classifications are similar to the specified regression equation will be treated as being more informative about the underlying agency preferences. This assumption is strong given the crudeness of available covariates.

5 Comparing the Agency Preference Estimates

The statistical measurement model yields estimates of agency ideal points ($\mathbf{x}$), the relationship between agency characteristics and agency preferences ($\boldsymbol{\gamma}$), and expert

[7]Regulatory data come from Dudley, Susan, and Melinda Warren. 2003. *Regulatory spending soars: An analysis of the U.S. budget for fiscal years 2003 and 2004*. Mercatus Center (George Mason University) and Weidenbaum Center (Washington University, St Louis). This regular publication of the Weidenbaum Center tracks regulatory spending over time on an agency-by-agency basis and is a useful tool for identifying which federal agencies engage in regulatory behavior (http://wc.wustl.edu). Information on whether or not agencies are engaged in defense or international issues is drawn from the federal budget. Source: Budget Analysis Branch, Office of Management and Budget. 2004. *Public Budget Database User's Guide: Budget of the United States Government, Fiscal Year 2005* (http://www.whitehouse.gov/omb/budget/fy2004/pdf/db_guide.pdf, February 2, 2004).

[8]In addition to the politics at the time of creation, we treat four other agency characteristics as raters. Specifically, we include a rating of "slants liberal" if the agency is tasked with social regulation, economic regulation, or international affairs and a rating of "slants conservative" for agencies dealing with defense issues.

characteristics ($\boldsymbol{\beta}, \boldsymbol{\tau}$). Only the first two are of substantive interest, but the estimated expert characteristics highlight the importance of accounting for expert differences when measuring agency preferences.

Figure 2 plots the agency preference estimates and standard errors from a multirater model using only expert opinion in the same order as in Figure 1.[9] The estimates use the same data as the mean ratings in Fig. 1, but the estimates in Fig. 2 permit the relationship between the experts' ratings to affect the weight given to each expert's determinations. The estimates are quite similar because the expert sample we utilize is of high quality—almost every expert is informative about agency preferences in the sense that their determinations discriminate between agencies in broadly consistent ways.

Despite the strong relationship, modest differences emerge as a consequence of permitting experts to vary in their determinations. For example, the departments of Transportation and Veterans Affairs are estimated to be more conservative than other departments relative to the mean rating by 4 and 13 spots, respectively, and the Office of Government Ethics and the Office of Science and Technology Policy move about 10 spots in a liberal direction. (The reason for the differences arises because of differences between the experts' discriminations—the six experts indicating that Veterans Affairs is more conservative are estimated to have more discrimination than the seven experts rating the agency as liberal.)[10]

Although it is tempting to conclude that the Commission on Civil Rights is more liberal than the Consumer Product Safety Commission based on the resulting point estimates, each agency's score is estimated with error, and the liberalism-conservatism of proximate agencies are often statistically indistinguishable. So although we are confident that the Department of Defense is more conservative than the Department of Labor and the Department of Energy according to our estimates, e.g., the confidence intervals for the Departments of Defense, Army, Navy, and Air Force overlap considerably, and we are much less certain about which one is the most conservative.

We can also compare the multirater item response model estimates to existing measures. Liberal agencies according to agency function are indicated with abbreviations following the agency's name in Fig. 2 (i.e., agencies tasked with social regulation are followed by Soc). Agencies created by a unified Democratic government are indicated with an open symbol, agencies created by a unified Republican government are indicated with a closed symbol, and agencies created during a period of divided government are indicated with an asterisk. Nineteen agencies are classified as engaging in social regulation (Gilmour and Lewis 2006a, 2006b), and 42 are created by a unified Democratic government.[11]

Figure 2 reveals that agency characteristics are only faintly related to the expert opinion. Agencies created under unified Democratic (Republican) control—plotted with open

[9]The model was estimated in *WinBUGS* 1.4.1 using slice updating. Two chains of 500,000 iterations each thinning by 100 were estimated, and the first 250,000 iterations were used as "burn-in" to find the posterior. Estimates are available online and in Appendix.

[10]The item discrimination estimates for the seven experts indicating "slants liberal" are $\beta_3 = 2.39$, $\beta_8 = 2.80$, $\beta_9 = 1.07$, $\beta_{10} = 1.18$, $\beta_{21} = 2.39$, $\beta_{22} = 3.51$, and $\beta_{29} = 2.58$. The item discrimination parameters for the six experts indicating "slants conservative" are $\beta_2 = 1.63$, $\beta_5 = 2.34$, $\beta_7 = 2.78$, $\beta_{12} = 4.18$, $\beta_{17} = 4.49$, and $\beta_{23} = 3.85$.

[11]Agencies coded as liberal by one method are not necessarily liberal according to another. For example, using the politics at the time of agency creation is the only method that suggests that the Council of Economic Advisors, the Federal Emergency Management Agency, and the General Services Administration are among the most liberal agencies. If we look at the list of agencies engaged in social regulation, this is the only list that includes the departments of Agriculture, Homeland Security, and Justice.

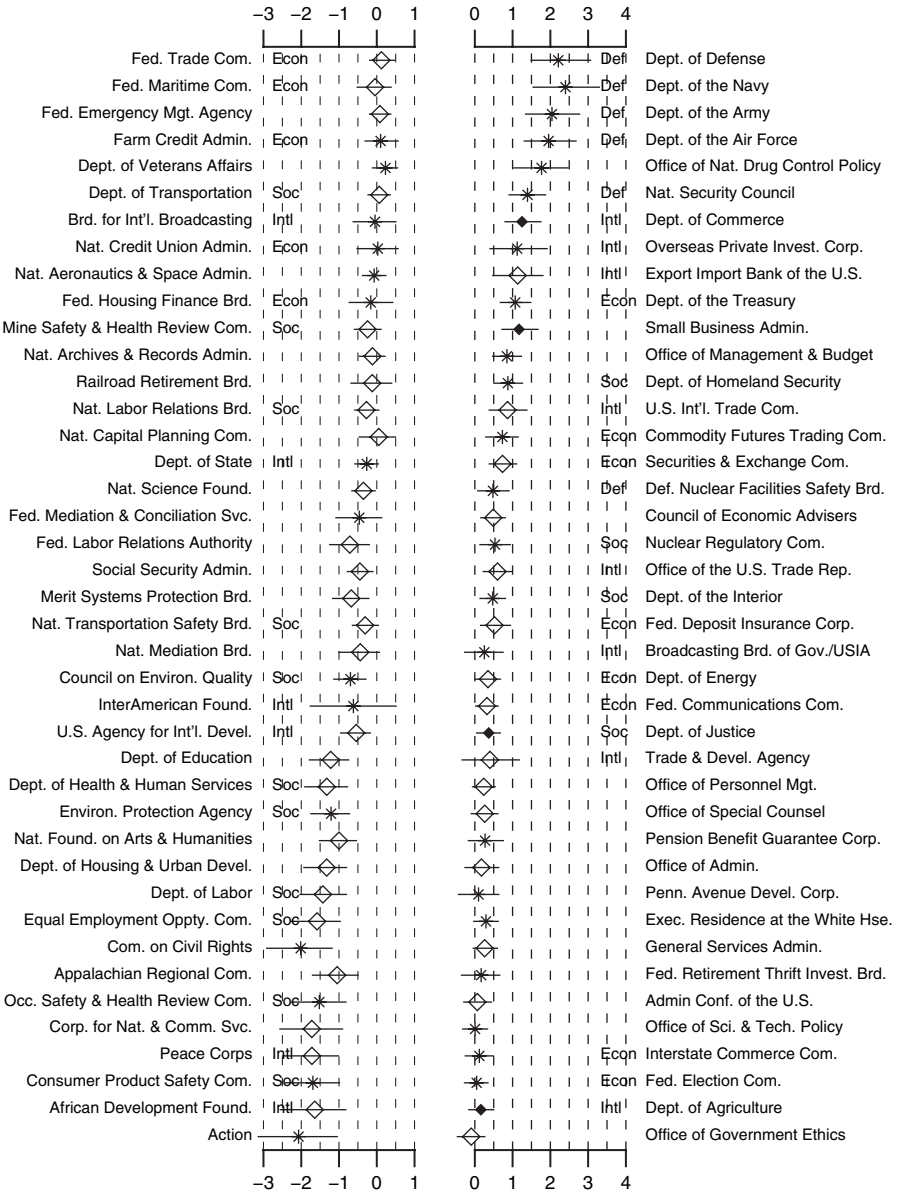

Fig. 2 Estimates from multirater model. Lines denote the 95% region of highest posterior density. Points for agencies created under unified Democratic (Republican) control are open (solid). Asterisks denote agencies created under divided government (or unified Federalist government). Agencies dealing with economic regulation, social regulation, international affairs, and defense issues are so labeled.

(solid) points—are not systematically more liberal or conservative than agencies created under divided government. Agencies dealing with defense issues are among the most
260 conservative according to the experts, but the relationship with agencies dealing with economic regulation, social regulation, and international affairs is unclear. Only our measure discriminates between agencies with any degree of precision. Using the politics at the

time of creation assumes all 42 agencies created under unified Democratic control or all 19 engaging in social regulation are liberal. It is likely that some agencies are more liberal than others within such a large group, and inspecting the lists reveals cause for concern. (The Department of Homeland Security, the National Archives and Records Administration, and Office of Personnel Management are not agencies associated with liberal policy preferences.) The most liberal agencies according to our experts would not be coded as liberal using the other measures of agency liberalism.[12] For example, looking at the politics at the time an agency is created would miss the Civil Rights Commission, the Consumer Product Safety Commission, and the Environmental Protection Agency. Using social regulatory functions as a proxy for liberalism would omit the Corporation for National and Community Service, the Peace Corps, and Department of Housing and Urban Development.

5.1 *Accounting for Expert Differences*

To highlight the importance of accounting for expert differences rather than assuming, like the mean rating does, that every expert is equally informative, we examine characteristic curves for four experts. (Treating the expert determinations differently explains the slight differences between the mean estimates and the multirater estimates.) The x axis in Fig. 3 plots the latent agency trait $\boldsymbol{\mu}$ for a hypothetical agency, and the y axis plots the probability that the expert rates an agency with the given policy preference as "conservative" (solid thin line), "liberal" (solid thick line), or "neutral" (dashed line).[13] To examine the experts' behavior over regions of the space in which the data have support, we determine the range of the hypothetical values of μ using the 95% region of highest posterior density of the agency preference estimates ($\mathbf{x}$).

Expert 22 represents a near ideal classification. Agencies with low latent ideologies (i.e., $\mu < -1$) are most likely to be classified as liberal and those with high latent ideologies (i.e., $\mu > -1$) are most likely to be designated as conservative. For an agency with $\mu = 0$, the expert is most likely to indicate the agency is "neither," but as μ increases (decreases) from 0, the probability of the expert denoting the agency as conservative (liberal) quickly increases.

For expert 6, unless $\mu > 1$, the expert is most likely to indicate that the agency is "neither consistently." In fact, the expert is never most likely to designate the agency as liberal regardless of the level of the latent trait. Expert 12 is always willing to designate an agency as either liberal or conservative; for no values of μ is the expert most likely to indicate that the agency is "neither consistently." Expert 19 is uninformative—for any realization of μ over which the data have nonnegligible support, expert 19 is most likely to rate the agency conservative.

These three expert characteristic curves provide the intuition for how the model accommodates expert differences. Because there is no realization for which expert 19 is not more likely to indicate that the agency is conservative, expert 19 is uninformative for discriminating between agency preferences. Whereas expert 6 may designate an agency with μ close to 0 as conservative, μ must be relatively large for expert 12 to do so. As a result,

[12] Measuring liberal agency preferences using only social regulatory agencies created under unified Democratic control does not greatly improve the alternative to the measure we propose. Such a list would miss the six most liberal agencies according to our measure, and name only three of the top 20 liberal agencies (Equal Employment Opportunity Commission [#7], Department of Labor [#9], and Department of Health and Human Services [#11]).

[13] Because only the ratios (τ_j/δ_j) and (β_j/δ_j) are identified, to evaluate the rater characteristic curves we assume $\delta_j = 1$ for all raters.

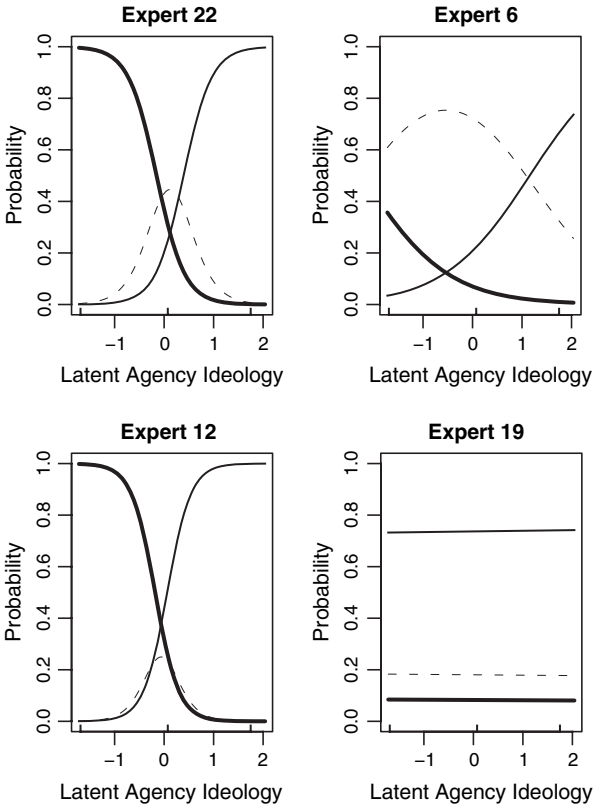

Fig. 3 Selected characteristic curves for experts. The thickest line denotes the probability of the expert designating an agency with the given policy preference as "slants liberal," the dashed line denotes the probability of the expert indicating "neither," and the slender solid line denotes the probability of a "slants conservative" rating.

conservative designations by expert 12 indicate a more extreme agency than the desig-
305 nations of expert 6. Finally, because expert 22's determinations partition the space of μ into three regions—a region where each response is most likely—expert 22's designations provide some sense of where μ must lie.

5.2 *Agency Preferences and Agency Characteristics*

In addition to the expert opinions we collect, we also know of agency characteristics plau-
310 sibly correlated with agency preferences. The relationship between characteristics and preferences in Fig. 2 suggests an imperfect relationship, but useful information may nonetheless be present in the agency characteristics that we can use to further refine our estimates.

Regressing the mean expert opinion and the estimates of the prior section on the agency characteristics of Table 1 yields the coefficient estimates in Table 2. The estimates are
315 largely in the expected direction: agencies created under unified Democratic (Republican) control are less (more) conservative than those created under divided control, defense agencies are more conservative, and those tasked with social regulation are estimated to be more liberal. Given the tremendous variation in the agencies created under unified

Table 2 Agency characteristics and expert opinions

	Mean rating	*Multirater estimates*
Constant	−.076 (.105)	−.165 (.180)
Unified Democratic creation	−.215 (.235)	−.376 (.404)
Unified Republican creation	.417 (.258)	.780* (.443)
Unified Federalist creation	.201 (.241)	.585 (.414)
Democratic president at creation	.078 (.236)	.236 (.406)
Social regulation	−.174 (.126)	−.341 (.217)
Economic regulation	.307* (.135)	.498* (.233)
Defense	.812* (.251)	1.60* (.431)
International	.076 (.144)	.086 (.247)
N	82	82
R^2	.355	.402
Residual standard error	.455	.783

Note. Standard errors in parentheses. Asterisks denote coefficients significant at .1 or less (two-tailed). The scales are not directly comparable.

Democratic control and among agencies engaged in different types of regulatory behavior, however, the relationship is imprecise.

There are two consequences of the modest relationships in Table 2. First, the imprecise relationship between agency characteristics and expert opinion and the face validity of the resulting agency preference measure suggests that expert opinion reflects more than just agency characteristics. Second, the unimpressive relationship between expert opinions and agency characteristics suggests that, at least in our application, a measure that integrates covariates and expert opinions will either fail to appreciatively change the estimates or else yield less credible estimates.[14]

Figure 4 plots four sets of agency preference estimates against one another. In addition to the mean rating and the multirater item response estimates plotted in Figs. 1 and 2, respectively, Figure 4 also includes two sets of estimates utilizing agency characteristics in three different ways.

Several conclusions emerge. First, the multirater item response model that ignores covariate information is highly correlated with the mean expert rating (correlation of .971), with the evident nonlinearity reflecting the assumption of a logistic error structure. Second, the decision how to model covariate information is largely inconsequential. Given the marginal relationships evident in Table 2, including covariate information in a way that does not privilege agency characteristics fails to change the estimates. Treating agency characteristics as an expert (Covariates as Experts) yield estimates that are nearly perfectly correlated with the estimates that analyze expert opinion alone (multirater model).

Only if we assume that true agency preferences are a linear function of agency characteristics—which is a very strong assumption given the available characteristics—do slight differences emerge (Latent Trait Reg.). Including agency characteristics in this way pushes agencies in the direction the covariates would predict. Nonregulatory agencies like

[14]Treating covariates as experts will change the estimates only a fraction because the lack of fit evident in Table 2 means that the contribution of the covariate "experts" to the agency estimate will effectively be down-weighted. The only difference in agency preference estimates likely to result from including agency characteristics is the model that assumes latent preferences are determined by agency characteristics.

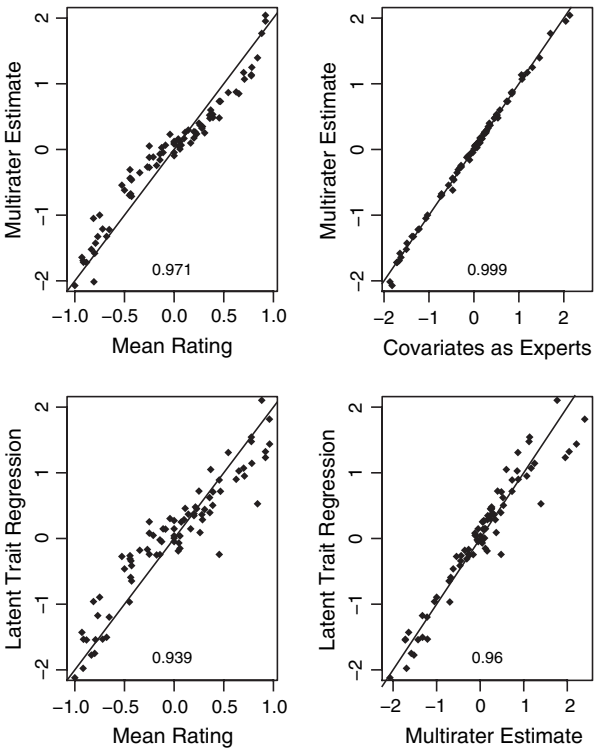

Fig. 4 Comparison of the estimates. The plotted estimates include the mean expert rating, the multirater model ignoring covariate information, the multirater model in which agency characteristics are recoded and treated as an expert and the multirater model in which the latent trait is modeled with agency characteristics. The number in each plot is the correlation coefficient for the multirater estimates ignoring agency characteristics.

Action, the African Development Foundation, and the Appalachian Regional Commission are estimated to be more moderate, whereas the departments of Agriculture and Justice, which do regulate, become significantly more liberal (by 18 and 17 spots, respectively). Agencies involved in foreign affairs and defense such as the Board for International Broadcasting, the Office of the United States Trade Representative, and the Overseas Private Investment Corporation are also estimated to be significantly more conservative (11, 6, 6 spots). Most of the movement in the agencies' relative rankings happens in the middle of the distribution; agencies in the tails are in the tails because they have the most expert agreement (so the covariates have the least influence).

5.3 *Limitations of the Measure*

Despite the face validity of our measure—especially relative to alternative measures—the method we use to measure agency preferences is imperfect. First, the reliability and usefulness of agency preference estimates depend critically on the quality of the raters and their ratings. Although we select experts well situated in the field, it is inevitable that some experts will be unable to rate all 82 agencies. The average expert evaluated two-thirds of the 82 agencies, and several experts mentioned the difficulty of the task (particularly rating agencies in the Executive Office of the President). Moreover, a few experts

appeared to have reasoned that because the current president was a Republican, almost all the agencies had to be either neutral or conservative. Not surprisingly, these were some of the least discriminating experts according to the multirater model. These two problems illustrate the benefits of employing a statistical measurement model that accounts for differences in the discrimination and liberalism-conservatism thresholds of experts.

This benefit, however, can also be a curse depending on the collective expertise of the experts and the validity of the assumption that the raters share a common conception of the underlying latent trait. Because outlying opinions are effectively treated as reflecting error rather than the latent agency preference, the model can accommodate a small number of deviant raters, but too much disagreement will result in uninformative estimates absent the incorporation of additional information. This does not appear to affect our application, but it may affect others depending on the difficulty of the task relative to the expertise and similarity of the raters.

Another potential shortcoming of the specific estimates we produce is that they are time invariant and they reflect long-term tendencies. This reflects an intentional choice because measures of agency preferences based on appointee preferences are sometimes inappropriate for testing political science theories. For example, Moe (1985b) argues that presidents politicize and centralize as a means of enhancing presidential control of the bureaucracy. If so, presidents are more likely to politicize or take policy formulation responsibilities away from agencies that do not share their preferences (Rudalevige 2002). Measures of agency preferences using appointee preferences cannot be used to test this proposition because a president's decision to politicize has to do with an agency's preferences prior to appointment.

The estimates are also not comparable to the preference estimates of other political actors (e.g., NOMINATE scores). Asking members of Congress to evaluate whether agencies tend to be more liberal, more conservative, or neither consistently compared to themselves and explicitly defining the member's threshold parameters in terms of the member's personal preference, however, may offer a potential solution.

6 Conclusion

A persistent and difficult problem in the study of American politics involves the measurement of administrative agency policy preferences. This is unfortunate because a measure of agency preferences would be useful for building and testing interinstitutional theories of American politics. Theories of bureaucratic control and delegation rely on claims about the degree of congruence between agency preferences and those of the three branches of government (Weingast and Moran 1983; Epstein and O'Halloran 1999; Huber and Shipan 2002). Our understanding of the administrative presidency suggests that presidents choose different control strategies for agencies that share or do not share their views about policy (Rudalevige 2002). Theories of lawmaking and budgetary politics suggest that gains from trade occur only when presidents and legislators can credibly commit to not changing agency policy preferences in the future (McCubbins, Noll, and Weingast 1989; McCarty 2004). Models of judicial oversight of administrative agencies depend fundamentally on claims about agency preferences relative to those of the courts (Canes-Wrone 2003).

We make two contributions. First, we demonstrate how the statistical model commonly used to assess legislative roll call voting can be extended for use in analyzing expert opinion about agency policy preferences. Analyzing an expert survey we administer to 37 experts using a multirater item-response model yields estimates of administrative agency liberalism-conservatism that significantly improve upon existing measures of

agency preferences (e.g., individual subjective judgments, measures of the politics at the time agencies were created, or details about agency functions). Although our substantive focus is U.S. administrative agencies, the approach is broadly applicable to other contexts and can be readily extended.

Second, we generate estimates of stable agency liberalism or conservatism useful for testing theories of political control and presidential administrative strategy. Given the importance of measurement to facilitating scientific advance—the notable impact of Poole and Rosenthal (1997) and Martin and Quinn (2002) is partially due to the estimates they provide to scholars—we hope that the provision of estimates that, although imperfect, nonetheless improves upon available measures of agency preferences provides a valuable contribution.

Appendix: Multirater item response estimates of Agencies Policy Preferences, 1988–2005 (most liberal to most conservative)

Agency	*Mean*	*SD*	*2.5% CI*	*97.5% CI*	*Politics at creation*	*Social Reg.*
Action	−2.07	0.55	−3.14	−1.05		
Commission on Civil Rights	−2.01	0.45	−2.91	−1.18		
Corporation for National and Community Service	−1.72	0.43	−2.56	−0.91	X	
Peace Corps	−1.72	0.38	−2.49	−1.02	X	
Consumer Product Safety Commission	−1.69	0.38	−2.42	−0.99		X
African Development Foundation	−1.64	0.45	−2.55	−0.82	X	
Equal Employment Opportunity Commission	−1.58	0.34	−2.28	−0.97	X	X
Occupational Safety and Health Review Commission	−1.52	0.37	−2.25	−0.82		X
Department of Labor	−1.43	0.32	−2.03	−0.81	X	X
Department of Housing and Urban Development	−1.33	0.30	−1.93	−0.80	X	
Department of Health and Human Services	−1.32	0.30	−1.91	−0.78	X	X
Department of Education	−1.22	0.27	−1.78	−0.75	X	
Environmental Protection Agency	−1.21	0.27	−1.74	−0.72		X
Appalachian Regional Commission	−1.05	0.31	−1.70	−0.50	X	
National Foundation on the Arts and the Humanities	−1.00	0.25	−1.52	−0.54	X	
Federal Labor Relations Authority	−0.71	0.27	−1.24	−0.20	X	
Council on Environmental Quality	−0.70	0.22	−1.14	−0.29		X
Merit Systems Protection Board	−0.68	0.25	−1.17	−0.21	X	
Inter-American Foundation	−0.62	0.58	−1.76	0.51		
U.S. Agency for International Development	−0.54	0.20	−0.96	−0.17	X	
Federal Mediation and Conciliation Service	−0.46	0.31	−1.08	0.13		
Social Security Administration	−0.45	0.17	−0.78	−0.10	X	
National Mediation Board	−0.44	0.28	−1.01	0.07	X	
National Science Foundation	−0.35	0.16	−0.66	−0.05	X	
National Transportation Safety Board	−0.31	0.18	−0.65	0.04	X	X
National Labor Relations Board	−0.27	0.16	−0.58	0.05	X	X
Department of State	−0.27	0.16	−0.58	0.04		
Federal Mine Safety and Health Review Commission	−0.24	0.18	−0.59	0.11	X	X

Continued

Appendix (continued)

Agency	*Mean*	*SD*	*2.5% CI*	*97.5% CI*	*Politics at creation*	*Social Reg.*
Federal Housing Finance Board	−0.16	0.29	−0.73	0.42		
Railroad Retirement Board	−0.12	0.28	−0.68	0.39	X	
National Archives and Records Administration	−0.12	0.17	−0.46	0.22	X	
Office of Government Ethics	−0.10	0.19	−0.46	0.26	X	
National Aeronautics and Space Administration	−0.07	0.16	−0.38	0.24		
Board for International Broadcasting	−0.05	0.29	−0.62	0.51		
Federal Maritime Commission	−0.05	0.23	−0.52	0.38	X	
Office of Science and Technology Policy	0.01	0.17	−0.32	0.34		
National Credit Union Administration	0.03	0.27	−0.52	0.56		
Federal Election Commission	0.05	0.16	−0.27	0.35		
National Capital Planning Commission	0.05	0.24	−0.46	0.49	X	
Department of Transportation	0.07	0.15	−0.23	0.36	X	X
Administrative Conference of the United States	0.07	0.19	−0.29	0.44	X	
Federal Emergency Management Agency	0.08	0.15	−0.19	0.37	X	
Pennsylvania Avenue Development Corporation	0.10	0.27	−0.43	0.63		
Farm Credit Administration	0.10	0.23	−0.31	0.57		
Interstate Commerce Commission	0.12	0.19	−0.25	0.49		X
Federal Trade Commission	0.12	0.17	−0.19	0.47	X	
Department of Agriculture	0.16	0.17	−0.16	0.50		X
Federal Retirement Thrift Investment Board	0.17	0.26	−0.35	0.66		
Office of Administration	0.17	0.23	−0.26	0.63	X	
Department of Veterans Affairs	0.23	0.17	−0.12	0.55		
Office of Personnel Management	0.24	0.15	−0.06	0.54	X	
Broadcasting Board of Governors/USIA	0.25	0.26	−0.27	0.75		
General Services Administration	0.26	0.16	−0.04	0.60	X	
Office of Special Counsel	0.26	0.18	−0.09	0.61	X	
Pension Benefit Guarantee Corporation	0.27	0.23	−0.17	0.75		
Executive Residence at the White House	0.30	0.17	−0.03	0.62		
Federal Communications Commission	0.32	0.15	0.02	0.62	X	
Department of Energy	0.35	0.17	0.01	0.68	X	X
Department of Justice	0.37	0.16	0.05	0.67		X
Trade and Development Agency	0.40	0.38	−0.34	1.18	X	
Department of the Interior	0.47	0.17	0.14	0.81		X
Defense Nuclear Facilities Safety Board	0.48	0.21	0.08	0.91		
Council of Economic Advisers	0.49	0.16	0.16	0.81	X	
Federal Deposit Insurance Corporation	0.52	0.20	0.15	0.94	X	
Nuclear Regulatory Commission	0.53	0.21	0.14	0.94		X
Office of the U.S. Trade Representative	0.60	0.19	0.22	0.97	X	
Commodity Futures Trading Commission	0.73	0.22	0.29	1.15		
Securities and Exchange Commission	0.73	0.18	0.38	1.11	X	
Office of Management and Budget	0.85	0.19	0.47	1.23		
U.S. International Trade Commission	0.87	0.25	0.38	1.38	X	
Department of Homeland Security	0.88	0.19	0.51	1.26		X

Continued

Appendix (continued)

Agency	*Mean*	*SD*	*2.5% CI*	*97.5% CI*	*Politics at creation*	*Social Reg.*
Department of the Treasury	1.07	0.21	0.68	1.48		X
Overseas Private Investment Corporation	1.12	0.39	0.40	1.91		
Export Import Bank of the United States	1.14	0.35	0.48	1.80	X	
Small Business Administration	1.17	0.24	0.72	1.67		
Department of Commerce	1.25	0.24	0.80	1.75		
National Security Council	1.40	0.25	0.91	1.87		
Office of National Drug Control Policy	1.77	0.37	1.04	2.47		
Department of the Air Force	1.95	0.35	1.31	2.68		
Department of the Army	2.04	0.38	1.34	2.77		
Department of Defense	2.21	0.41	1.49	3.06		
Department of the Navy	2.40	0.46	1.54	3.30		

Note. Agencies are ordering from most liberal to most conservative according to (posterior) mean estimate; 95% of highest posterior density interval also reported. Estimates reflect the ratings of 26 experts. The last two columns indicate whether the agency would be coded as "liberal" by other common means of measuring agency preferences.

References

Aberbach, Joel D., and Bert A. Rockman. 2000. *In the web of politics: Three decades of the U.S. Federal Executive.* Washington, DC: Brookings.

Bafumi, Joseph, Andrew Gelman, David Park, and Noah Kaplan. 2005. Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis* 13(2):171–87.

Bertelli, Anthony M., and Christian R. Grose. 2006. Secretaries of Pork? Executive ideology, multiple bureaucratic principals, and distributive public policy. Unpublished manuscript, University of Georgia.

Canes-Wrone, Brandice. 2003. Bureaucratic decisions and the composition of the lower courts. *American Journal of Political Science* 47(2):205–14.

Cohen, Jeffrey E. 1986. The dynamics of the "Revolving Door" on the FCC. *American Journal of Political Science* 30(4):689–708.

Clinton, Joshua D., and John S. Lapinski. 2006. Measuring legislative accomplishment, 1877–1994. *American Journal of Political Science* 50(1):232–49.

Clinton, Joshua D., Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of legislative behavior: A unified approach. *American Political Science Review* 98(2):355–70.

Epstein, David, and Sharyn O'Halloran. 1999. *Delegating powers.* New York, NY: Cambridge University Press.

Gilmour, John B., and David E. Lewis. 2006a. Political appointees and the competence of Federal Program Management. *American Politics Research* 34(1):22–50.

———. 2006b. Assessing performance assessment for budgeting: The influence of politics, performance, and program size in FY2005. *Journal of Public Administration Research and Theory* 16(2):169–86.

Hillygus, D. Sunshine, and Shawn Treier. 2006. The contours of policy attitudes in the mass public. Unpublished manuscript, Harvard University.

Huber, John D., and Charles R. Shipan. 2002. *Deliberate discretion?* New York: Cambridge University Press.

Jackman, Simon. 2004. "What do we learn from Graduate Admissions Committees? A multiple-rater, latent variable model with incomplete discrete and continuous indicators. *Political Analysis* 12(4):400–24.

Jackman, Simon, and Shawn Treier. 2006. Democracy as a latent variable. Unpublished manuscript, Stanford University.

Johnson, Valen E., and James H. Albert. 1999. *Ordinal data modeling.* New York: Springer.

Lewis, David E. Forthcoming. *Politicizing administration.* Princeton, NJ: Princeton University Press.

Martin, Andrew D., and Kevin M. Quinn. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10(2):134–53.

McCarty, Nolan M. 2004. The appointments dilemma. *American Journal of Political Science* 48(3):413–28.

McCubbins, Mathew D., Roger Noll, and Barry Weingast. 1987. Administrative procedures as instruments of political control. *Journal of Law, Economics, and Organization* 3:243–77.

———. 1989. Structure and process, politics and policy: Administrative arrangements and the political control of agencies. *Virginia Law Review* 75(2):431−82.

Moe, Terry M. 1985a. Control and feedback in economic regulation: The case of the NLRB. *American Political Science Review* 79(4):1094−116.

———. 1985b. The politicized presidency. In *The new direction in American politics*, eds. J. E. Chubb and P. E. Peterson, 235–71. Washington, DC: Brookings Institution Press.

———. 1989. The politics of bureaucratic structure. In *Can the government govern?* eds. J. E. Chubb and P. E. Peterson, 267–329. Washington, DC: Brookings.

Nixon, David C. 2004. Separation of powers and appointee ideology. *Journal of Law, Economics, and Organization* 20(2):438−57.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A political-economic history of roll call voting*. New York: Oxford University Press.

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

Rudalevige, Andrew. 2002. *Managing the president's program: Centralization and Legislative Policy Formulation, 1949−1996.* Princeton, NJ: Princeton University Press.

Quinn, Kevin M. 2004. Bayesian factor analysis for mixed ordinal and continuous responses. *Political Analysis* 12(4):338–53.

Snyder, Susan K., and Barry R. Weingast. 2000. The American system of shared powers: the President, Congress, and the NLRB. *Journal of Law, Economics, and Organization* 16(2):269−305.

Voeten, Erik. 2000. Clashes in the assembly. *International Organization* 54(2):185−217.

Weingast, Barry W., and Mark J. Moran. 1983. Bureaucratic discretion or congressional control? Regulatory policymaking by the Federal Trade Commission. *Journal of Political Economy* 91(5):765−800.